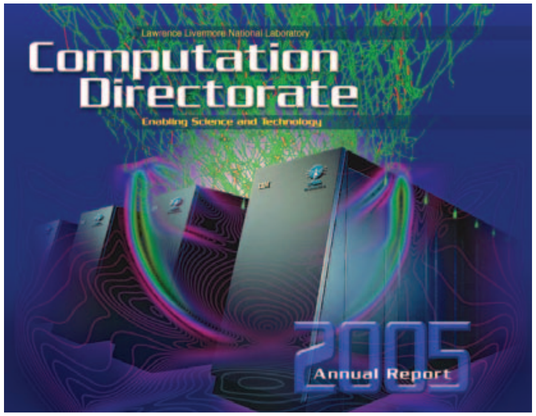
Lawrence Livermore National Laboratory

Computation Directorate

Enabling Science and Technology

2005 Annual Report



About the Cover

The BlueGene/L (BG/L) supercomputer shown on the cover is an icon for the nation's Advanced Simulation and Computing (ASC) program. Developed in partnership with IBM, BG/L was delivered and installed at Lawrence Livermore over the course of 2005. The full machine uses 131,072 commercial microprocessors to achieve unprecedented computing power. In 2005, it achieved a Linpack result of 280.6 trillion floating-point operations per second (TF), and a sustained performance of more than 100 TF on an important calculation for the ASC program. It ranks first on the Top500 Supercomputers list. Shown in the background are results from various simulations run on the high-performance computers maintained by Livermore's Computation Directorate. (Cover design by Daniel S. Moore.)

Enabling Science and Technology

LLNL Computation Directorate Annual Report 2005

UCRL-TR-219935

Online: http://www.llnl.gov/comp/publications/annual_report/2005/Comp_Annual_Report.pdf

LLNL Associate Director for Computation: .... **Dona L. Crawford**

Deputy Associate Directors: .... **Theodore C. Michels (Principal Deputy) Steven F. Ashby, and Michel G. McCoy**

Scientific Editors: .... **Steve R. Anderson, Mary E. Zosel, and Mark C. Miller**

Art Director: .... **Daniel S. Moore**

Production Editor: .... **Carolin Middleton**

Editors: .... **Arnold N. Gatilao, James R. Kohl, Deanna K. Midtaune, and Maurina S. Sherman**

Compositor: .... **J. Louisa Cardoza**

Photographer: .... **Joseph Martinez**

Proofreader: .... **Gabriele Rennie**

Authors: .... **See contact information after each article**

Print Production: .... **Charlie M. Arteago, Jr. and Haig's Quality Printing**

About the Computation Directorate

Our Vision

Computation aspires to be the preeminent high-performance computing and computer science organization in order to enable scientific discovery and Laboratory missions.

Our Mission

Computation assures Laboratory mission and program goals are attained by delivering outstanding computer science expertise, world-class high-performance computing capabilities, and creative technology and software solutions.

Computation Directorate

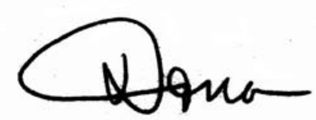
NNSA National Nuclear Security Administration

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

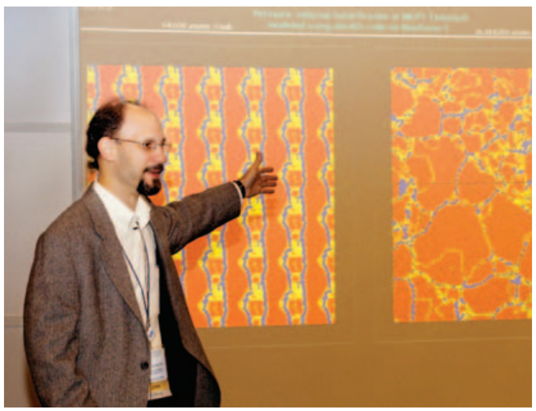
Contents

Awards and Recognitions

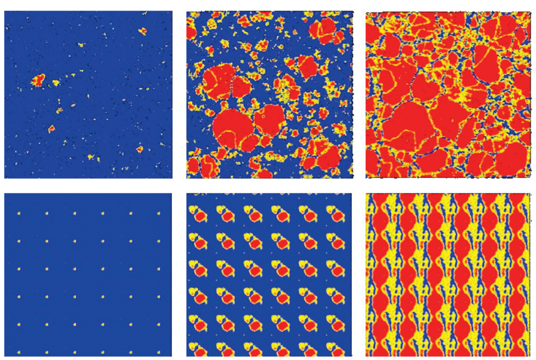
In addition to the directorate-wide accomplishments presented in this annual report, I want to recognize the awards received by Computation personnel and projects during the past year. On behalf of Lawrence Livermore National Laboratory (LLNL) and the directorate, I congratulate the award winners. They set the bar a little higher for all of us.



Dona Crawford
Associate Director, Computation



Physicist Fred Streitz from LLNL's Physics and Advanced Technologies Directorate discusses the results from his team's Gordon Bell Prize–winning calculation.



These results are part of the 3D simulation on BG/L that won the 2005 Gordon Bell Prize. The calculations modeled the pressure-induced solidification of molten tantalum using (top row) 16 million atoms and (bottom row) 64,000 atoms.

2005 Gordon Bell Prize

A team of scientists led by LLNL physicist Fred Streitz received the 2005 Gordon Bell Prize for pioneering materials science simulations conducted on the world's fastest supercomputer—the IBM BlueGene/L (BG/L) at Livermore. Other team members included James Glosli, Mehul Patel, Bor Chan, Robert Yates, and Bronis de Supinski of LLNL, and James Sexton and John Gunnels of IBM. Running a three-dimensional (3D) molecular-dynamics code on BG/L, the team investigated solidification in tantalum and uranium at extreme temperatures and pressures. Simulations ranged from 64,000 atoms to 524 million atoms.

Two other LLNL teams were finalists in the competition. François Gygi in Computation led a team that attacked a difficult molecular-dynamics problem using the full quantum-mechanical formulation and achieved an unprecedented 64 trillion floating-point operations per second (TF) on BG/L. Gygi's team included Erik Draeger, Bronis de Supinski, and Robert Yates of LLNL; Franz Franchetti of Carnegie Mellon University; Stefan Kral, Juergen Lorenz, and Christoph Ueberhuber of Vienna University of Technology in Austria; and John Gunnels and James Sexton of IBM's Thomas Watson Research Center.

The third team, led by Andy Yoo also in Computation, achieved the first known search of a massive graph with more than 30 billion connections. This team included Edmond Chow and Keith Henderson of LLNL, William McClendon and Bruce Hendrikson of Sandia National Laboratories, and Umit Catalyurek of Ohio State University.

At the Top of the List

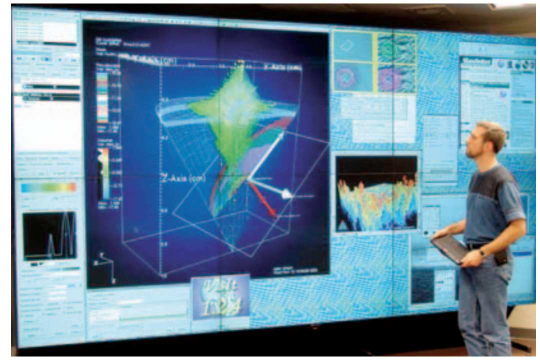
Two LLNL supercomputers rank in the top 5 of the Top500 Supercomputer Sites, announced at the 2005 Supercomputing Conference (SC|05). Number 1 on the list is BG/L. This IBM supercomputer reached 280.6 TF on the Linpack benchmark, the industry standard to measure computing speed. In a demonstration of its capability, BG/L ran a record-setting materials science application, which was an important calculation for the Department of Energy's (DOE's) Advanced Simulation and Computing (ASC) program, at 101.5 TF sustained over 7 hours on the machine's 131,072 processors. BG/L remains the only system ever to exceed the 100-TF mark. At number 3 on the Top500 list is LLNL's Purple supercomputer, which was built by IBM and supports the ASC program. In Purple's demonstration, it reached a peak performance of 63.4 TF.

Meeting the Challenge

The BG/L supercomputer also dominated the first annual High-Performance Computing Challenge Competition at SC|05, sweeping all four Class 1 best-performance awards and tying (with a team from Cray) for the Class 2 "elegant" implementation award. The best-performance benchmarks involve seven tests on a supercomputing run, including Global HPL, which measures the floating-point rate of execution for solving a linear system of equations; EP STREAM, a simple synthetic benchmark program that measures sustainable memory bandwidth in gigabytes per second and the corresponding computation rate for simple vector kernel; Global RandomAccess, which measures the rate of integer random updates of memory; and Global FFT, the floating-point rate of execution for double precision complex one-dimensional discrete Fourier transform.

Picture Perfect with VisIt

VisIt, a flexible visualization and analysis tool developed by Computation scientists, received a 2005 R&D 100 Award in *R&D Magazine*'s annual competition for the top 100 industrial, high-technology inventions. This software, which was released for unlimited noncommercial use, has a scalable architecture. It can thus process the enormous data sets generated by the world's most powerful supercomputers and visualize the results in seconds. It also has a plug-in architecture, allowing scientists to easily add new capabilities. The project team, which received funding from DOE's ASC program, was led by LLNL computer scientist Eric Brugger.



VisIt, the flexible visualization and analysis tool developed by Computation scientists, received a 2005 R&D 100 Award. Here, Sean Ahern uses VisIt to view results from a simulated experiment for the National Ignition Facility.

Three Patents for the Sapphire Project

Sapphire, the LLNL data-mining project led by computational scientist Chandrika Kamath, received three patents for new technologies during the past year: using histograms to introduce randomization in the generation of ensembles of decision trees (U.S. Patent 6,859,804); creating ensembles of decision trees through sampling (U.S. Patent 6,938,049); and parallel object-oriented, denoising patents using wavelet multiresolution analysis (U.S. Patent 6,879,729). All three of these patents contribute to efforts by the Center for Applied Scientific Computing to develop scalable algorithms for exploring large, complex, scientific data sets. By applying and extending ideas from data mining, image and video processing, statistics, and pattern recognition, the Sapphire project is developing a new generation of computational tools and techniques to improve the way in which scientists extract useful information from data.

VisIt development team: (left to right, standing) Mark Miller, Sean Ahern, Eric Brugger, Kathleen Bonnell, and Brad Whitlock; (sitting) Hank Childs and Jeremy Meredith; and (kneeling) Linnea Cook.

Serving the Mathematics Community

David Keyes, a computational mathematician and acting director of LLNL's Institute for Scientific Computing Research, was elected as vice president at large for the Society of Industrial and Applied Mathematics (SIAM) for 2006. Keyes is also the Fu Foundation Professor of Applied Mathematics at Columbia University. The mission of SIAM is to build cooperation between mathematics and other scientific and technological communities through publications, research, and conferences.

David Keyes

1.00 — Strategic Solutions to Meet Program Needs

Dona Crawford,
Associate Director
Computation Directorate

Computing at LLNL in 2005

Even before Lawrence Livermore National Laboratory (LLNL) opened in September 1952, cofounders E. O. Lawrence and Edward Teller recognized the need for a computer and placed an order for one of the first production Univacs. Equipped with 5,600 vacuum tubes, the Univac had impressive calculational power for its time, although much less than that contained in today's five-dollar calculator. Computing machines quickly demonstrated to Livermore staff the ability not only to perform complicated calculations but also to simulate physical processes. Today, computational excellence is a major reason for our continued success as one of the world's preeminent scientific research institutions.

Supercomputer simulations allow us to understand the intricacies of physical phenomena that occur at exceedingly short time scales and at extreme pressures and temperatures. In some research areas, an experiment cannot be conducted because it would be too costly or too difficult to perform or diagnose, or it would involve toxic materials. The computer, in effect, serves as a virtual laboratory in which calculational results—the simulated experiment—can point to scientific "truth" as effectively as other means. Physical experiments and computational simulations often complement one another, in that physical experiments validate computational simulations, and simulations help researchers design laboratory experiments.

For the past decade, the driving force behind increasingly realistic simulations has been stockpile stewardship, which is the Department of Energy's (DOE's) National Nuclear Security Administration (NNSA) program to ensure the safety and reliability of the nation's nuclear weapons stockpile. A major element of stockpile stewardship is the Advanced Simulation and Computing (ASC) program, which had an initial 10-year goal to obtain computers that could run simulations at 100 trillion floating-point operations per second (TF). Working on stockpile stewardship pushes us to develop new computer codes, better simulation methods, and improved visualization technologies. It also sets the requirements for new generations of machines. The resulting capabilities, in turn, are applicable to other disciplines. Researchers in chemistry, materials science, climate change, and physics have been among the leaders in using supercomputers to advance their fields.

Computing capability at LLNL took a giant leap forward in 2005 with the installation and dedication of two world-class computers, one for stockpile science and the other for weapons performance. Together, they total almost 500 TF in peak performance, all at the disposal of the three DOE/NNSA weapons laboratories.

LLNL supercomputers listed on the Top500 Supercomputer Sites (November 2005).

Rank	Speed	System Name	Description
1	280.6 TF	BlueGene/L	BlueGene/L is a 65,536-node supercomputer built in partnership with IBM.
3	63.4 TF	ASC Purple	Five-sixths of the final Purple system, also built by IBM. When combined with pURPURA (below), the final Purple will have a peak performance of 100 TF.
11	19.9 TF	M&IC Thunder	Thunder is a highly integrated, well-balanced capability computing resource with 1,024 nodes. It is used for unclassified Multiprogrammatic & Institutional Computing (M&IC) simulations.
19	13.0 TF	ASC pURPURA	One-sixth of ASC Purple delivered to LLNL for early production use.
45	7.6 TF	MCR Linux Cluster	MCR, a tightly coupled Linux cluster for the M&IC community, has 1,152 nodes and 4 gigabytes of memory.
47	7.3 TF	ASC White	ASC White was the third step in the six-stage DOE plan to achieve a 100-TF system.

The Purple machine is a robust, massively parallel system built by IBM that will have a peak performance of 100 TF. Purple represents the culmination of more than a decade of careful planning and work across three laboratories and the ASC Program Office. It symbolizes the ability of ASC to deliver on a grand vision. We expect full-scale production runs on Purple to begin in spring 2006.

BlueGene/L (BG/L) is an icon for the future and the new ASC, moving toward better predictive models. Developed in partnership with IBM, BG/L was delivered and installed over the course of the year: a quarter of the machine at first, then half, and finally the complete machine in the fall of 2005. At each phase of the installation, the machine ranked first on the list of Top500 Supercomputer Sites, which is based on the best performance running the Linpack benchmark. The full machine achieved a Linpack result of 280.6 TF, making it currently the fastest supercomputer in the world. We call its design "disruptive" because it uses 131,072 commercial microprocessors to achieve unprecedented computing power. Clever packaging allows the machine to use less floor space and consume less energy than previous computers.

With the BG/L and Purple machines, the ASC program has clearly exceeded its 10-year goal of 100 TF. In Section 2, we report on the directorate's successes with these new machines. We provide details on our move into the Terascale Simulation Facility, which houses the ASC capability computers and the personnel who support these machines. We also describe our efforts to integrate BG/L and Purple into LLNL's research community and describe the other services provided by our directorate to researchers at LLNL and other DOE facilities in the area of high-performance computing.

The Terascale Simulation Facility at LLNL houses the ASC capability machines, Purple and BG/L.

In addition to our many services for computational science, the directorate supports LLNL's Institutional Information Technology Program. In this role, we manage much of the information technology (IT) infrastructure at LLNL. Our responsibilities include overseeing the Laboratory's networks and providing security solutions both for LLNL and other DOE sites. In 2005, we continued to increase the software and services available to managed desktops, and we improved configuration management at Livermore by deploying Active Directory and other automated tools. We also completed a major upgrade of the Secure Communication and Teleconference (SCAT) system. This LLNL-developed system has been used throughout the DOE complex for more than 20 years and is a vital component in the nation's emergency response system. With the upgrade, SCAT is now a Web-based application that can span multiple networks. In addition, work continued in LLNL's Cyber Security Program to better combat the threat of unwanted network intrusions. Section 3 highlights our activities in this area.

In 2005, BG/L achieved a Linpack result of 280.6 TF, making it the fastest supercomputer in the world according to the Top500 Supercomputers list.

Computation employees provide computer science expertise to all directorates at the Laboratory. This work encompasses a wide range of applications, such as analyzing pathogens in support of national biodefense, developing large-scale modeling and simulation software in support of the nation's Stockpile Stewardship Program, and designing an integrated control system with the related software for the National Ignition Facility. LLNL computer specialists working in multidisciplinary teams have achieved some impressive results using an array of skills from desktop support expertise to complex applications development to advanced research. Section 4 describes a subset of the applications development work within Computation.

Beyond our efforts in support of the Laboratory's mission, the Computation Directorate has established a world-class research organization. Similar to the applications development area, this organization cuts a wide swath across Laboratory programs. In this year's report, Section 5 highlights our research efforts in molecular dynamics, climate modeling, information extraction from large data sets, and scientific visualization.

In the future, we expect to continue our leadership role in high-performance computing. We envision an evolution from our current focus on computing and information management to predictive simulation and knowledge discovery. This effort will include new approaches to analyzing massive amounts of heterogeneous data and exploiting discrete simulations. We anticipate that research efforts in real-time image analysis, natural language processing, and multi-modal information analysis will be accelerated.

The Computation Directorate delivers high-performance computing and creative technology and software solutions in support of national security and all the other important Laboratory missions. As we have done in the past, we used outstanding computer science, high-level IT, and desktop support expertise along with the world-class high-performance computing capabilities to help LLNL achieve its mission and program goals in 2005. Success in our support function depends on a continued use of interdisciplinary teams. No one person or single type of expertise can guarantee success. LLNL's programs depend on Computation to provide the necessary hardware infrastructure and accompanying levels of software and computer science expertise to help them reach their scientific milestones.

In closing, I invite you to review this report of the directorate's projects, visit the Web sites listed here for more information, and contact the investigators whose work interests you. Thank you for your interest in the Computation Directorate at Lawrence Livermore.

Dona Crawford
Associate Director, Computation

RETURN TO TABLE OF CONTENTS

SECTION 2

High-Performance Computing

2.00 — A Decade to an Apogee

A decade ago, when the Department of Energy (DOE) began its Advanced Strategic Computing Initiative (ASCI), the program's ambitious goal was to develop the capability to run integrated weapons performance calculations on computers that could process up to 100 trillion floating-point operations per second (TF). Today, that goal is being achieved with the successful deployment of the Purple machine for use by the National Nuclear Security Administration's (NNSA's) three weapons laboratories—Lawrence Livermore (LLNL), Los Alamos, and Sandia national laboratories. Purple is housed in LLNL's Terascale Simulation Facility (TSF), which was completed in 2004 and now houses the capability machines and related personnel for the Advanced Simulation and Computing (ASC) program. Even more impressive is that computational scientists have developed the million-plus-line codes needed to exploit the power of the ASC machines in complex simulations of nuclear weapons performance.

With the completion of the ASCI vision, DOE and NNSA have transitioned efforts to the ASC program. This program's strategy is to use the newly developed computing technologies in support of stockpile stewardship and to advance our understanding of fundamental science. The basic science research is supported by ASC Advanced Architecture platforms such as BlueGene/L (BG/L), which is the first in a planned series of breakthrough computing technologies.

Purple and BG/L are the signature achievements for Livermore Computing (LC) during 2005. We also continue to support high-performance computing beyond the nuclear weapons complex. LLNL's Multiprogrammatic and Institutional Computing (M&IC) platforms, such as the Linux-based Thunder and MCR cluster, provide scientists with the capacity computing they need while balancing the performance cost.

Our 2005 accomplishments demonstrate the successful implementation of LC's three-curve strategy, which was first conceived in 2000. At that time, we identified two factors as critical to the success of ASCI as it evolved into the full-fledged ASC program.

First, the emphasis in computing would broaden. Ten years ago, we focused our efforts primarily on developing a scalable computing platform such as Purple for high-fidelity calculations of weapons performance. However, we realized that as computing technologies matured, researchers would want support for capacity machines, such as the M&IC system, as well as breakthrough capability for fundamental science applications.

Second, to achieve this diversification within the budget constraints of a mature ASC program, we would need to optimize the cost and cost performance for a spectrum of computers so that all of them meet the requirements of their designated workload. We could not rely on a magic-happens-here operational plan. Instead, this combined focus required thorough long-term planning, and early returns on our implementation of the three-curve strategy have been impressive.

In this section, we report on these accomplishments: The ASC capability computers are now located in the TSF. ASC Purple has been commissioned and is being used to run demanding three-dimensional weapons codes. BG/L won the Gordon Bell Prize for its sustained performance of more than 100 TF on a significant molecular-dynamics calculation.

Clearly, LC's strategy has paid off. Recent discussions with DOE's Office of Science managers indicate that they recognize the three-curve strategy as a rational way to manage balanced computing thrusts in the future. In particular, success requires that mission drivers are clearly defined; funding is provided for the computers, infrastructure, and applications needed for scientific research; and verification and validation efforts demonstrate that measurable progress is being made toward well-defined goals. With the budget pressures facing the ASC program, multilaboratory investments with proven industrial partners (such as IBM) are even more critical than they have been in the past.

The ASC and LC leadership at LLNL are working together with IBM, the Office of Science, and NNSA in a cost-sharing partnership to design follow-on systems to BG/L, culminating in a BlueGene/Q system in 2011. The next goal is to develop the capacity to process calculations at a speed of several thousand trillion floating-point operations per second. This ambitious venture will not only help the nation keep its edge against foreign competition. It also will allow DOE to attract and retain talented computer scientists and engineers who can provide the nation with cost-effective computing solutions to meet important national needs, from stockpile stewardship to nuclear power fuel-cycle research and development—to the as-yet-unknown problem in need of a world-class solution.

RETURN TO TABLE OF CONTENTS

2.01 — Deploying the Terascale Simulation Facility

Overview

In 2005, Computation moved into the TSF, achieving an important milestone for the ASC program. Planning for a computing facility to house multiple world-class supercomputers began in 1996. After extensive design and approval processes, LLNL hosted a ground-breaking ceremony for the $91-million facility in April 2002.

The builder, M. A. Mortenson Company, completed construction of the TSF well ahead of schedule and on budget. At a ribbon-cutting ceremony on July 8, 2004, then Secretary of Energy Spencer Abraham officially opened the facility. In October, LC began relocating computer systems from the Building 113 complex and installing new systems in the TSF. Staff moved into the TSF between January and April 2005. All operational functions of LC, including Computer Operations and the hotline, were moved to the TSF during this period.

Progress in 2005

Although installation and relocation activities started in 2004, most of the moves occurred during the first half of 2005. The top table at right provides the final schedule of major computer systems and equipment moved. Relocating infrastructure components such as network-attached storage (NAS) services is a complex operation and often results in significant downtimes. Careful planning enabled LC to take only two major downtimes, one each for the high-performance unclassified and classified networks. The unclassified high-performance network was down for about one day, and most critical services were restored in less than 24 hours. The classified high-performance network was also down for about one day.

Staff moves were organized by group or work unit. Two units were moved each week starting in late January until the moves were completed. In addition, support facilities and equipment such as conference rooms, mail facilities, copiers, and printers were moved at this time.

The magnitude and scope of effort involved in relocating a world-class supercomputer center are difficult to convey. Computer floor layouts must be planned, computer system electrical installation completed, network connectivity established, downtimes negotiated, office occupancy and furniture layouts determined, and labor scheduled to execute the actual moves. The bottom table at right provides statistics to quantify this effort.

Significance

The TSF has positioned LLNL as a premier site for capability computing for NNSA. The facility has 48,000 square feet of free-span computer floor and is designed to provide 15 megawatts of power for computing equipment (12 megawatts are currently available) plus the power necessary for cooling and office facilities (25 megawatts total). The TSF is a modern, flexible facility that can house the most demanding (power- and cooling-wise) computing equipment. In addition, it is designed to handle future high-density systems that may require liquid cooling.

Moving to the TSF has improved interactions within the LC organization because staff members once located in four separate buildings are now housed in one facility. Most importantly, moving was accomplished with minimal disruption to LC customers, thereby helping them to continue their quest for scientific validation and discovery through computer simulation throughout the move period.

System move schedule for the TSF.

System	Install or relocation dates	New (N) or relocated (R)
TSF, East Room		
Classified Archive Storage	October 5, 2004–May 23, 2005	N/R
BG/L*	November 8, 2004–February 4, 2005	N
iLX	December 7–9, 2004	R
GPS	December 13–17, 2004	R
Unclassified centerwide NAS	February 4–5, 2005	R
TSF, West Room		
Cub	March 15–17, 2005	R
SC	March 22–23, 2005	R
Tempest	April 5–8, 2005	R
Ace	April 12–15, 2005	R
Purple**	April 14–May 5, 2005	N
Lilac	April 18–May 16, 2005	R
Classified centerwide NAS	April 29–30, 2005	R

* Period includes only the first two deliveries totaling 32 racks, or half of the final system.
** Period includes delivery and initial installation of only the first 256 nodes.

TSF deployment statistics.

Total equipment racks moved or installed	559
Total nodes moved (excluding BG/L)	2,670
Miles of computer system electrical wire installed	53.3
Network connections established	~3,000
Total staff (and offices) moved	216

Contact Information

Douglas R. East, dre@llnl.gov

2.02 — Purple Integration

Overview

ASC Purple is the fulfillment of an early program goal for a 100-TF supercomputer. This computing capability is equivalent to the combined capacity of 25,000 high-end personal computers. Unlike a collection of separate personal computers, ASC Purple can concentrate its entire computing resources on a single problem. It will be used to run the most demanding three-dimensional (3D) weapons performance codes. In addition, as a major experimental user facility, it will allow the NNSA weapons laboratories to prioritize the largest problems and schedule them on the machine.

Purple can operate in this manner because of the ASC program's foresight. The program is striving to procure sufficient capacity computing, consisting of midsized clusters, throughout the NNSA complex to accommodate the myriad smaller jobs that must be run. If this approach is successful, the large machine can be available for its primary task: executing large-memory applications that require tremendous internal bandwidth and possibly weeks to complete. Purple will be operated as an über facility; that is, only the most demanding and important calculations for the NNSA weapons laboratories will be run on this machine.

ASC Purple installed in LLNL's Terascale Simulation Facility.

Progress in 2005

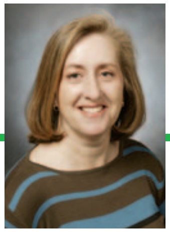
Livermore Computing began taking delivery of the ASC Purple system in late April 2005. The first phase consisted of 252 nodes, or one-sixth of the final Purple configuration, plus one-sixth of the global parallel file system. The Phase 1 system, nicknamed pURPURA (Latin for purple), was available for use in the classified environment in late June. This early delivery allowed code groups to begin preparing their applications for the Purple environment.

Subsequent deliveries of Purple hardware arrived in June (252 nodes), July (252 nodes), and September (776 nodes) for a total of 1,280 nodes. These nodes were installed and tested in the unclassified environment during 2005, and they will be merged with pURPURA in the classified environment in February 2006. Limited production availability for the full Purple system (1,532 nodes) is expected in the spring of 2006. An additional benefit of having such a large unclassified system available, if only for a short time, has been its availability for science runs.

ASC Purple is an enormous system by most metrics. The federation input/output interconnect requires more than 17,000 cables totaling over 140 miles. Purple provides 2 million gigabytes of storage from more than 11,000 Serial ATA and Fibre Channel disks, and it has 131 node racks, 90 disk racks, and 48 switch racks. The final configuration includes 1,528 8-way Squadron IH compute/input/output nodes and two 32-way Squadron H nodes partitioned into four 16-way login nodes. Purple uses 4.8 megawatts of power in running the system and an additional 3 megawatts for cooling.

Significance

The ASC program's IBM Purple supercomputing system is the realization of a bold vision formed a decade earlier—to provide the computing capability needed to routinely run complex, 3D integrated weapons performance applications in days or weeks. Using the Purple machine, scientists can validate these simulations against experimental data because in many cases the numerical errors caused by limited mesh resolution can be avoided. In the past, such errors have interfered with results produced in stockpile simulations. Because of the system's speed and memory, the accuracy of the simulations it runs is tied more directly to the fidelity of the physics and material models inherent in the codes instead of to the resolution used. The computing power of ASC Purple will enable simulations that more accurately predict stockpile performance and will better ensure the safety and reliability of the nation's stockpile.



Contact Information

Pam Hamilton, pgh@llnl.gov
Mark Seager, seager@llnl.gov

Project Web site:
http://www.llnl.gov/asci/platforms/purple/

RETURN TO TABLE OF CONTENTS

2.03 — BlueGene/L Integration

Overview

BlueGene/L is a 65,536-node supercomputer that was fully integrated at LLNL over a year's time. BG/L uses a revolutionary densely packed architecture that includes diskless nodes with 700-megahertz processors interconnected via high-speed networks. It delivers unparalleled ratios of cost, floor space, and power per trillion floating-point operations per second. BG/L differs from past vendor-proprietary supercomputers in that the vendor did not supply a job scheduler, resource manager, debugger, or parallel file system. The BG/L integration team at LLNL was responsible for integrating these components as well as a federated network infrastructure during 2004 and 2005.

Progress in 2005

The past year was one of tremendous achievement for integrators and users of BG/L. In February, the second quarter of the system arrived and was successfully integrated by the end of March. In early science runs on the first half of the machine (from April through July), scientists achieved first-time results on important molecular-dynamics calculations (Sections 2.04, 5.01, and 5.02). These science runs were critical to the machine's successful integration because the simulations helped the integration team uncover many issues that IBM, LLNL, and other collaborators addressed throughout the year.

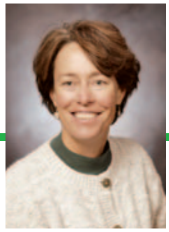
Early integration work concentrated primarily on system software, namely the control system, resource manager, job scheduler, debugger, and parallel file system software. When the system went down for full machine integration in August, most system software issues had been resolved. At year end, the complete BG/L was the fastest machine in the world, having achieved a 280.6-TF Linpack result.

One area of continuing work on BG/L is file system performance and stability. LLNL's parallel file system runs Lustre software developed by Cluster File Systems in collaboration with LLNL and others. Tuning input/output performance on BG/L is a complex issue. Many tuning controls are available in the complex network of input/output nodes, switches, Lustre servers, and disk controllers that form the 896-terabyte (TB) file system. Measured performance writing to the file system is approximately 26 gigabytes per second (GB/s) of the maximum of 32 GB/s. System performance reading from the file system has been as high as 21 GB/s and is still being tuned. Reliability and performance are steadily improving, and the BG/L integration team continues to focus on these areas.

Significance

The BG/L integration demonstrated both the cost and performance characteristics required to obtain a 1,000-TF (or 1-PF) supercomputer, including reasonable power consumption and floor-space requirements. This system is a radical departure from supercomputers of the past. Most significantly, recent results from a domain decomposition molecular-dynamics simulation of tantalum solidification (Section 5.01) demonstrate that BG/L's performance reaches far beyond existing limits for an important scientific problem. A mature, sophisticated software environment must be proven during the next several years to definitively determine the scientific and commercial capabilities of this architecture; however, we are highly encouraged by our early experience.

The BG/L integration team stands in front of the award-winning system.



Contact Information

Kim Cupps, cupps2@llnl.gov

BG/L Web site:
http://www.llnl.gov/asc/platforms/bluegenel/

2.04 — BlueGene/L Early Application Results

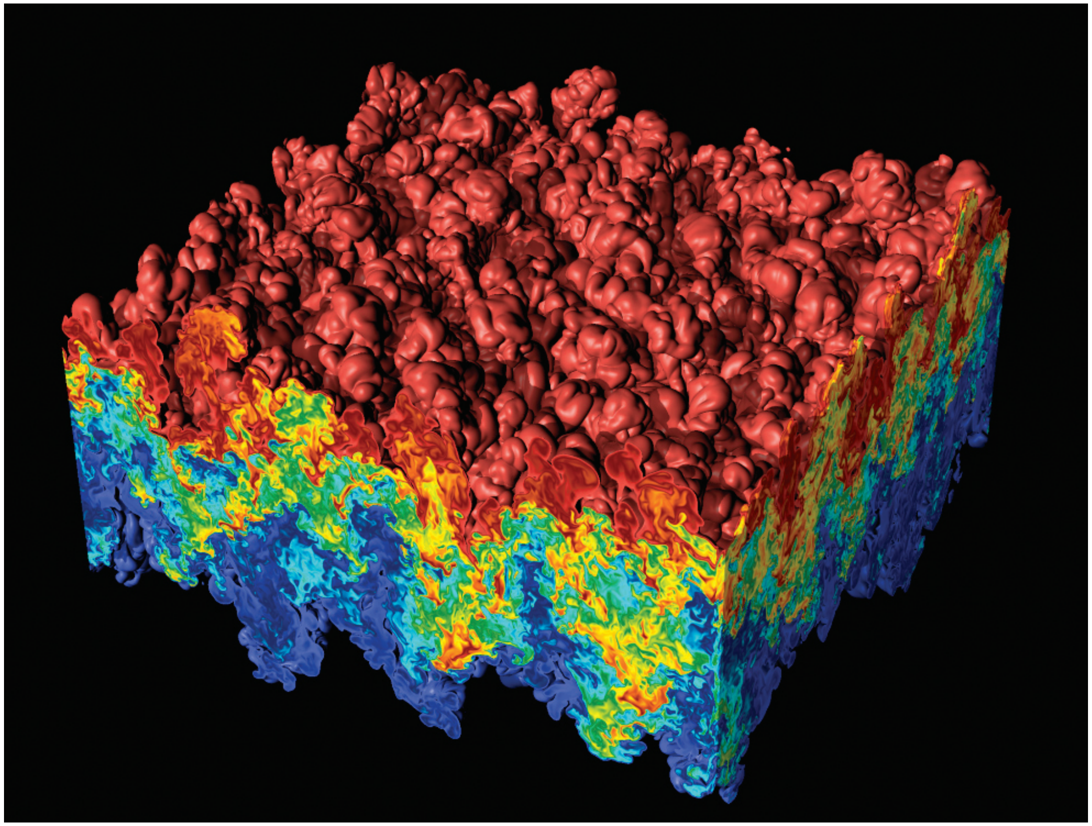
Overview

Initial BG/L applications have brought qualitative changes to ASC materials science and physics modeling and engineering. Results from the most recent code runs on BG/L represent important scientific achievements and a huge step forward to achieving the capabilities needed for NNSA's Stockpile Stewardship Program to succeed. BG/L allows ASC scientists to address computationally taxing stockpile science issues, and the early code runs have provided a glimpse of the exciting and important science results to come. Several code teams, selected for high programmatic relevance, worked tirelessly over the past two years getting ready to run full-scale science calculations on BG/L's 131,072 processors. Specific areas of interest to ASC and the Stockpile Stewardship Program include classical molecular dynamics, dislocation dynamics, hydrodynamic turbulence, and first-principles atomistic calculations.

BG/L will move into classified production in February 2006 to address critical problems of material aging, including equation-of-state characterization, damage evolution, and shock effects. The system's unique combination of high peak speed, scalability, small space footprint, and low power consumption make it an ideal solution for these science areas. With BG/L, problem spaces can overlap. Thus, the largest problem for a quantum-chemistry code can exceed the size of the smallest dislocation dynamics problem, and the largest dislocation dynamics problem can span the range of several crystal grains. Overlap of problem spaces will permit scientists, for the first time, to directly compare different techniques for modeling materials. Such comparisons will help researchers improve model fidelity over larger physical domains and thus better understand the relevant physics.

Progress in 2005

In late August, scientists had their first opportunities to run full-scale codes on BG/L. Running on all 131,072 processors, the LLNL 3D molecular-dynamics code ddcMD achieved a world-record performance for the largest simulation of its kind. The code performed at up to 107 TF with a sustained rate of 101.7 TF over seven hours. In November, this calculation won the prestigious 2005 Gordon Bell Prize for innovation in advanced high-performance computing at the Supercomputing Conference SC|05, held in Seattle. The Qbox first-principles quantum-level chemistry code from LLNL,



Results from full-machine simulations on BG/L using Miranda, a hydrodynamics code developed by LLNL physicists Andrew Cook and Bill Cabot. Miranda studies the behavior of Rayleigh–Taylor instabilities, revealing fluid flow at high Reynolds numbers. These challenging scientific calculations allow physicists to examine situations that are difficult to measure in experiments.

the SPaSM classical molecular-dynamics code from Los Alamos, and a new scalable, breadth-first search (BFS) algorithm demonstration on BG/L were also finalists for the 2005 Gordon Bell Prize. Results from a full-scale BG/L run of Miranda, an LLNL massively parallel code that explores hydrodynamic turbulence, were also presented at SC|05.

The prize-winning ddcMD (see Section 5.01) calculations investigated solidification in tantalum and uranium. Simulation size ranged from 64,000 to 524 million atoms using realistic, but computationally expensive, quantum-based interaction potentials. The ddcMD simulations allow scientists, for the first time, to examine the process of solid formation as a function of temperature and pressure from the atomistic level, helping them better understand the properties of these metals. This ability to model natural nucleation and growth and to create grain structures that reflect directional quantum-mechanical bonding of the atoms was unobtainable before BG/L.

Qbox (see Section 5.02) calculates accurate physical properties of interacting molecules. Simulations on BG/L were performed on up to 32,768 nodes (half of the machine), with performance measurements yielding a strong scaling efficiency of 86 percent. Qbox calculations in 2005 built on the previous year's work to determine optimal layouts of matrices on the BG/L torus. These calculations demonstrated that the immense computing power provided by BG/L can be efficiently harnessed to perform accurate first-principles molecular-dynamics (FPMD) simulations. The wide applicability of FPMD makes these initial Qbox demonstration calculations relevant to research in physics, chemistry, biochemistry, and nanoscience.

Many emerging large-scale-data science applications require searching large graphs distributed across multiple memories and processors. The distributed parallel BFS algorithm, in part developed at LLNL (see Section 5.03), demonstrates scalability on BG/L for random graphs with up to 3 billion vertices and 30 billion edges. Scalability was obtained through a series of special optimizations, in particular, those that ensure the scalable use of memory and efficient collective communication functions for BG/L's torus interconnect architecture.

The SPaSM code runs also demonstrated excellent performance and scalability of a classical molecular-dynamics code on the BG/L platform. Early SPaSM simulations involving up to 160 billion atoms (micrometer-size cubic samples) on 65,536 processors were described at the SC|05 Gordon Bell special session. Los Alamos researchers also reported on two extended production simulations (one lasting 8 hours and the other 13 hours of wall-clock time) of the shock compression and release of porous copper using a realistic many-body potential. These simulations demonstrated the capability for long-running SPaSM simulations, including on-the-fly parallel analysis and visualization of massive data sets.

Miranda, an LLNL massively parallel spectral/compact solver for variable-density incompressible flow, simulates viscosity and species diffusivity effects. Its most recent full-machine calculations, which were performed in part to prepare for the 2006 International Workshop on the Physics of Compressible Turbulent Mixing, reached Reynolds numbers above 20,000, which denotes extremely turbulent flow. Miranda uses fast Fourier transforms and band-diagonal matrix solvers to compute spatial derivatives to at least 10th-order accuracy. Leveraging previous work in 2004, the Miranda team implemented a mapping strategy resulting in virtually perfect scaling up to all 131,072 processors of the BG/L machine.

Significance

The early success of codes on BG/L, the 2005 Gordon Bell Prize, and the ability of BG/L to win all four of the 2005 High-Performance Computing (HPC) Challenge Awards at SC|05 show that the system architecture is clearly well-matched to materials science efforts. These achievements also prove that progress made in 2004 on code porting, scaling, and tuning is bearing fruit. A key part of the NNSA mission is to better understand how aging will affect the nuclear weapons systems that are being maintained in the nation's stockpile beyond their intended design life. The 3D materials and physics simulations on BG/L will provide NNSA scientists with the detailed information they need about the material properties of weapons components, allowing NNSA to meet this important national security mission.

Contact Information

Steve Louis, louis3@llnl.gov
ddcMD: Fred Streitz, streitz1@llnl.gov
Qbox: Erik Draeger, draeger1@llnl.gov
BFS: Andy Yoo, yoo2@llnl.gov
SPaSM: Tim Germann, tcg@lanl.gov
Miranda: Andy Cook, cook33@llnl.gov

2.05 — Remote Customer Services

Overview

In this new era of high-speed networking and easy Internet availability, access to the fastest computers in the world is no longer a concern to HPC customers. The Internet has brought us at LC closer to customers and eliminated the issue of inaccessibility and remoteness. The problem that remains, however, is to eliminate the sociological remoteness that is a consequence of geographic separation. Our intent, and what our customers have come to expect, is that we provide a full spectrum of consulting and customer service to support their research efforts. In particular, we emphasize technical customer service through personalized interactions.

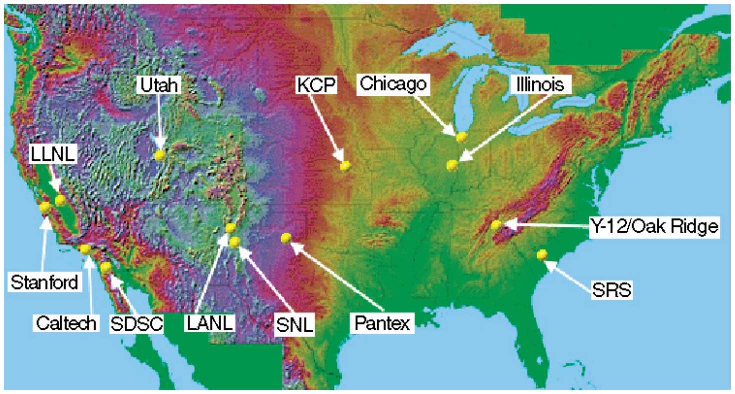
Our HPC customer base consists of 3,189 active users, of which 1,655 compute in the closed environment and 2,474 in the open environment. An increasing number of customers can be categorized as remote—but only in the geographic sense that the LLNL computing resources are some distance away. The 731 remote users are from Sandia and Los Alamos national laboratories, the Academic Strategic Alliance Program (ASAP) centers, the nuclear weapons complex (NWC), and other locations.

The tools, system software, and application infrastructure we develop enable efficient and effective use of LLNL computing resources from any location at any time. We provide extensive user services that improve the accessibility, usability, and reliable operation of LLNL's computing resources.

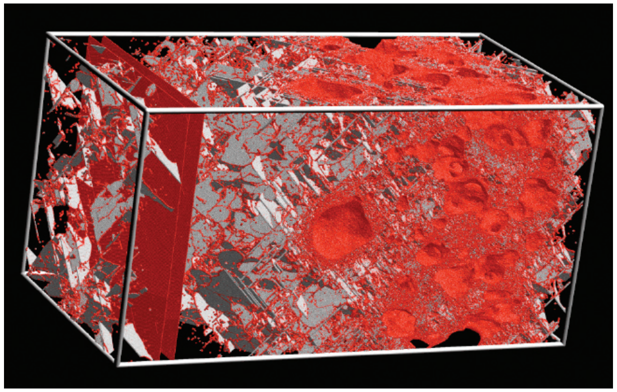
Progress in 2005

Smoothly transitioning to a new trouble-ticket system and successfully moving the LC Hotline to the new TSF with no interruption in customer service were major triumphs. With the successful installation of the two most powerful computers in the world, BG/L and Purple, we trained the hotline staff on using these machines so they, in turn, can assist customers. In addition to continuing our service on all other platforms, we issued accounts on the machines; provided documentation, Web pages, tutorials, and workshops; and began user training. Researchers from the ASAP centers and Los Alamos attended a two-day workshop developed by Computation staff shortly after BG/L became available. During this short workshop, all attendees, with the expert help of the staff, ran codes on 32,000 processors. The ASAP teams have successfully run codes during scheduled dedicated time.

Steve Valone, a Los Alamos code developer who attended the workshop, ported the molecular-dynamics code SPaSM to BG/L. During the workshop, a SPaSM code run that simulated 160 billion atoms on 65,536 processors consistently achieved 24.4 to 25.5 TF, for which the Los Alamos team was chosen as a finalist for the 2005 Gordon Bell Prize.



Remote LLNL customer locations.



Results from a BG/L calculation using the molecular-dynamics SPaSM code developed at Los Alamos. This code is used in atomic-scale studies of aging effects on metals.

In his workshop comments, Valone said, "ASC-[BG/L] users at Los Alamos have felt the benefit of all of the support functions at the LC. We had essential interactions with support staff on architectural issues. We have had access [to] and support from IBM personal. There has been aid in porting codes to [BG/L] and debugging. The training workshop was a big help. And, of course, the all important machine access has been steady and dependable. We all think that the [BG/L] support has been excellent in these many respects."

Monthly ASAP teleconference calls, organized and led by LC, culminated in the annual tri-laboratory outreach visit to each ASAP center. Problems and issues raised at these face-to-face meetings were addressed promptly to provide customers with a more stable computing environment. The annual ASAP site reviews in October 2005 required large machine runs. The ASAP researchers used over 1 million CPU hours on ALC, uP (the 108-node unclassified part of the Purple system), Thunder, and MCR in August. This year, LLNL delivered more than 75 percent of the compute cycles to ASAP on the Laboratory's ASC machines.

A researcher from the University of Utah's Center for Simulation of Accident and Fire Environments (C-SAFE) commented, "C-SAFE center personnel have made considerable use of the supercomputing capability provided at LLNL. Their ALC, Thunder, and [unclassified] Purple systems have provided us with the means to perform our large-scale simulations, without which much of the progress we have made would not have been possible. However, access to the hardware is only one part of the overall service provided by LLNL that is necessary for us to get our job done. While we have extensive experience in porting our large Uintah simulation code to new machines, LLNL has provided us access to their personnel who have in-depth knowledge of each of their architectures. Additionally, more mundane, but equally valuable services we have received from LLNL include expeditious creation of new accounts and Unix groups as we need them, help with management of the large volume of data we produce, monitoring the status of our long running jobs to ensure maximal use of the machine, keeping us informed on the current status of the machines as well as long-term plans for upgrades and phasing in of new architectures."

The figure at right shows an image from a C-SAFE calculation of industrial flares used to vent and burn waste gases from oil and gas drilling and refining operations. This simulation is the first in the world to show the time-dependent characteristics of such flows and to accurately model the changes that result from different wind conditions.

An important customer service is the NWC help desk support, which is an outgrowth of the tri-laboratory user support effort between Lawrence Livermore, Sandia, and Los Alamos national laboratories. LC's role in this application support has increased substantially as the number of NWC application codes available to the engineers continues to grow. We collaborate closely with Sandia, Los Alamos, and the NWC plants to provide support for these codes.

An increasingly important goal is for the help desk staffs at the NWC plants to work together to improve code reliability and accessibility and to extricate the application code teams from account and access issues. In support of this goal, we facilitate sitewide integrated processes and communication to NWC applications and infrastructure among plants. We also provide a central help desk for customers of NWC applications. Our trouble-ticket system, which we use to field requests, questions, and suggestions, allows us to track issues from the time they are reported until they are resolved and closed.

Significance

The world's fastest computers have value only if they can be used. Our mission is to provide the necessary service to enable world-class science. We measure our success by our users' ability to exploit the power of our hardware and software. Our staff provides the required interface with the scientific community to achieve this mission. The power of our hardware is fully realized through human interaction and our personal intervention.

Results from a C-SAFE simulation of industrial flares used to vent and burn waste gases produced by oil and gas drilling and refining operations.

Contact Information

Jean Shuler, jshuler@llnl.gov
Greg Tomaschke, tomaschke1@llnl.gov

2.06 — Strategy for Multiprogrammatic and Institutional Computing

Overview

The strategy for M&IC has extended leading-edge simulation on scalable platforms beyond the ASC program at LLNL. By working closely with the Stockpile Stewardship Program and leveraging mutual resources, M&IC and ASC have made LLNL a premier institution for computational and simulation science. Such standing is vital to the continued success of LLNL science programs and the recruitment and retention of top scientists. M&IC continually works at balancing investments to meet the widespread demand for capacity computing and to provide cost-effective capability platforms.

Progress in 2005

Lawrence Livermore has unquestionably reached a position of distinction in the large, competitive HPC community. Our unprecedented HPC environment is helping to establish simulation as a peer to theory and experiment in the process of scientific discovery. Our 3D simulations model physical phenomena that were previously out of reach. Computation is essential when analytically intractable models are explored or complex multiphysics phenomena must be understood quantitatively. As scientists improve their understanding of truly basic science, the Laboratory needs computational resources to make the vital quantitative connections among disparate phenomena that constitute the foundation of both pure and applied science.

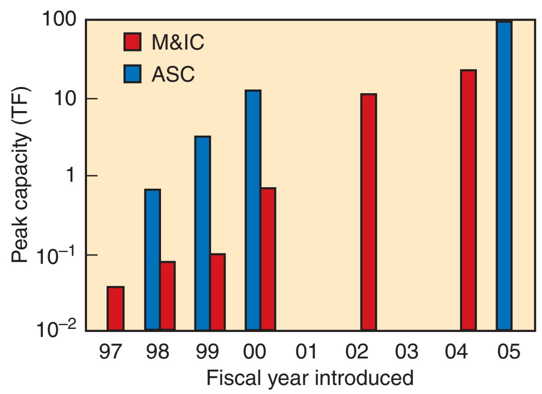
Because of the Laboratory's strong and consistent investments, M&IC has grown into a powerful unclassified computing resource that is being used to push the limits of computing and its application to simulation science. Computation is now a mainstream method in theoretical science at LLNL, and the Thunder and MCR clusters provide Livermore scientists with access to world-class unclassified systems.

It is one thing to reach a stature of preeminence, and an altogether different undertaking to sustain that position. M&IC has developed a FY05–FY08 plan for leveraging ASC developments so we can continue to provide a stellar environment for scientific discovery at LLNL. Competitive advantage can be achieved by maintaining the quality of the simulation environment, thereby enabling seasoned disciplines to integrate theory and experiment through validation and verification.

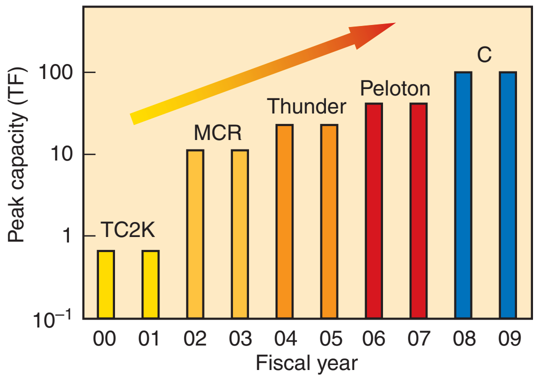
ASC has provided research and development funds that allow M&IC to invest in high-performance but low-cost infrastructure systems—often before other sites can do so. Because of ASC investments, LLNL benefits from one of the most experienced and well-staffed scientific computing centers in the world. An investment in hardware is leveraged by attention from experienced integrators, operators, and services staff and from a well-engineered foundation in networks and storage. Experience gained on M&IC resources can likewise be used to further ASC goals. All of these efforts mitigate the considerable risks inherent in investing in the newest and best cost-performance technologies.

An essential part of this strategy is to continue to acquire new computer systems. Although more than 50 important classified and unclassified simulation codes are available to our computational scientists, we are challenged by the high demand for access and the long wait times to queue code runs. Simply stated, the machines are saturated.

To meet the needs of the institutional computing community, M&IC received funding over several years to procure a new cluster system. Called Peloton, this system will provide approximately 40 TF of capability cycles and an additional 5 TF of capacity cycles in addition to those available on MCR and Thunder. We also are encouraging external collaborators to access LLNL resources to expand and sharpen our scientific thrusts on important science



Comparison of M&IC to ASC capabilities from 1997 to 2005. Throughout this time, LLNL's institutional systems have remained close to the computational power available on ASC machines.



As M&IC plans for the future, we have identified the platform requirements to maintain LLNL's preeminence in high-performance computing.

and technology priorities, including those established under the Laboratory's Aurora project. To this end, we are procuring a disk farm, known as a collaborative Data Oasis, which will reside on the open network. The Data Oasis file system will allow us to better accommodate external collaboration and data sharing with foreign nationals, who are an important and growing sector of the nation's scientific community. Healthy partnerships with our industrial and academic colleagues in the HPC community are also essential to success. Sustaining LLNL's preeminence in high-performance computing is vital for the Laboratory's science programs and dramatically improves our ability to recruit and retain top scientists.

This past year was one of substantial accomplishments for the M&IC program:

- The Thunder system is running grand challenge problems. That is, the system is available to a limited number of projects, which are allocated large banks of computing cycles to run aggressive problems for breakthrough science.
- We completed a 2005 white paper on institutional capability computing requirements, and our request for Peloton funding was approved.
- We procured the Data Oasis file system to enable external collaborations.
- Research on M&IC platforms resulted in a significant number of publications, presentations, and external collaborations for LLNL scientists. The complete M&IC bibliography, available as a separate document, includes journal articles (282), presentations (95), other publications (32), and ongoing collaborations related to terascale resources.
- The M&IC environment is comparable to the nation's best unclassified environments, and by partnering with the ASC program, we have achieved this success with a smaller expenditure than might be expected.

Significance

High-performance computing is the integrating element of the scientific discovery triad formed by theory, experiment, and computing. At LLNL, HPC resources are in high demand by many programs. Evidence suggests that researchers will rely even more heavily on computational resources in the future than they do today. For instance, several of the Laboratory's proposed Aurora initiatives depend on and assume a robust HPC infrastructure. Therefore, sustaining our distinction is more important than ever before.

Simulation and related data-analysis efforts will be a key in the ongoing competition between nation-states, industries, and research institutions. The quality of the simulation environment and the ability of seasoned disciplines to integrate theory and experiment through verification and validation will help the Laboratory maintain its competitive advantage. Strategic foresight and a talented staff have allowed us to enjoy great success in the past, and we are positioned to experience similar success well into the future.

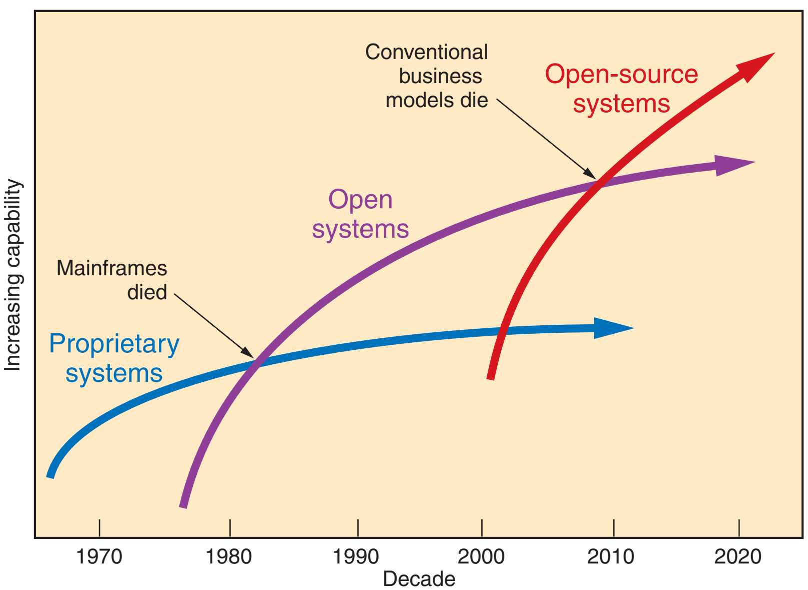
Contact Information

Brian Carnes, carnes1@llnl.gov

2.07 — LC Spectrum of Open-Source Activities

Overview

Livermore Computing (LC) is an active promoter of open-source software. The current collection of open-source activities spans a surprisingly wide range, from external collaborations to local code development to product testing. LC acts as an open-source sponsor (funding source), licensee, developer, user, and tester. With each product, a different aspect of open-source software emerges, making up a spectrum of software integration models. The project areas encompass operating systems, resource management, file systems, and software tools. Many, but not all, of the projects listed in Section 6.01 (Industrial Collaborators) involve open-source development, and specific project contact names are given there.



Disruptive software technology trends. During this period, software technology has been changing as rapidly as hardware technology.

Progress in 2005

The Clustered High Availability Operating System (CHAOS) project deployed release 3, based on Red Hat Enterprise Linux 4, with support for Infiniband and Quadrics networks. CHAOS is installed on all of the Linux-based LC systems. LC licenses the Red Hat software and serves on the Red Hat customer advisory board, which allows input into Red Hat's strategic directions. LC continues active involvement in the ASC PathForward collaboration with Hewlett-Packard and Cluster File Systems in deployment of the Lustre file system. In 2005, Lustre was deployed as the BG/L global file system, supporting 900 TB of disk capacity at 30-GB/s bandwidth. Lustre software is released as open source under a general public license. LC also made extensive use of the locally developed open-source software package IOR, a tool for benchmarking parallel file systems, to identify many issues with the parallel file system on Purple and to test the performance of the Lustre system on BG/L.

The recent deployment of LC's Simple Linux Utility for Resource Management (SLURM) software on both Purple and BG/L has been an important contribution of LLNL-developed open-source software. SLURM replaced IBM software, giving users additional portability with the Linux capacity systems. SLURM is now deployed on hundreds of clusters.

Software for message-passing interface profiling is now available on additional systems, including the Cray X1. LLNL is an active member in the developer community for the open-source software Python, particularly in its use for parallel systems. Python is being used more widely in HPC. LC, together with Los Alamos and Sandia, sponsors two external open-source tool projects: Valgrind and Open SpeedShop. The Valgrind system emulator is a useful tool for memory debugging. LC is helping to develop it into a stable product for use with parallel programs on all of the major ASC systems.

The Open SpeedShop project is a PathForward effort with Silicon Graphics, Inc. (SGI), to develop a software performance instrumentation infrastructure for Linux systems. Beta release software was made available in late 2005. According to the SGI business model for this software, SGI will release the infrastructure and commonly used performance experiment plug-ins as open source. Subsequently developed proprietary plug-ins will be offered for purchase.

Significance

LC has a broad portfolio of open-source projects over a range of development models. As a result, LC can harness the power of the open-source development process and leverage huge external investments to meet local programmatic deliverables. Overall, the open-source projects form important partnerships, matching LLNL expertise with external resources to produce software that benefits the entire HPC community.

Contact Information

Mary Zosel, mzosel@llnl.gov
Mark Seager, seager@llnl.gov

Open-source projects: http://www.llnl.gov/computing/hpc/resources/os_projects.html

SECTION 3

IT Services and Computer Security

3.00 — IT Services and Computer Security

The Computation Directorate provides support for much of the information technology (IT) infrastructure at Lawrence Livermore National Laboratory (LLNL). This work involves managing the Laboratory's networks and desktops and providing security solutions for LLNL and other Department of Energy (DOE) sites. Executing this responsibility requires a great deal of coordination and cooperation between the organizations and the personnel in the directorate as well as interaction with other programs at LLNL and DOE.

The directorate is proud of its 2005 accomplishments. In the IT services area, we focused on four goals: security, manageability, efficiency, and productivity. With this approach, we continued to increase the software and services available to managed desktops. At the same time, we deployed Active Directory and automated tools, while establishing best practices for configuration management. A benchmarking study of other scientific organizations helped to validate our current IT strategy, and through it, we established working relationships with other sites. The progress made in 2005 resulted in greater efficiency and security for our desktop environment and allowed our users to be more productive.

Securing LLNL's facilities and the information produced is extremely important to us. The Computation Directorate supplies the personnel that support, create, and research new solutions for cyber security to be used both at LLNL and at other DOE sites. The Secure Communication and Teleconference (SCAT) system is used for nationwide secure communication. It is both an information management system and a communication system developed at LLNL and has been operating for the past 20 years. In 2005, we completed a major upgrade to make the SCAT system Web-based and to allow it to span multiple networks. Once we transfer data to the new system, the aging hardware infrastructure will be retired, and automated processes will replace the manual processes currently used to distribute information. This system has been vital to DOE for secure communication, especially during times of emergency preparedness.

Personnel from the Computation Directorate helped LLNL's Cyber Security Program (CSP) make significant progress. With a defense-in-depth philosophy, CSP has not allowed even one major disruption in service, although LLNL continues to be attacked at an ever-increasing rate. In 2005, more automation was employed to allow automatic blocking and unblocking of unregistered hosts. Additionally, firewall egress filtering was implemented, spam filtering software was deployed, and remote access services were updated. The DOE accreditation of LLNL's unclassified systems was a significant accomplishment of 2005.

The cyber security threat we face from network intrusions has continued to grow at LLNL and across the Internet at other sites. To better understand the possible threat, we are using visualization techniques to determine whether patterns of activity can be catalogued. Using wavelet scalograms, we characterized patterns produced by large network scans. Using this technique is helping us investigate the hostile network activity that attacks our networks and systems.

3.01 — Secure Communication and Teleconferencing

Overview

The SCAT system is a nationwide, secure, communication and information management system. It was developed by Computation personnel who support the rapid-response operational mission requirements of the Nonproliferation, Arms Control, and International Security Directorate's Nuclear Assessment Program (NAP). SCAT is also used by other emergency preparedness and response organizations. The system has been in continuous operation since it was deployed in 1986 and is now in its sixth major software version (SCAT V6).

During the summer of 2004, NAP initiated a Web application project to upgrade SCAT to version 7 (SCAT V7) so it will better meet customer needs. The SCAT V7 architecture is Web-based, with specific information management, security, and usability features designed to support NAP and other emergency response mission requirements. Specialized hardware linkages and software interfaces will allow different instantiations, or clones, of the system to exist on and span multiple networks.

The SCAT project consists of two phases: upgrading the existing V6 system to the V7 Web application, and subsequently cloning and linking SCAT V7 servers on other networks.

Progress in 2005

Significant progress was made on the first phase of the SCAT project in 2005. The Web application code moved from early development through alpha testing to final beta testing. Security planning and administration for SCAT V7 is also nearing completion; only final accreditation by the cognizant authority is needed before initial production operation begins in early 2006.

Also in 2005, a new server infrastructure was specified, designed, acquired, installed, and tested. This infrastructure includes a high-reliability server farm at LLNL, including the V7 Web server, a relational backend database server, a backup server, and router and firewall appliances.

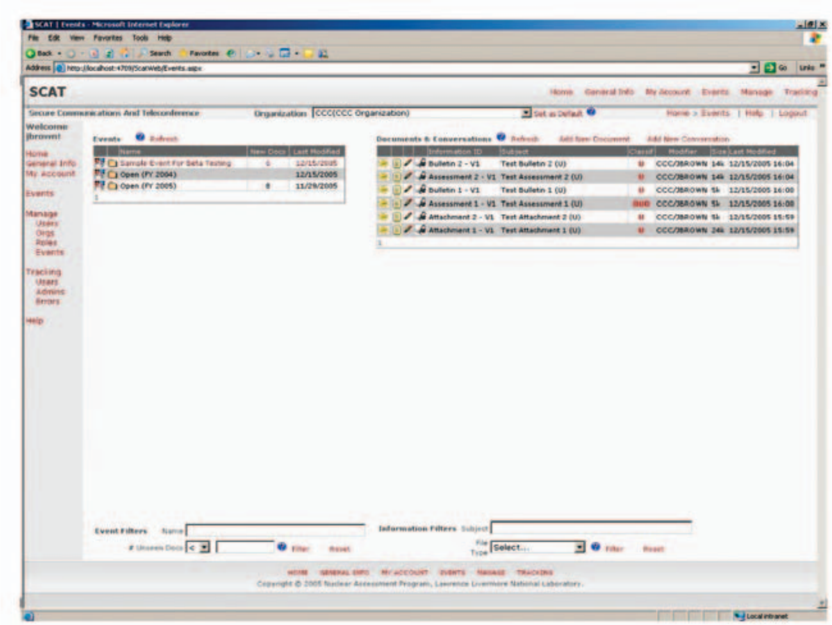
A key part of the V6-to-V7 evolution was migrating and converting all existing data to the new version. In this process, thousands of documents were transferred from the existing SCAT V6 DEC Alpha/VMS system and digital relational database to the SCAT V7 Windows servers and Oracle database, and SCAT V6 documents were converted from text to portable document file (PDF) format.

The SCAT V7 Web application is also being deployed in an unclassified environment to help the Department of Homeland Security evaluate radiation detection alarms. By replacing current e-mail-based pathways for disseminating information with the SCAT V7 Web application, we can significantly enhance information management of this important technical data while streamlining overhead and administration costs.

Significance

When SCAT V7 enters production, SCAT V6 will be phased out after a brief period of parallel operation. Once the legacy V6 hardware is retired, a new and robust architecture will be in place for this mission-essential national capability. The SCAT V7 system encompasses significant user interface and functionality enhancements, which new and future users will find much more intuitive than the legacy system. System performance and information throughput will be greatly improved. Administration and operation of the new system will be significantly streamlined.

Once the main SCAT system has been implemented and its clones are spanning multiple networks, automated processes will replace the manual processes that are currently used to disseminate important information at a variety of classification levels across the networks. This automation will allow users on each network to readily access the appropriate information without the time delays imposed by manual processes.



Screen shot of the SCAT version 7 Web application.

Contact Information

John Brown, brown174@llnl.gov

RETURN TO TABLE OF CONTENTS

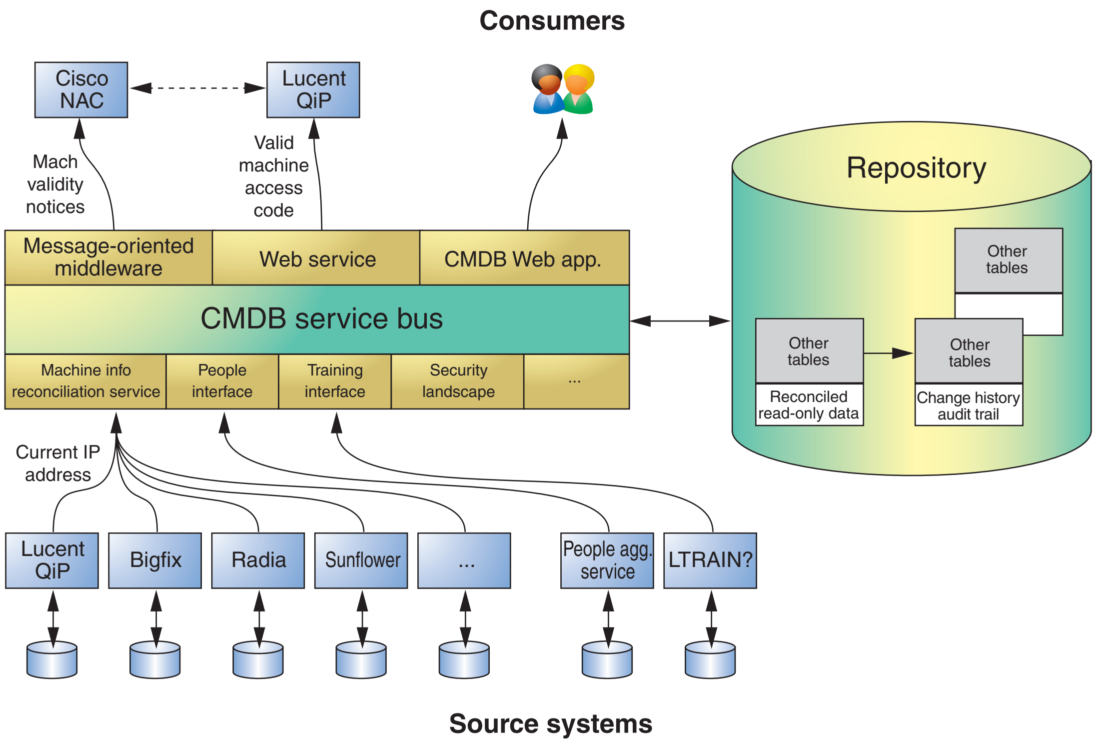
3.02 — Managed Systems

Overview

As the responsible organization for most of the Laboratory's IT environment, including desktop and server infrastructure, the Computation Directorate provides system administration and networking personnel to programs throughout LLNL. The suite of software and services offered on the desktop computing platform grows in sophistication each year. At the same time, this set of desktop services has become an integral part in enabling more than 10,000 users to significantly contribute to the Laboratory's science and business each day.

In 2005, we continued to pursue the four goals that form part of the vision we established in 2004: security, manageability, efficiency, and productivity. Computation supports a combination of more than 20,000 Windows, Macintosh, and Linux platforms, including desktops, networks, and servers that, behind the scenes, are crucial to smooth and continuous Laboratory operations. The efforts of many Computation employees, operating as a cohesive team, allow for secure and efficient solutions to be developed and deployed to end users, resulting in a productive computing environment.

Although progress was made across several fronts in 2005, an ongoing challenge is software vulnerabilities. Not only has the number of software vulnerabilities increased, but also the timeframe permitted to remedy them has decreased. The current environment dictates that a response be made in days or hours rather than weeks or months. Given the criticality of responding quickly to a diverse set of challenges, we have made configuration management one of our major thrust areas. A focus on automation, common tool sets, and improved support processes has resulted in more effective responses and more efficient compliance reporting.



Configuration management database architecture.

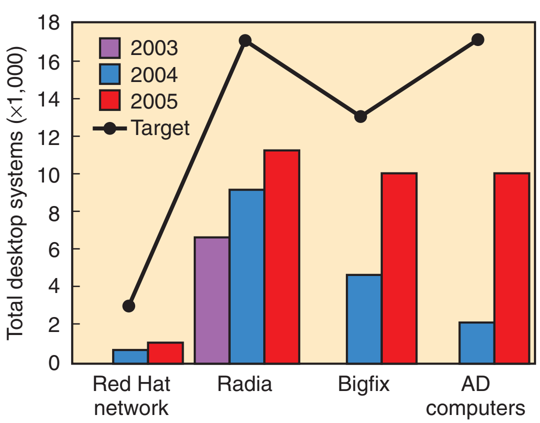
Progress in 2005

Effectively meeting the desktop and server management challenges requires the ability to automatically manage system configurations and user accounts. In 2005, desktop management efforts continued to focus on automated tool deployment and establishing best practices around configuration management.

We initiated two major activities in 2005 to further define and implement the strategic vision for desktop management. The first was a benchmark study in which directorate personnel visited external scientific organizations similar to LLNL to gather information about best practices for IT support. The results of the study validated LLNL's existing desktop strategy in areas such as wireless networking, automated tools, and centralized support. The relationships established with the external organizations will be important resources as we pursue other best practices identified during the study, such as effective collaboration suites and efficient IT policy management.

The second activity, sponsored by LLNL's Chief Information Officer, is an initiative to develop a configuration management database (CMDB) for the Laboratory. The approach is to leverage existing sources of information about systems and users to provide a seamless, integrated view of the information needed to meet compliance requirements and support business decisions about the IT environment. Products purchased for server and network management complement the existing desktop management tools. The result is a suite of tools (Bigfix, Radia, and Opsware) to manage and collect information about configurations. Deployment of these tools has reached critical mass on the desktop and will continue to grow in the server and network arenas in 2006.

The Active Directory (AD) infrastructure provides the foundation for managing desktop computers and users, and is leveraged by the desktop management tools (Radia and Bigfix). With the majority of Windows systems already in AD, the focus in 2006 will be to integrate Macintosh and Linux desktops into AD for account and policy management. Also in 2006, we will integrate AD with LLNL's identity management system to enable automatic user account provisioning and deprovisioning.



Automated tools deployed since 2003.

The institutional help desk pilot study, completed in 2005, demonstrated that centralizing first-level desktop support greatly increases the efficiency of user support. The pilot showed that the central help desk could handle 35 percent of the work currently done by technicians in the field in about half the time. These efficiency gains occurred with no loss in customer satisfaction. Because of our success with this pilot study, additional directorates will use the central help desk for their first-level support, beginning in 2006.

Significance

The progress made in 2005 resulted in more efficient and secure management of desktops and a more stable and productive environment for users. The automated tools we deployed enabled us to automatically collect configuration evidence. As a result, our restricted open networks were successfully certified and accredited, which allowed the directorates to continue to operate and meet their missions. Using automation to patch Windows system vulnerabilities enables thousands of systems to be patched within hours rather than days and requires significantly less administrator effort.

The strategic implementation of a CMDB will provide more effective system and infrastructure management. The information it collects will support business decisions and help LLNL plan and manage hardware refresh cycles, software site licenses, and computer security risks. The successful central help desk pilot study paved the way for increased efficiency and improved customer support. Our vision provided a roadmap for prioritizing and investing in IT infrastructure improvements. We continue to move LLNL toward the objectives that enable security, manageability, efficiency, and productivity.

Contact Information

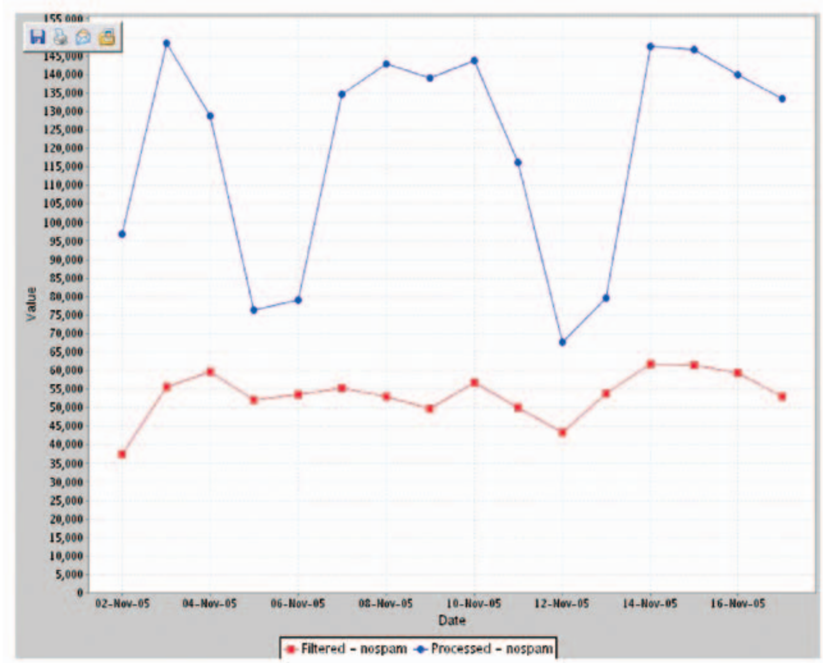
Greg Herweg, gherweg@llnl.gov
Robyne Teslich, teslich2@llnl.gov

3.03 — Cyber Security Program

Overview

The Cyber Security Program ensures LLNL compliance with DOE orders relating to computer, network, and telecommunications security. CSP is also responsible for providing training on and promoting awareness of cyber security issues for the Laboratory's technical and administrative staff. CSP coordinates and represents the Laboratory in negotiations between LLNL and other government agencies on issues of cyber and communications security.

During 2005, Computation Directorate personnel helped CSP make significant progress in enhancing the security of LLNL's computing assets. Cyber Security personnel use a variety of approaches to coordinate and manage cyber security functions and to protect LLNL's networks and computer infrastructure from compromise and attack. Through risk analysis, assessments, and management, we can identify risks to LLNL systems and implement an optimum mitigation strategy. In addition, CSP provides technical expertise to Laboratory personnel to help resolve computer security problems or issues that affect computing operations.



During a typical workweek at LLNL, Brightmail™ filters about 40 percent of all inbound e-mail as spam. On weekends, the amount of spam increases to about 80 percent.

Progress in 2005

CSP personnel are challenged daily to enable the use of new technologies despite the ever-increasing security vulnerabilities and requirements. Key achievements in 2005 include the first accreditations of directorate-specific unclassified systems and the development of a long-term strategy for complying with National Nuclear Security Administration (NNSA) policy documents.

The past year also brought more automation to CSP, such as the ability to automatically block unregistered hosts that appear on any of LLNL's networks and automatically unblock the systems once they become registered. CSP improved the process (defined in the computer security plan) used to approve computer access for foreign national employees. Improvements included the deployment of a Fast Track approval application, which allows directorate-level approval for approximately 50 percent of the plans. The Fast Track streamlined process greatly reduces delays associated with the legacy CyberTrak application. CSP also hosted a series of process improvement workshops to guide the development of a replacement application.

E-mail filtering is an ongoing concern. In early 2003, CSP began using Brightmail™, an e-mail filtering software product. In its first year of use, the software filtered an average of 10,000 unwanted e-mail messages per day. By the end of 2005, this number had grown to an astounding 60,000. On any given day during the workweek at LLNL, Brightmail™ filters approximately 40 percent of all inbound e-mail messages as spam. This number increases significantly over the weekend when approximately 80 percent of all inbound e-mail is filtered as spam. As the amount of spam continues to increase, so does the effort placed on preventing it. In 2005, CSP asked LLNL employees to assist in the effort by submitting messages missed by the spam filter to a Web site where the unwanted messages can be analyzed and used to construct additional filters.

In 2005, an institutional Lightweight Directory Access Protocol (LDAP) server provisioned with the official user names and personal access codes of employees was released. This authentication silo is being used for LDAP-aware applications such as Remedy and for open network authentications into services such as U-Learn. CSP is working with the Laboratory's Administrative Information Services (AIS) to migrate all institutional servers managed by AIS to this infrastructure for authenticating logins. In addition, CSP is working to integrate Macintosh and Unix desktops into this environment.

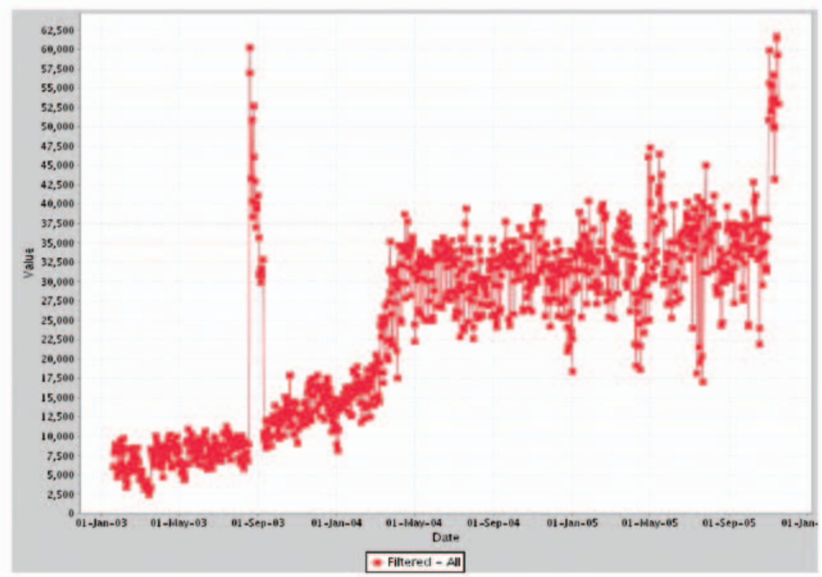
This year, CSP contributed to improvements for both classified and unclassified Entrust communications between DOE sites. CSP worked with Livermore Computing system administrators to implement a classified instance of the Entrust public-key infrastructure (PKI). This implementation will enable the Laboratory to reduce the amount of classified accountable removable electronic media it generates and communicate in a secure manner with other national laboratories. Additionally, CSP worked with all DOE sites to ensure that the unclassified Entrust PKI was registered with the Federal Bridge. This effort helped facilitate work with Homeland Security Presidential Directive HSPD-12.

The Laboratory's remote-access services were updated in 2005. Older systems were retired and all users were converted to the new Cisco VPN 3000. A Secure Socket Layer–based VPN appliance was also purchased and deployed in a test environment. This new offering will help consolidate remote-access services and will provide a better experience for remote-access customers.

National Computer Awareness Day was highlighted by CSP in November. Four short presentations, given by staff experts, provided employees with tips to heighten their security awareness when using LLNL computers at work and personal computers outside the work environment. The talks were videotaped and made available as a QuickTime movie on the Laboratory's internal Web site.

Significance

In addition to protecting LLNL from cyber security incidents, CSP delivers on its mission to provide training for employees and raise awareness of cyber security issues across the Laboratory. CSP has been instrumental in upgrading systems and services, which will ensure that the high level of quality and security required by LLNL and its sponsors will be maintained into the future.



This chart on the e-mail filtered by the application Brightmail™ shows a dramatic increase in unwanted messages (spam) between 2003 and 2005.

Contact Information

Edward J. Matsche, matsche1@llnl.gov

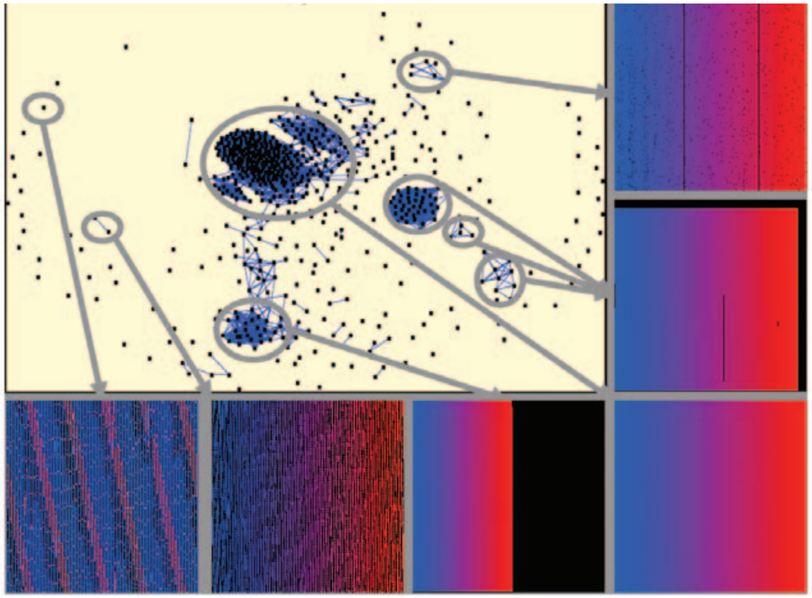
3.04 — Internet Ballistics Demonstrates Cyber Forensic Capability

Overview

Every day, millions of hostile network probes sweep the Laboratory in search of vulnerable systems and services. These scans are often conducted from systems that are compromised by worms or by our adversaries, through "botnets" (a collection of malicious software robots), anonymizing relays (which mask a hacker's identity), or other indirect methods. All of these methods make the source Internet Protocol (IP) address of the scans an unreliable indicator for tracking adversarial behavior.

Preliminary scan visualizations conducted by the Computer Incident Advisory Capability (CIAC) revealed that these scans inhabit a rich population of distinctive patterns in packet target sequencing and interpacket arrival timing. Moreover, an uncanny correlation appeared between similar visual patterns and the apparent source IP addressing.

The Internet ballistics study determines the degree to which this pattern similarity can be attributed to the actual network location of the scan source or is characteristic of the scan methodology of the adversary (such as schedules, tools, or platforms) independent of transient-source IP addressing. Effective analysis requires both a mathematical means for assessing pattern similarity and a large body of controlled scans generated as a baseline for validating models and refining measures.



Linlog force-directed clustering of wavelet scalograms as *n*-dimensional vectors serve to group and isolate scans.

Progress in 2005

Our studies used wavelet scalograms to characterize the patterns produced by large network scans. These scalograms represent a 500- to 1,000-fold reduction in volume of data, yet remain faithful to the underlying pattern. We investigated more than 40 wavelets to determine the most useful for encoding the multiscale variations present in the scans.

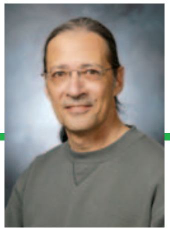
Collaboration with Kwan-Liu Ma from the University of California at Davis and graduate student Chris Muelder produced a visualization tool that helps judge the effectiveness of competing measures in distinguishing scan characteristics. The tool applies Linlog force-directed clustering on the characteristic wavelet scalograms and allows clusters to be selectively investigated to validate the accuracy of the scalogram representation (see the figure at right).

A closed test network provided the first baseline data for the controlled scanning activity. We designed the network to directly capture the microsecond-level behavior of various hostile scanning tools and will use it to compare identical control scans directed through varied Internet routing. A temporal IP sensor we developed and optimized captures high-quality time-stamped scan traffic in support of the investigation. In initial tests involving hundreds of scans using Nmap, Hping2, and Strobe scanning tools, Internet ballistics methodology distinguished the tools and even produced distinct clusters according to the command line settings used by the (simulated) adversary. When we applied these methods to almost 1,000 unlabeled field scans, many large clusters formed that were constrained to specific destination ports—a strong indication that specific attack tools, crafted to ply particular ports, were at work.

Significance

The ability to consistently identify Internet adversaries, in both cyber security and Internet counterintelligence, is of paramount importance. In the realm of cyber security, the consistent appearance of adversaries can be used as forward indicators of undiscovered vulnerabilities. In the mission of counterintelligence, the ability to partition the world of hostile network activities into "players" consistently identified by features other than source IP address would provide a wealth of support, helping us to connect the dots on activities that might otherwise appear random and disconnected.

Contact Information



Tony Bartoletti, azb@llnl.gov

SECTION

4

Developing Applications Software

4.00 — Developing Applications Software

Every major program at Lawrence Livermore National Laboratory (LLNL) partners with the Computation Directorate in developing the software needed to meet its responsibilities to the National Nuclear Security Administration (NNSA), the Department of Homeland Security (DHS), and other agencies. Highly skilled computer science professionals in Computation are matrixed to the programs and work closely with domain experts to design and implement these applications. Applications span a range of scientific and technical disciplines: fluid dynamics; particle transport; genome sequencing; efficient, scalable, parallel multiphysics frameworks; computational modeling of materials under extreme conditions; real-time control systems for large science experiments; and management and analysis of large sets of information from disparate sources.

The directorate employs about 550 computer scientists and applications programmers, making it one of the largest applied computer science organizations in the world. The partnership arrangements between Computation and LLNL programs vary with each program's needs, but some features are common to most of them. For example, programs often rely on a small number of computer codes that grow and evolve over time to meet specific needs. Computer scientists assigned to these programs work with domain experts to define, implement, and optimize these codes. Projects range from small teams with 1 to 5 members, to large, interdisciplinary groups with up to 20 or more computer scientists, engineers, and domain experts.

The required computer science expertise differs from one project to the next. For example, stockpile stewardship applications demand expertise in parallel computing. In contrast, control-system applications involve developing interfaces to manage hardware and to meet unique testing and quality-assurance requirements. Other applications, such as those for biodefense, require expertise in managing large, domain-specific databases and scalable algorithms for searching and pattern matching.

In this section, we highlight a few of the programmatic efforts for which the Computation Directorate provides expertise. We describe our 2005 work on applications for the National Ignition Facility's (NIF's) control system to automate major aspects of shot preparation, execution, and postprocessing. We also report on the efforts of our software quality engineers who are working with programmatic code teams to help LLNL achieve the goals of its Institutional Software Quality Assurance Implementation Plan. In the past year, we made significant contributions to the hardware and software used for detecting and analyzing biothreat pathogens. The codes Purgatorio and XEOS, which simulate the thermodynamic properties of materials, have matured and in 2005 made substantial contributions to equation-of-state data that are fundamental to hydrodynamics simulations. In addition, we have helped develop several information management and analysis tools for seismic research at LLNL. Computation scientists are also collaborating with scientists and engineers from the Stanford Linear Accelerator Center (SLAC), Argonne National Laboratory, and the University of California at Los Angeles (UCLA) to build the Linac Coherent Light Source (LCLS), which will be the world's first x-ray free-electron laser. Each article highlights how we have applied our skills and expertise to contribute to solving problems.

4.01 — Shot Automation for the National Ignition Facility

Overview

A shot automation software framework developed at LLNL was deployed during the past year to automate shots performed on NIF, using the Integrated Computer Control System. The new model-based software automates the entire complex shot sequence from the initial setting of operating parameters for the experiment to the final countdown to charge and fire the laser.

Progress in 2005

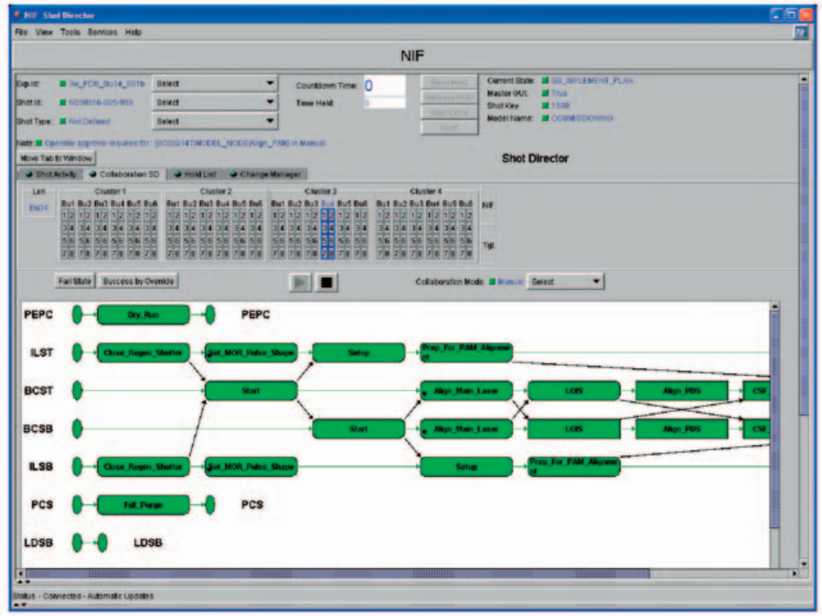
The new shot automation framework automates a four- to eight-hour shot sequence that includes entering shot goals from a physics model, setting up the laser and diagnostics, automatically aligning the laser beams, and verifying status. This sequence consists of a set of preparatory verification shots, leading to amplified system shots. The high-precision timing system has a four-minute countdown procedure and, during the last two seconds, triggers a shot, which is followed by post-shot analysis and archiving. The framework provides a flexible, model-based execution controlled by scriptable automation called "macro steps."

The framework is driven by high-level shot director software. This software provides a restricted set of shot life-cycle state transitions to 25 collaboration supervisor systems, each of which automates a bundle of eight laser beams. A common set of resources is shared by all. Each collaboration supervisor commands approximately 10 subsystem shot supervisors—systems that perform automated control and status verification. A collaboration supervisor translates shot life-cycle state commands from the shot director into sequences of macro steps to be distributed to its subsystem shot supervisors. Each shot supervisor maintains the order of macro steps for the subsystem it controls and supports collaboration between steps. The collaboration supervisor also manages failure, restarts, rejoining the shot cycle (if necessary), auto or manual macro-step execution, and collaborations between other collaboration supervisors.

Shot supervisors execute the macro-step shot functions as commanded by the collaboration supervisor. Every macro step has database-driven verification phases and a scripted perform phase. This process provides for a highly flexible methodology for performing a variety of NIF shot types. Database tables define the order of work and dependencies (workflow) of macro steps to be performed for a shot. A graphical model editor (shown to the right) allows NIF personnel to define and view an execution model. A change manager tool enables "de-participation" of individual devices, of entire laser segments (beams, quads, or bundles of beams), or of individual diagnostics.

Significance

The shot automation software deployed to NIF is being used to support laser commissioning shots and build-out of the NIF laser. The new software is a fundamental change to the operational paradigm of NIF and enabled the first-ever multiple quad shots. This system will be used to automate all future target and experimental shot campaigns.



Graphical user interface for the automated NIF shot director shows options available along with the scheduled operations in the model-based workflow.

Contact Information

Larry Lagin, lagin1@llnl.gov

RETURN TO TABLE OF CONTENTS

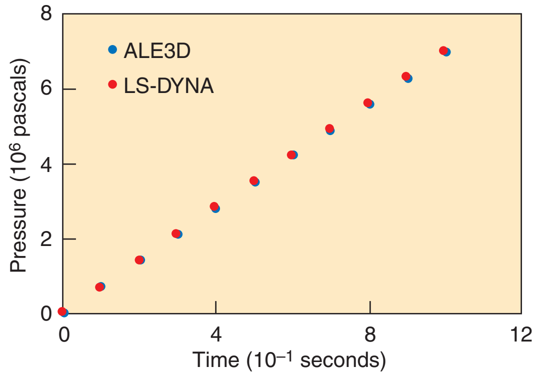
4.02 — ASC Software Quality Engineering

Overview

The Advanced Simulation and Computing (ASC) program in the Department of Energy's (DOE's) NNSA is a vital component of the nation's stockpile stewardship. This mission is to give the U.S. adequate confidence in the safety and readiness of the nuclear weapons stockpile without nuclear testing. The ASC program provides the simulation and modeling capabilities and technologies needed to combine new and old experimental data, past nuclear-test data, and past design and engineering experience into powerful tools that can be used to assess design modifications, certify nuclear weapons and their components, and analyze relevant experiments.

The ASC Verification and Validation (V&V) project ensures that the models used to simulate the physical world can be verified against textbook problems and analytic sources. The V&V project must also validate codes against prior experimental data to ensure that the simulated physics agrees with known experiments. This information depends on other data input to the software, which correspond to how physicists set up a problem space (or mesh); define the materials; apply various physics functionality, such as hydro and thermal dynamics, incompressible flow, and advection; and select material models.

The goal of the ASC V&V Software Quality Engineering (SQE) team is to ensure that the ASC codes are of adequate quality to perform their intended mission. To achieve this goal, highly skilled software quality engineers are embedded with the code teams to monitor



Functional equivalency testing allows developers to compare some ASC codes such as ALE3D with commercial applications such as LS-DYNA.

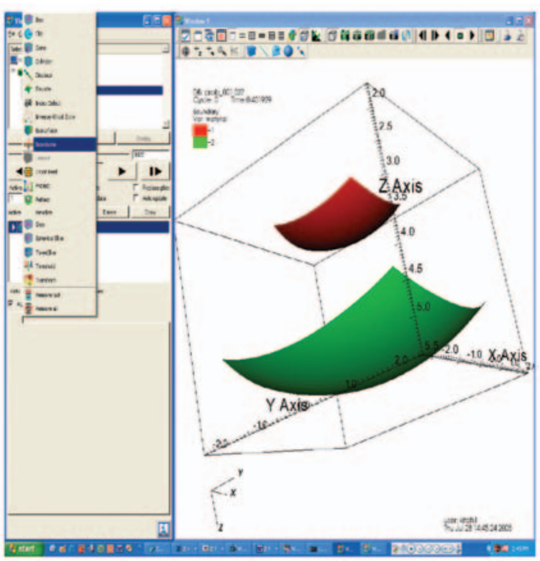
compliance with the established quality assurance, configuration-management, and testing requirements as well as with LLNL and ASC contractual requirements. The engineers also help implement quality improvements, such as issue tracking tools, test case improvements, and test coverage measurements, and they create the regression tests, user surveys, reports, and metrics required by the code teams.

Progress in 2005

This past year was an important one for the SQE team because NNSA audited the ASC software to ensure that improvements recommended in 2002 had been made. The auditors received updated plans on the activities of the code teams for software quality, risk assessment, software testing, software configuration management, and other topics such as disaster recovery and process improvement. Code teams completed a risk analysis in compliance with the new LLNL Institutional Software Quality Assurance Implementation Plan.

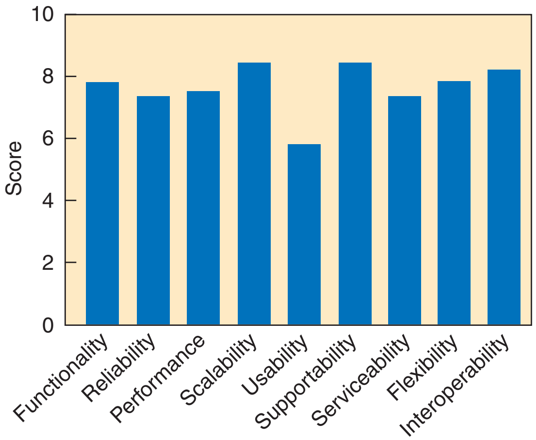
Best-practices forums for the ASC code team leaders, which began in late 2004, continued through 2005, allowing the code teams to share effective practices and warn about ineffective ones. As a result of the forums, improvements to the configuration-management and build processes have evolved, and new tools are being implemented, such as Tapestry for test management and Purcov for test coverage measurement. The teams also established software requirements for third-party suppliers.

Measuring the test coverage of statements and functions while executing hundreds of individual regression tests allows the code teams to



Graphical user interface for the output tool VisIt.

determine which tests are most efficient. Using this information and other improvements in platform speeds and parallelism, code teams were able to cut automated regression testing time by 30 percent or more while maintaining or improving test coverage. The LLNL-developed Tapestry test tool provides test management and reporting functions and checks for expected results during the entire test run instead of a discrete time slice. Test coverage analysis also allowed the engineers to modify test problems so they could examine areas of code that previous tests had missed. With metrics on code coverage over time, code teams can dynamically track net test coverage as new code is added, allowing them to determine where additional testing is required. In addition, functional equivalence testing added in 2005 allows developers to compare the results of some LLNL ASC codes to commercial modeling codes. (See the top figure on p. 30.)



Results from a 2005 survey of ASC users.

Another test capability added was randomized automated testing to the ALE3D code using the commercial testing tool WinRunner. WinRunner can process test problems with randomized values for selected inputs for hundreds of hours, allowing a team to exercise a code over a wide range of data values. WinRunner is also used to demonstrate code reliability and to refine reliability predictions using statistical methods such as the Musa–Everett model. The graphical user interface for the input tool Cyclops and the output display VisIt (bottom figure on p. 30) were both used as part of the automated randomized testing with WinRunner.

User guide traceability tests were piloted in 2005. This technique develops test cases directly from the user guide to exercise specific functions of the code, such as input block parsing and syntactical rules.

In 2005, the SQE team began helping to expand the regression test suites in third-party codes such as LEOS and PMesh, making it less likely that defects in third-party codes will affect the simulation codes that use them.

Defect tracking tools were added or updated to ASC codes. Surveys conducted in 2005 asked users to rate various quality factors. As shown in the figure below left, the survey results underscored the user approval of ASC codes. However, they indicated that users needed more information on the codes. As a result, user classes were scheduled in 2005, and updated software documentation was provided.

Significance

The tasks pursued in 2005 by the ASC V&V SQE team increased the assurance of ASC software being developed and allowed code developers to take advantage of new scientific insights. Promising pilot studies demonstrated new software technologies and techniques for improving the codes. In 2006, the SQE team will build on these accomplishments, working directly with the code teams to improve software quality assurance.

Contact Information

Greg Pope, pope12@llnl.gov

4.03 — Computational Support for National Biodefense

Overview

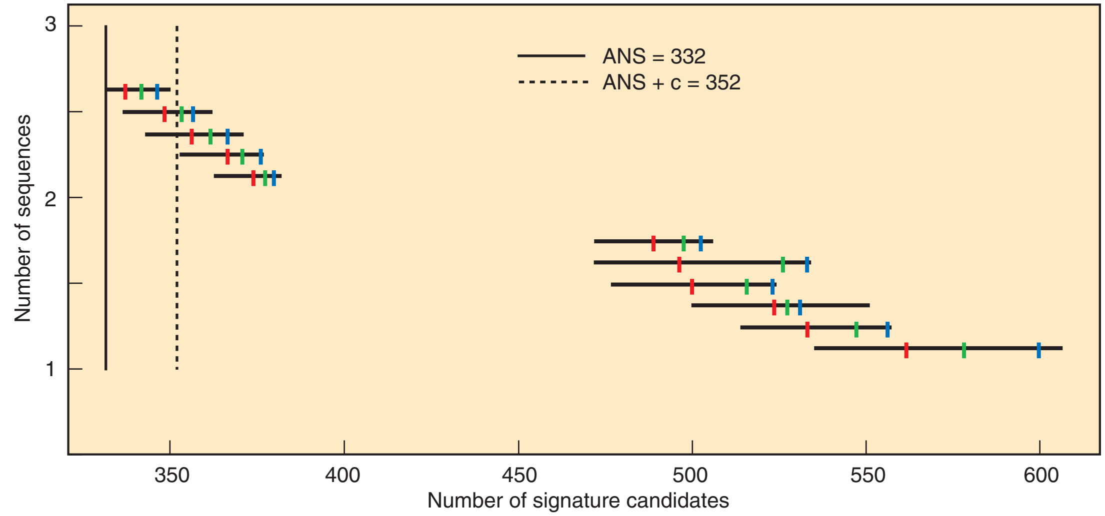
The Pathogen Bioinformatics Group (PBG) at LLNL helps the nation prepare against the threat of biological outbreaks, whether of natural or nefarious origin. This effort involves computational sequence analyses of DNA and proteins to predict signature regions for the detection and diagnosis of human and agricultural infectious threats as well as bioforensic discrimination among strains. The PBG has a lead role in helping DHS plan for FY06. The group also works in multiple capacities to support the national BioWatch system, which samples the air in major U.S. metropolitan areas for biothreat pathogens. The PBG has taken a multi-pronged approach to bioinformatics research and development. The following section describes some of the group's diverse projects.

Progress in 2005

The PBG is designing a virulence detection chip, a microarray to detect genes that confer virulence and antibiotic resistance, and received new funding from DHS to continue the development effort. The chip is being designed to detect hundreds of genes and all known homologues of those genes, numbering in the hundreds of thousands of sequence variants. The PBG is building a database of known virulence and antibiotic resistance genes, and to date, it contains more than 800 gene families. Computational techniques used for chip design include sensitive Hidden Markov Model searches running on the LLNL Thunder supercomputer followed by combinatorial methods to select the best minimal sets of conserved oligonucleotides required for detection.

The U.S. Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, provides LLNL access to more than 100 unpublished pathogen genomes. The PBG performs analyses on a number of pathogens at the request of the CDC, including avian influenza, variola (the causative agent of smallpox), *Brucella*, Rift Valley fever, and the coronavirus SARS-Cov (the etiological agent for severe acute respiratory syndrome). Analyses range from predicting peptide targets for vaccine development and modeling the three-dimensional (3D) structure of protein variations between virulent and avirulent strains, to designing signatures for detection and forensic discrimination of pathogen genomes.

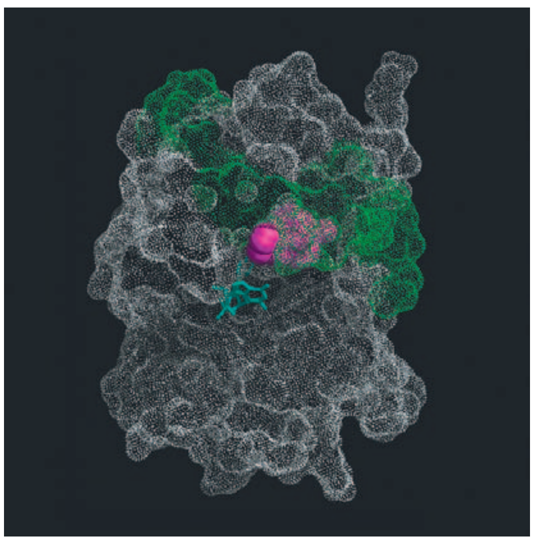
Because of the success of LLNL's collaboration with CDC, a similar collaborative agreement has been arranged between LLNL and DHS Plum Island, the nation's lead research laboratory for high-threat agricultural pathogens. In 2005, the PBG provided Plum Island with bioinformatics analyses of organisms such as foot-and-mouth disease virus. Moreover, the PBG is working with commercial and academic organizations funded by the DHS Homeland Security Advanced Research Project Agency to develop signatures compatible with a variety of competing assay platforms for use in comparative testing. The PBG also assists on several Laboratory Directed Research and Development projects, for example, developing a panel to detect and discriminate respiratory viruses.



A range plot shows that sequence data from *Clostridium perfringens* is critical in determining conserved regions suitable for diagnostic signatures and that more target sequencing is recommended. Genomes of close phylogenetic relatives are also important but contribute less to narrowing the pool of signatures to be screened in the laboratory.

Work continues to update assessments on the number of genome sequences required to generate high-quality diagnostics, as requested by an intelligence community sponsor. Using Monte Carlo simulations and the Signature Pipeline, the Sequencing Analysis Pipeline reevaluates the need for additional investments in sequencing dangerous pathogens and their innocuous close relatives based on newly sequenced genomes. (See the figure on p. 32.) Recommendations based on the PBG's computational analyses are used to set national sequencing priorities for the CDC Category A–C threat list of pathogens.

The PBG collaborates with researchers from several national laboratories to develop reagents for protein-based detection of toxins and pathogenic organisms. This past year, the PBG designed and built a protein signature evaluation software system to rapidly identify protein targets. The system combines advanced methods in modeling the 3D structure of proteins with computational methods for identifying optimal surface features to guide reagent development. (See the figure below left.)



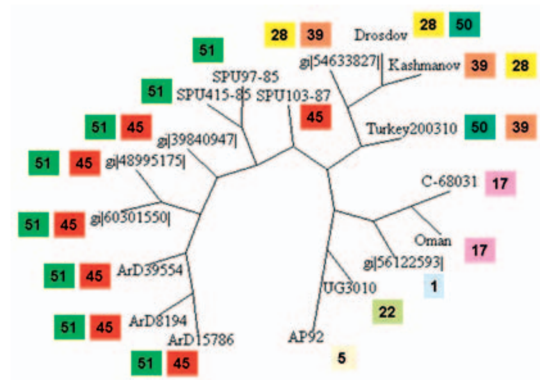
The 3D protein structure of ricin-A chain shows the optimal binding site (magenta spheres) for directing detection reagent development.

In 2005, the PBG enhanced the KPATH signature system, which checks for uniqueness of signature regions against a database that, as of November 2005, contained about 900 bacterial and 16,000 viral complete genomes. Diagnostic signature prediction using the KPATH Signature Pipeline for a bacterial genome requires less than 2 hours, while that for a virus takes less than 30 minutes. This turnaround time provides the capacity for a rapid response to new sequence data. New tools developed by the PBG facilitate rapid in silico screening. With this screening, users can examine whether signatures are likely to yield false-positive or false-negative results by comparing them against the National Center for Biotechnology Information GenBank nucleotide sequence database, which contained 15.4 billion bases as of November 2005.

The PBG continued to work with collaborators to computationally evaluate the specificity and reliability of signatures developed elsewhere. In addition, the group designed software to handle signature prediction for highly divergent viral genomes, such as many of the viral hemorrhagic fevers and influenza. For such viruses, no signature regions are conserved across all strains. Combinatoric algorithms yield minimal signature sets that in combination should detect all strains or distinguish among serotypes. (See the figure below.)

Significance

DHS and other agencies recognize the PBG as providing a national core capability for biodefense. The group's ability to apply the best available computational techniques strengthens the nation's biosecurity against the threat of natural or terrorist release of dangerous pathogens.



Diagnostic signature clusters (colored numbers) mapped onto a phylogenetic tree for Crimean Congo hemorrhagic fever show the concordance between evolutionary branches and shared signatures for detection.

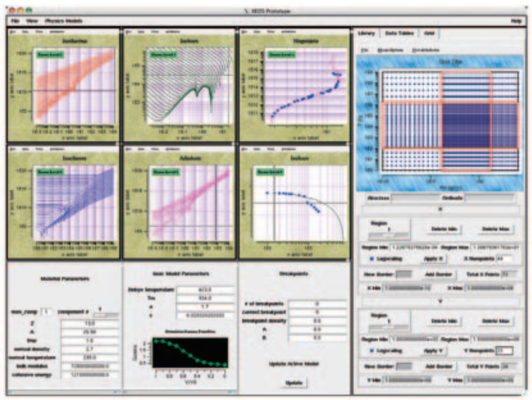
Contact Information

Shea Gardner, gardner26@llnl.gov
Tom Slezak, slezak1@llnl.gov

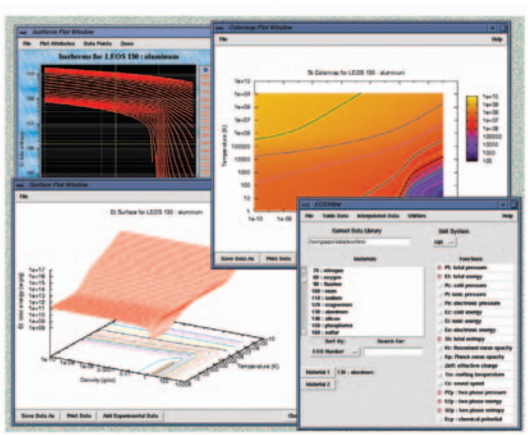
4.04 — Next-Generation Equation-of-State Data

Overview

An equation of state (EOS) describes a material by the relationships among its thermodynamic quantities. Many applications require accurate EOS data for extreme states of matter, and LLNL scientists often use the Livermore Equation of State (LEOS) system to obtain EOS data.



The XEOS user interface.



The EOSView equation-of-state browser application.

A new electronic structure program, called Purgatorio, produces improved data for the electronic component of an EOS. XEOS, a new global EOS program, constructs a complete EOS based on the Purgatorio electronic component, along with ionic and cold-curve components. A prototype EOS data library has been generated with XEOS. The EOS data in this library can be used with the LEOS access code and EOSView browser and are available for use in the Laboratory's hydrodynamics codes.

Progress in 2005

This year, both Purgatorio and XEOS matured from experimental prototypes to sophisticated, near-production-level resources. These codes produced EOS data for several materials, making available a new LEOS-format data library.

Purgatorio is a numerical implementation of the earlier Inferno EOS model with two major innovations: a phase-amplitude method, for accurately calculating radial wave functions at high energies; and a novel automatic quadrature scheme, for resolving the structure of the continuum density of states and sharp resonances. Purgatorio treats continuum and bound electrons in the same manner, and the solution is fully relativistic.

XEOS is based on a model that separates an EOS into electronic, ionic, and cold-curve components. The Purgatorio electronic component is a major upgrade over the Thomas–Fermi model used previously. The other components are provided by physics models implemented in XEOS itself. The graphical user interface for XEOS allows a user to construct an EOS interactively. A user can modify parameters and view the results of these changes in real time on several plots, including comparisons with experimental data. The final results are written directly to a LEOS-format database.

The LEOS system includes a set of software functions to read and write data from the LEOS database. This software was extensively tested in 2005, resulting in a more robust, reliable software library with which to access EOS data.

Significance

The quality of simulation results depends on the quality of the input data. Because EOS data are a fundamental input to hydrodynamics code simulations, data accuracy is a key component for producing accurate simulations. In addition, many LLNL researchers are interested in accurate EOS data. For these reasons, the development of Purgatorio and XEOS represents a substantial contribution to many efforts at the Laboratory.

Contact Information

Greg Streletz, streletz1@llnl.gov
David Young, young5@llnl.gov

4.05 — New Software Tools Improve Seismic Research

Overview

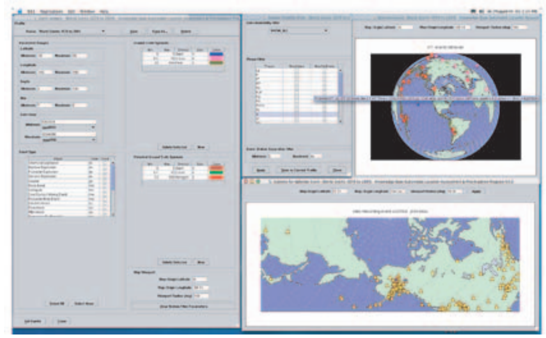
The Ground-Based Nuclear Explosion Monitoring Research and Engineering program at LLNL, which is funded by NNSA, made significant progress this past year. We enhanced the process of deriving seismic calibrations and performing scientific integration, analysis, and information management with software automation tools. We also completed several achievements in schema design, data visualization, synthesis, and analysis. Our tool efforts address the issues of very large data sets and the varied formats used during seismic calibration research. Because data volumes have increased, scientific information management issues, such as data quality assessment, ontology mapping, and metadata collection—all of which are essential for producing and validating derived calibrations—can hinder the effectiveness of researchers.

Progress in 2005

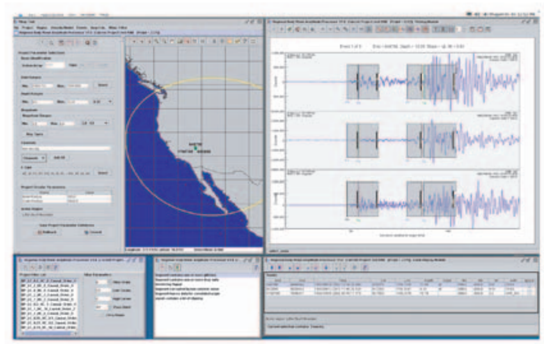
Our software development efforts produced an object-oriented database framework that provides database-centric coordination between scientific tools, users, and data. Nearly 500 million parameters, signals, measurements, and metadata entries are stored in a relational database that can be accessed by an extensive object-oriented multitechnology software framework. This framework includes elements of schema design, stored procedures, real-time transactional database triggers, constraints, and coupled Java and C++ software libraries to handle the information interchange and validation requirements. A core feature is the ability to rapidly select and present subsets of related signals and measurements to the researchers for analysis and distillation both visually (Java graphical user interface [GUI] client applications) and in batch mode (instantiation of multithreaded applications on clusters of processors).

The first of two information management and analysis tools is the Knowledge Base Automated Location Assessment and Prioritization (KBALAP) program (top figure at right). KBALAP is a set of database services and GUI applications that combine to produce location ground-truth data critical for nuclear explosion monitoring research. KBALAP can also evaluate raw signals and relevant measurements as they enter the system to identify what, if any, further analysis or data are required to improve accuracy. Without software automation and project management capabilities, these tasks would not be feasible, given the volumes of available data and the required analysis tasks.

The Regional Body-Wave Amplitude Processor (RBAP) program (bottom figure at right) automates the process of measuring the amplitude of regional seismic phases to improve seismic-event characterization and identification at each station. Compared with previous tools, RBAP is faster and easier to use, and it scales to a larger number of events. It also permits project revision and updating through the database by tracking new data entering the system. RBAP allows researchers to recalculate derived calibrations using different processing assumptions. In addition, ground-truth data produced within the KBALAP tool are instantly available to researchers using the RBAP tool because both access the same database for parameter management and coordination.



The Knowledge Base Automated Location Assessment and Prioritization program is a set of database services and a Java client application that combine to produce location ground-truth data for use in producing travel-time correction surfaces.



The Regional Body-Wave Amplitude Processor helps to automate the process of measuring the amplitude of regional seismic phases. These measurements can then be used to calibrate seismic discriminants at each station.

Significance

The new information management and analysis tools improve efficiency in producing scientific data products and the accuracy of derived seismic calibrations, which are essential to LLNL's research on nuclear explosion monitoring. These tools also provide a synergistic framework for further software development and have produced one of the largest integrated, quality-controlled seismic databases for seismic research.

Contact Information

Stanley Ruppert, ruppert1@llnl.gov

RETURN TO TABLE OF CONTENTS

4.06 — Simulations and Modeling for the Linac Coherent Light Source

Overview

When the LCLS becomes operational in 2009, it will be the world's first x-ray free-electron laser (FEL). This dramatically new x-ray source, which is being constructed at SLAC, will have exceptionally bright, coherent, short, and energetic x-ray pulses with wavelengths of 1.5 to 15 angstroms. Its fast "shutter" and brilliance will enable new scientific research, such as investigating new states of matter, understanding and following chemical reactions and biological processes in real time, and imaging nanoscale and noncrystalline biological materials.

Funded by the DOE's Office of Science, the LCLS effort involves a collaboration of national laboratories and universities. SLAC is responsible for the linear accelerator, Argonne National Laboratory is providing the magnetic undulator system, and LLNL is responsible for the x-ray beam transport, optics, and diagnostics (XTOD). UCLA is providing theoretical and numerical predictions.

Progress in 2005

Computation members of the XTOD team work in two areas: modeling and controls. The major challenge for the team is to design and build components and optics that can only be tested in the LCLS. To support this design process, we develop the applications needed to accurately model the beam and components.

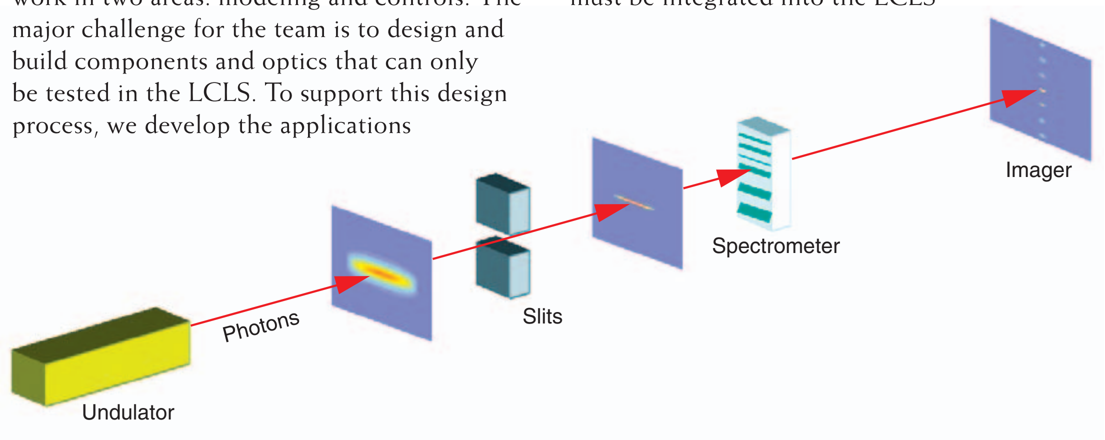
In 2005, the XTOD team began designing and prototyping each component. The team is responsible for delivering the vacuum system, slits, mirrors, attenuators, imagers, and other diagnostic instruments needed to characterize the x-ray beam. The diagnostic instruments will be used during commissioning to measure the intensity of the x-ray pulse and its spatial, temporal, and photon-spectral properties.

We made significant progress developing a suite of modeling and simulation applications to propagate both the spontaneous and FEL photons down the beamline and through each component. Detailed radiation flux data sets, provided by UCLA, are projected back to the undulator. Monte Carlo codes then track individual photons through pipes, masks, slits, attenuators, and spectrometers. We use the flux field at specific locations to predict detection thresholds for each instrument. Our calculations help determine appropriate sizes and materials for the various components.

The beam transport, optics, and diagnostics must be integrated into the LCLS global control system, which is based on the Experimental Physics and Industrial Control System (EPICS) architecture. This year, we concentrated on identifying the number of control points (about 4,000), determining a naming convention, and defining the configuration databases. We also selected commercially available hardware and estimated the cost and effort to integrate the XTOD components into the global system.

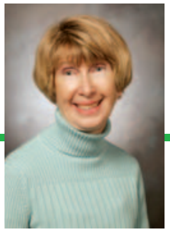
Significance

When commissioning begins in 2007, the LLNL x-ray diagnostic instruments will provide the first answers to the questions concerning the overall performance of LCLS. Instrument design will be based on modeling and simulations. The successful transition of LCLS from its construction phase into an operational facility will depend on how well we can integrate the XTOD components into the EPICS control system. After commissioning, LCLS will provide important data for several LLNL projects, in particular, research on biomolecular structure and warm, dense matter.



A model of LCLS spontaneous radiation from the undulator through the off-axis spectrometer.

Contact Information



Linda Ott, ott2@llnl.gov
Kirby Fong, fong5@llnl.gov
Steve Lewis, lewis84@llnl.gov

LCLS Web site: www-ssrl.slac.stanford.edu/lcls

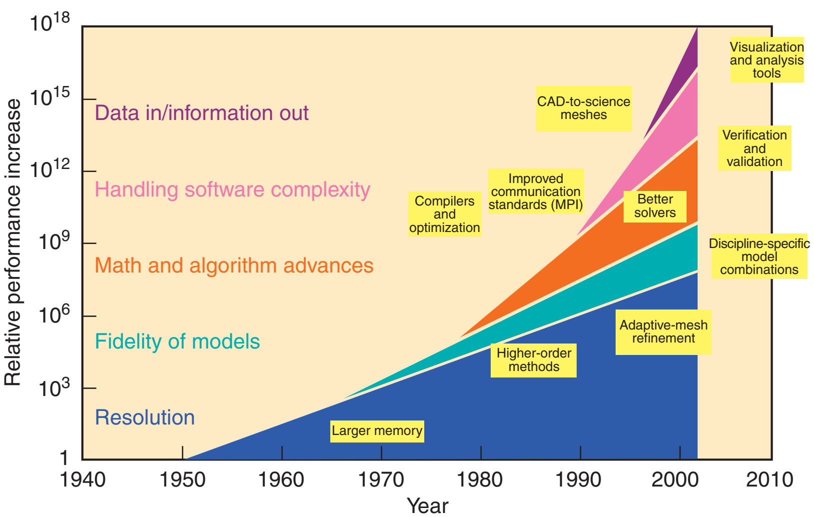
SECTION 5

Research and Advanced Development

5.00 — Computing Research Drives Predictive Capability

Many programs at Lawrence Livermore National Laboratory (LLNL) have extreme computing demands, which drive the research activities in the Computation Directorate. The directorate's research, in turn, enables use of massively parallel computing to solve problems of national interest. The recent increases in hardware capability at LLNL are applied in concert with equally important computational science advances. The figure below gives a high-level historical view of research areas that contribute to LLNL's current simulation capabilities.

Especially noteworthy for 2005 is the leadership and participation of Computation personnel in the BlueGene/L (BG/L) science calculations, which culminated in three finalists for the 2005 Gordon Bell Prize. The winning team—which included scientists from Computation, LLNL's Physics and Advanced Technologies (PAT) Directorate, and IBM—achieved more than 100-trillion floating-point operations per second (TF) on a molecular-dynamics (MD) problem applied to solidification in metals. A second team attacked a difficult MD problem using the full quantum-mechanical formulation and achieved 64 TF on BG/L. The third team achieved the first known search of a massive graph with more than 30 billion connections, proving that gloomy predictions of restricted scaling performance for such problems on BG/L were unwarranted.

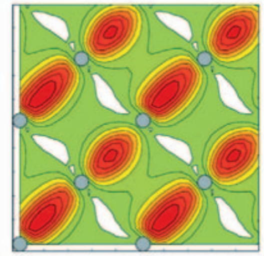


The relative simulation performance increases as a result of computational science research are plotted versus time. Although the numerical factors for the five research areas shown are qualitative, they represent large classes of scientific modeling efforts at LLNL and elsewhere.

Several topics in this section focus on the use of computations to address problems that deal with data management and the need to understand information derived from different sources. The Computation Directorate continues to advance our ability to simulate physical systems. At the same time, we are adding capabilities to address data science challenges central to 21st century issues. Both areas are represented in this section.

Our BG/L simulations of metals featured two MD applications that presented different challenges to scientists. In the first effort, which involved classical MD, interactions are formulated in terms of local forces. The second effort used quantum mechanics to describe the electronic structure that lies at the heart of metallic property characterizations.

An essential characteristic of metals is that some electrons can freely move around the material. The classical approach describes the approximate behavior of the electrons using local interaction potentials. Because of locality, the parallel treatment of the problem is not encountered at the core of the physical formulation. The problem thus becomes one of marshalling the computational resources available. The first team accomplished this tour de force with attention to all conceivable inefficiencies in the parallel implementation of the domain decomposition molecular-dynamics (ddcMD) code. Not only did this calculation achieve a world record 100 TF in an MD simulation, but it also provided new understanding on metallic solidification.



In metals, conduction electrons interact with the atomic latticework of other electrons and ions that form the bulk of a material. A completely accurate quantum calculation involves a more intricate computation than does the classical formulation.

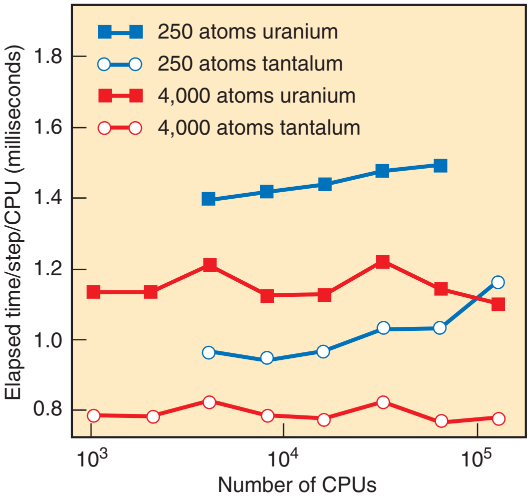
The second team dealt with overlapping (nonlocal) electronic structures, which complicate the formulation. This calculation essentially tracked every conducting electron throughout the entire material sample being simulated. (See the figure above.) Such detail requires a more complex solution process than is used in classical MD, which restricts the solution space to many fewer than the millions of electrons that must be treated for an accurate calculation. The 64 TF achieved on this extremely difficult problem was unprecedented.

5.01 — Classical Molecular-Dynamics Simulation of Material Solidification

Overview

Understanding the properties of matter under extreme conditions is fundamental to researchers in astrophysics, planetary science, and nuclear physics. Because experiments are often impracticable at these pressures and temperatures, scientists rely on computer modeling to understand the processes.

The ddcMD code is an MD simulation developed by Fred Streitz and Jim Glosli of LLNL's PAT Directorate, with tuning and system support provided by the Computation Directorate and IBM. The particle-based domain decomposition scheme used in the ddcMD code allows processors to compute potentials for overlapping spatial regions. This approach results in better scaling compared to traditional geometry-based schemes and enables ddcMD to achieve unprecedented performance on BG/L.



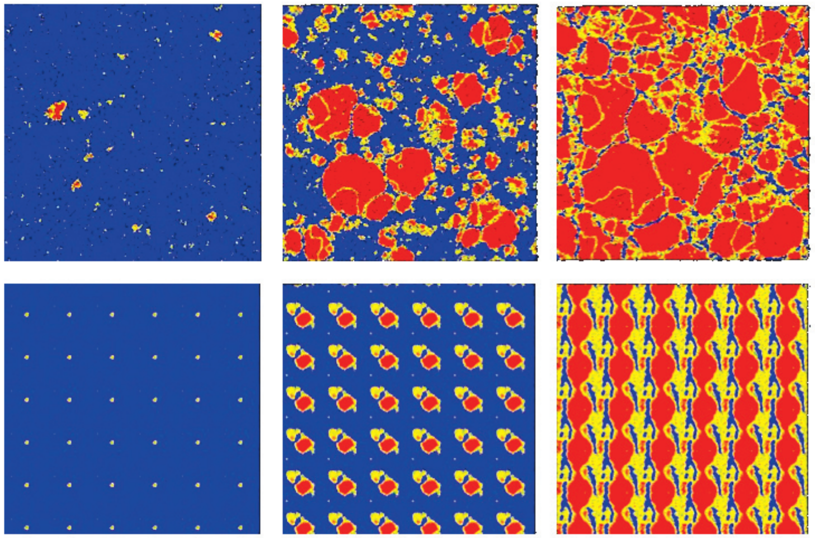
Excellent scaling of ddcMD on BG/L. CPU = central processing unit.

Progress in 2005

This was a year of tremendous progress for both BG/L and ddcMD. In 2005, BG/L doubled in size twice to its final configuration of 65,536 nodes and was integrated into the Terascale Simulation Facility (TSF). As a result of intensive testing and design discussions with IBM, the system software (operating systems, compilers, message-passing library, scientific libraries, and performance-analysis tools) improved significantly in reliability and performance. The ddcMD team worked closely with BG/L experts and other applications teams.

In 2005, the ddcMD team performed scaling studies and tuned the code for higher performance. Single-node operation was improved by hand tuning numeric kernels that perform combined matrix multiplication and dot products of small matrices. This approach to scalability via particle-based decomposition proved to be excellent for both "strong" and "weak" scaling. (See the figure at left.) These results were achieved even as system software was being updated.

The ddcMD simulations on BG/L in 2005 modeled the pressure-induced solidification of molten tantalum and investigated solidification in quenched uranium. The tantalum simulations yielded the first determination of the minimum size needed for model solidification, from the rapid nucleation and growth phase during which individual grains are formed to the coarsening phase during which the metal solidifies into a characteristic microstructure. (See the figure above right.) We expect the uranium simulations, which are ongoing, to reach solidification in early 2006.



Cross sections of simulations using (top row) 16 million atoms and (bottom row) 64,000 atoms taken at equivalent times during solidification.

Significance

The ddcMD results are a historic achievement, and the LLNL–IBM team received the 2005 Gordon Bell Prize for this accomplishment. The tantalum calculations provide the first model of metal solidification without approximations because of system size. The uranium calculations have already achieved more than 100 TF. In 2006, we expect to find nucleation and growth in this important actinide, allowing direct comparison to a transition metal system.

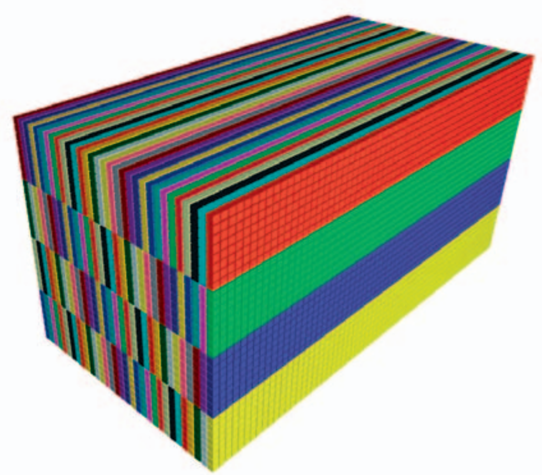
Contact Information

BG/L system software: Kim Yates, kimyates@llnl.gov
BG/L project: Kim Cupps, kimcupps@llnl.gov
ddcMD application: Fred Streitz, streitz@llnl.gov

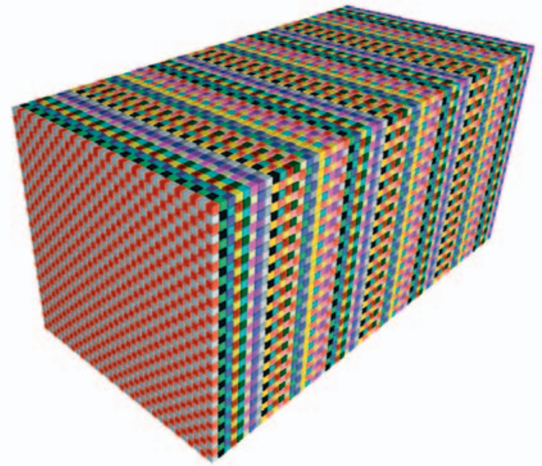
5.02 — First-Principles Molecular-Dynamics Simulations on BG/L Using Qbox

Overview

Qbox is a first-principles molecular-dynamics (FPMD) code developed at the Center for Applied Scientific Computing to enable the next generation of materials simulations relevant to stockpile stewardship. We present the results from porting Qbox to BG/L and optimizing the code's performance to take advantage of the supercomputer's full 65,536 nodes. The FPMD method is an accurate, atomistic simulation approach that uses a quantum mechanical description of electrons. Although computationally expensive, FPMD enables scientists to perform predictive materials simulations in which no empirical or adjustable parameters describe a given system of atoms.



The default node mapping for the full 65,536-node BG/L 3D torus interconnect. Parts of the calculation are done in parallel, with communication both within and between subcommunicators in the message-passing interface. Colors represent the different subcommunicators. With this mapping, the peak performance was 39.5 TF.



An improved node mapping for the 1,000-atom molybdenum system on 65,536 nodes. This mapping resulted in a peak performance of 64 TF, an improvement of over 60 percent from the default mapping shown in the figure above.

Progress in 2005

Predicting the properties of metals under extreme conditions of temperature and pressure has been a long-standing goal for researchers in materials science and high-energy-density physics. Because of the high computational cost of FPMD, tractable system sizes are typically limited to less than a few hundred atoms on most currently available supercomputers. Under extreme conditions, some important properties of metals require simulations to use a significantly larger number of atoms than has been feasible on previous systems. The calculation of melting temperatures using the two-phase simulation technique and the calculation of defect energies and defect migration processes are two examples.

Qbox simulations of 1,000 molybdenum atoms (12,000 electrons) showed that a strong-scaling parallel efficiency of 86 percent can be obtained between 1,024 and 32,768 processors. The same problem runs more than 27 times faster on 32,000 nodes than on 1,000 nodes. Using 65,536 processors, we observed sustained performance of up to 64 TF, indicating efficient use of the machine.

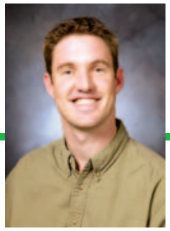
Qbox was designed specifically to take advantage of LLNL's large parallel computers. An efficient data layout achieves good load balance, and the data flow is carefully managed during the most time-consuming operations. In addition, choosing an efficient mapping of tasks to processors on BG/L's three-dimensional (3D) torus at run-time interconnect minimizes communication time and obtains good performance.

For example, using the default node mapping for 65,536 nodes (top figure at left) resulted in sustained performance of 39.5 TF. By exploring a range of mapping strategies, we found that the same calculation could be mapped onto the machine to achieve a sustained performance of 64 TF (bottom figure), an increase of over 60 percent.

Significance

The ability to perform predictive simulations of metals under extreme conditions is of great importance to both stockpile stewardship and computational materials science. These results represent unprecedented scaling and performance for an FPMD calculation. By enabling scientists to study systems that are considerably larger than was previously feasible, Qbox running on BG/L paves the way for the next generation of predictive materials simulations at LLNL.

Contact Information



Erik Draeger, draeger1@llnl.gov
François Gygi, gygi2@llnl.gov

RETURN TO TABLE OF CONTENTS

5.03 — Scalable Parallel Performance for Complex Graph Analysis on BG/L

Overview

Search capabilities play an important role in analyzing the large, complex sets of data that comprise graphs. Breadth-first search (BFS) is of particular importance because it is used in numerous applications to answer user queries. Searching graphs with billions of vertices and edges poses challenges and calls for a distributed, parallel BFS algorithm.

However, the scalability of the distributed BFS algorithm for very large graphs becomes a critical issue because the demand for local memory and interprocessor communication increases as graph sizes increase. We are developing a scalable and efficient distributed BFS scheme that can handle graphs with billions of vertices and edges. IBM's BG/L, with its large memory and fast interconnect, provides us with an ideal platform for scalability testing.

Progress in 2005

We achieve high scalability through a series of innovative optimization techniques. First, we use two-dimensional (2D) edge partitioning instead of the more conventional one-dimensional (1D) vertex partitioning to partition graphs for parallel processing. In 2D edge partitioning, a level expansion of the BFS is achieved by performing column- and row-wise communications. With 2D partitioning, the number of processes (P) involved in the collective communications is $O(P^{1/2})$ in contrast to O(P) with 1D partitioning. Hence, with 2D partitioning, we can reduce the communication time significantly.

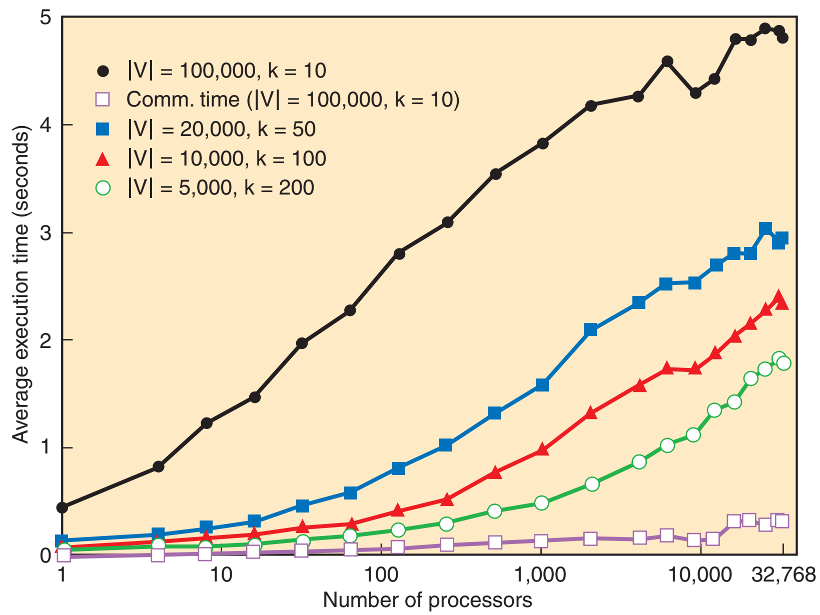
In our initial experiments on BG/L, we identified message buffers used in collective communications as a bottleneck for scalability because their size increases with the number of processors. We learned that for a class of graphs called Poisson random graphs, the expected message length is bounded by the local problem size. Using this information, we can control the buffer sizes and, hence, relieve the memory bottleneck. This optimization allows us to manage the memory more efficiently and improve scalability.

We also optimized the ring-based collective communications to take advantage of BG/L's 3D torus interconnect, which parallelizes much of the ring communications. For this implementation, we explored the use of "reduce scatter" to reduce message volume.

When we tested the BFS algorithm on 32,000 nodes of BG/L, we found that it exhibited good scalability in terms of the total problem size. It searched a random graph with 3.2 billion vertices and 32 billion edges in just 4.9 seconds. (See the figure at right.) To the best of our knowledge, this is the largest explicitly formed graph ever handled by a graph-search algorithm.

Significance

This work demonstrated our capabilities to search very large, complex graphs with billions of vertices and edges, and it helped us identify new approaches to scale to even bigger graphs. Our results provide us with insight on other types of scalable graph-search algorithms that can be developed in the future. This work was selected as a finalist for the 2005 Gordon Bell Prize.



Weak scaling results of the distributed BFS on 32,768-node BG/L system. |V| = number of vertices assigned to each processor, and k = their average degree.

Contact Information

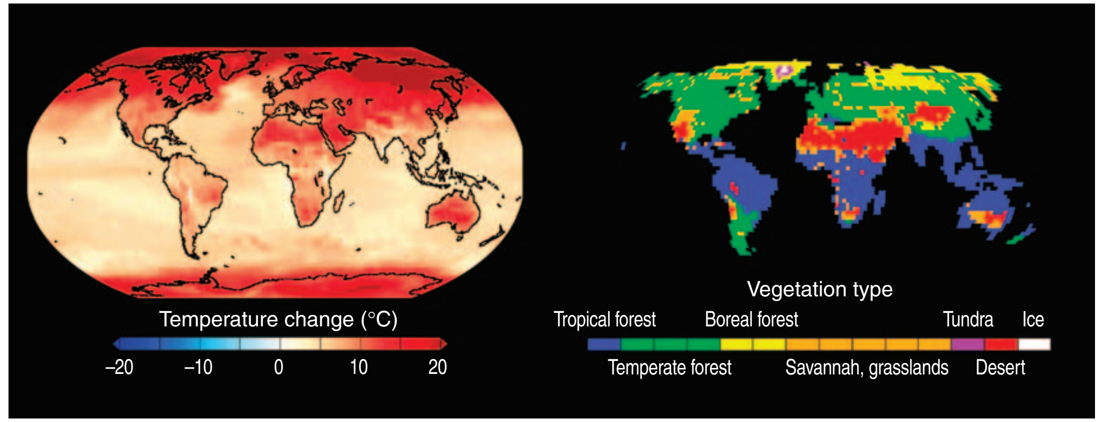
Andy Yoo, ayoo@llnl.gov

5.04 — Long-Term Climate Simulations

Overview

We developed a 3D climate model that can predict atmospheric conditions by projecting future greenhouse gas emissions based on quantities dictated by current energy policies. Using these forecasts, the model can then determine how much the climate will change over time. This approach requires incorporating detailed treatments of the carbon cycle and other biogeochemical cycles. Researchers in the Center for Applied Scientific Computing (CASC) and the Energy and Environment (E&E) Directorate developed the comprehensive Integrated Climate and Carbon (INCCA) model, the first of its type in the U.S.

The INCCA model is being used to evaluate the climatic impact of proposed scenarios for greenhouse gas emissions, carbon sequestration, and policy-relevant land-use changes. Previous studies with INCCA evaluated how limited carbon dioxide (CO_2) fertilization will affect global climate change and how climate sensitivity will affect carbon-cycle feedback.



Results from LLNL's model (left) show the predicted climate in the year 2300 assuming the continued use of all available fossil fuel resources. The change in global mean surface temperature, which was 0.8°C in 2000, will be 7.8°C by 2300. Land areas will warm more than the oceans, and Arctic and Antarctic regions will warm more than the tropics. The simulated global vegetation distribution (right) shows the poleward expansion of tropical and temperate forests as well as the poleward migration of boreal forests over time. Tundra and land ice will nearly vanish by the year 2300.

Progress in 2005

In 2005, we used the INCCA model to simulate the potential global climate change over several centuries that would result with release of CO_2 emissions from all the currently estimated fossil fuel resources. The model predicts that if humans continue to use fossil fuels at this rate for the next few centuries, the polar ice caps will be depleted, ocean sea levels will rise by 7 meters, and median air temperatures will be 8°C warmer than present day. The model predicts that by the year 2300, the level of atmospheric CO_2 would nearly quadruple from 380 parts per million (ppm) to 1,423 ppm. The simulated warming also indicated that species of trees now growing in tropical and temperate latitudes will migrate toward the poles and global forest cover will expand by more than 40 percent.

Land-management practices can help mitigate climate change. Our climate models allow us to examine the consequences of such changes before they are deployed on a large scale. For example, forest regrowth is a proposed method for storing carbon. However, the dark forest canopy absorbs solar radiation, and our simulations showed that replacing current vegetation with trees leads to a global mean warming of 1.3°C. The simulations indicate that planting trees across the U.S. and Europe to absorb CO_2 may not be effective at mitigating warming. To better understand these issues, we are simulating various land-use changes with the fully coupled INCCA model.

Significance

In 2006, CASC and E&E will work together to improve future climate predictions by applying the best available computational methods and computer resources to this problem. The collaboration of computational and climate scientists continues to put LLNL at the forefront of climate modeling.

Contact Information

Michael Wickett, wickett@llnl.gov

RETURN TO TABLE OF CONTENTS

5.05 — MESA Demonstrates Scaling to National Level

Overview

The Multiscale Epidemiological/Economic Simulation and Analysis (MESA) decision support system is a project funded by the Department of Homeland Security (DHS) to develop a coupled foot-and-mouth disease (FMD) epidemiological model and economic impact model. The national-scale system will provide an environment to explore various scenarios for the intentional introduction of FMD. Users will be able to assess the efficacy of response options and countermeasures, such as diagnostics, vaccines, and therapeutics, in controlling the extent and duration of simulated outbreaks.

MESA is an individual-based model, with each agricultural facility (such as a herd or stockyard) modeled as a discrete entity. The Java-based framework includes components for scenario definition and setup, data analysis, and visualization, including links to a geographic information system. MESA is designed to run on supercomputers or large clusters and can be extended to simulate other foreign animal diseases.

Progress in 2005

The initial effort focused on gathering requirements for the decision support system from key stakeholders in the veterinary and economic communities. Key gaps now being addressed in MESA include the following:

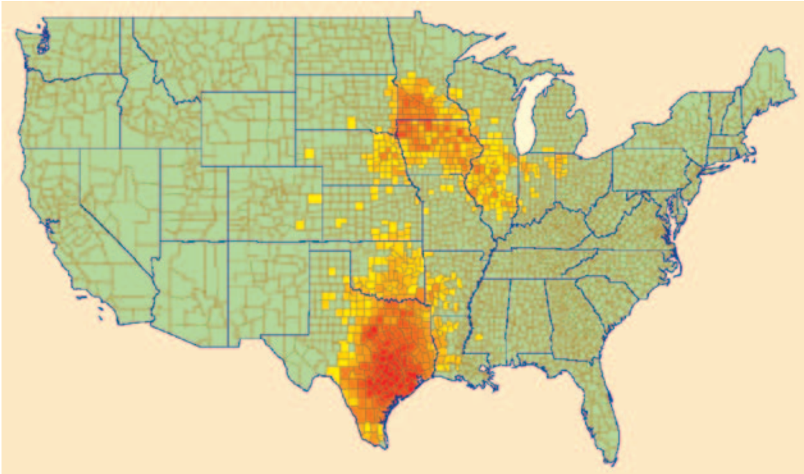
- Regionalizing the parameters because livestock industries vary geographically.
- Independently varying control measures between states because each state may operate independently.
- Varying temporal parameters to represent seasonal changes in facility behavior.
- Explicitly representing facilities that operate as part of a larger corporate structure to capture these integrated relationships.
- Scaling the model to simulate nationwide outbreaks.

A full economic and epidemiological model specification was written based on the identified gaps. The software architecture was engineered to enable extensibility in the future by using component design principles, and modeling and analysis components were composed inside a BeanShell scripting engine to provide dynamic and flexible scenario control.

In parallel with the formal design processes, a prototype system was developed to quickly investigate the basic behavior of the model. The notion of an aggregated facility representing a collection of facilities in a state or county was introduced to address scaling to national-level outbreaks. With this approach, interactions at long distances can be modeled, which significantly reduces the computational complexity of the problem. Nationwide outbreak scenarios were successfully run on 1,024 processors, demonstrating the viability of the approach. The team used the prototype system to run a set of scenarios as part of the DHS Harvest Resolve FMD tabletop exercise.

Significance

Foot-and-mouth disease is one of the most economically threatening livestock diseases in the world. Although it has not affected U.S. livestock since 1929, the disease is endemic in many areas around the world and is a constant threat to U.S. herds and flocks. Although FMD does not affect humans, its economic impact if reintroduced to the U.S. could be disastrous. A decision support system to estimate the effect of FMD occurring under various scenarios and its evolution following different response measures will enable decision makers to plan response strategies more effectively, thus potentially lessening the economic impact of an intentional introduction of the virus.



Example output summarizing 510 Monte Carlo MESA runs shows the likelihood of FMD spreading throughout the country after a herd in Texas is infected. Colors indicate the relative likelihood that a county would be infected, ranging from highest (red) to lowest (green).

Contact Information

Steven Smith, sgsmith@llnl.gov
Tom Bates, twbates@llnl.gov

5.06 — Scalable Information Extraction for Homeland Security

Overview

Our goal in this project is to build on existing natural language processing (NLP) tools so they will effectively work within real-world DHS programs. In particular, we are focusing on the biodefense domain, working closely with the Biodefense Knowledge Center (BKC), to develop scalable NLP technology. With these improved tools, BKC analysts can effectively monitor biological outbreaks and identify and track usage of bioweapons-relevant technologies.

To meet this goal, we developed a prototype text-analysis pipeline that ingests documents and inserts the relevant facts, relationships, and events into the BKC semantic graph based on the Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE) architecture. This pipeline identifies entities of interest in text (such as diseases and pathogens), determines important relationships between identified entities (such as a causal relationship between pathogen and disease), and extracts complex events (such as outbreaks, which include disease, victim, date, and location information) based on these relationships.

From: ProMED-mail
[Indonesia]: [Farm Worker] is [First] [Confirmed] [Human Case] of [Avian Influenza]

JAKARTA: A [farm worker] in [eastern Indonesia] has [tested positive] for [avian influenza virus], marking him the country's [first] [human case] of the virus. The worker from [southern Sulawesi] island is [healthy] but [two] [tests] at a [Hong Kong] [laboratory] [confirmed] that he [had been infected] by [avian influenza virus]. [Since 2003], the highly lethal disease has struck [chickens], [quail] and [other birds] in [18 Indonesian provinces] on [seven islands]. Researchers have also determined that [avian influenza virus] has [infected] [Indonesian pigs], an ominous development because swine can catch both avian and [mammalian] influenza viruses.

Excerpt from a ProMed article with interesting facts highlighted.

Progress in 2005

We participated in the DHS-sponsored Advanced Simulation and Computing (ASC) Text Analysis Workshop, where the goals and text challenges facing DHS were defined. During this workshop, we identified potential synergy among various DHS-funded projects and subsequently organized two highly successful, multi-institutional working group meetings. As a result of the meetings, all text-analysis efforts in the DHS Office of Research and Development have been standardized on IBM's open-source Unstructured Information Management Architecture (UIMA), which in turn resulted in a common development infrastructure across the program.

Our keyword-based parser, which we deployed in 2005, examines the performance demands of a system that processes millions of documents. We applied this technology to several BKC-relevant corpora, including a 15.5-million document set (PubMed corpus) that generated 835,000 relationships and a 30,000 document set (ProMed corpus) that generated 10,000 relationships.

In an initial survey of NLP systems, we also formed strategic partnerships with Nagiza Samatova's group at Oak Ridge National Laboratory and Ellen Riloff at the University of Utah to leverage their relevant work. Our Oak Ridge partnership focused on developing an entity-tagging system customized to BKC-specific data types. Through our Utah partnership, we extended the Sundance system, a leading research system in development for more than a decade, to improve its accuracy on BKC- and DHS-relevant corpora.

We combined the two systems with customized dictionaries, which were defined in collaboration with the BKC analysts, and with specialized components we developed. We then integrated them into UIMA and created an initial prototype of the text-analysis pipeline. This pipeline is being evaluated using outbreak event templates extracted from ProMed reports. Our future work will focus on improving the accuracy and scalability of this pipeline.

Significance

The intelligence community receives tens of thousands of free-text documents per day. Rapidly discovering potential terrorist threats within this large collection of documents is vital to the nation's homeland security. Meeting this grand challenge can only be accomplished using automated NLP techniques that selectively provide information to intelligence analysts. By leveraging existing efforts and developing scalable NLP technology that can be applied in real-world environments, this project is helping DHS to effectively use information contained within its large free-text corpora.

Contact Information

Terence Critchlow,
critchlow@llnl.gov
David Buttler, buttler1@llnl.gov
David Hysom, hysom1@llnl.gov

RETURN TO TABLE OF CONTENTS

5.07 — Multigrid Technology for Electromagnetics Systems

Overview

The Scalable Linear Solvers (SLS) project is developing parallel algorithms and software to solve large linear systems of equations that arise from physical phenomena with wavelike behavior. These models occur in many important electromagnetics (EM) applications at LLNL, such as magnetohydrodynamics (MHD), laser optics, and circuit design, and those at other laboratories, such as defense radar detection and medical imaging. New scalable linear solvers can greatly improve the ability to solve these problems by allowing codes to simulate models faster and with higher resolution than ever before.

The new solvers are being developed in the SLS *hypre* library, which is already used by several codes developed for the ASC program. Although existing solvers in the *hypre* library can be used with EM systems, they generally do not perform as well on EM systems as they have on diffusion systems. This degradation is due to a characteristic difference between diffusion and EM processes. Because the existing solvers were developed using intrinsic mathematical properties of diffusion, performance degradation for EM equations is expected.

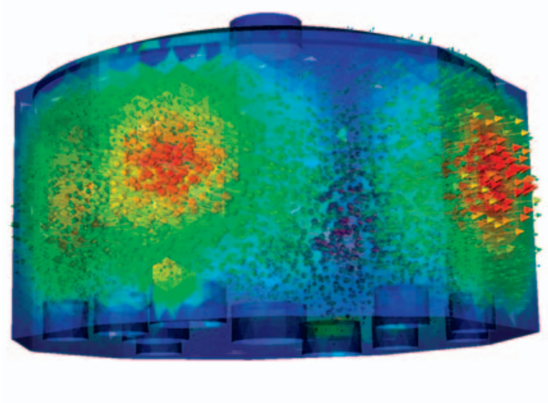
To obtain fast, scalable solvers for systems with wavelike behavior, such as those found in EM applications, the SLS project is developing multigrid technology that incorporates the mathematical properties of EM equations. The new technology dramatically improves solver performance and even resolves complicated EM models in highly varying material media where other solvers fail.

Progress in 2005

The SLS project developed several new multigrid algorithms for solving linear systems derived from EM models, in particular, models described by Maxwell's equations. The algorithms exploit basic properties of Maxwell's equations and their discretization, integrating them directly into the solvers. These properties are used to decompose the solution into different characteristic components, each of which is quickly resolved using different multigrid techniques. During 2005, the project team successfully developed technology to handle problems discretized on structured, semi-structured, and unstructured grids. Algorithms tested on EM problems in highly inhomogeneous media provided a substantial increase in computational speed over existing solvers. Since these new methods are scalable, the speed-up factors can be dramatic (as much as 10-fold) for large-scale simulations.

To deliver the solvers to the ASC code groups, the project team added the Maxwell solvers to *hypre*. A parallel solver implemented through the *hypre* semi-structured conceptual interface allows LLNL code groups to solve EM problems discretized on structured 2D and 3D domains composed of complicated block parts. This structured-grid software fulfilled a key milestone in a Laboratory Directed Research and Development–funded project involving the SLS team and the HYDRA code group in LLNL's PAT Directorate.

Algorithms for unstructured discretizations have been implemented separately in several SLS research codes. These algorithms use cutting-edge algebraic multigrid techniques developed by the SLS team in collaboration with colleagues in academia. The team is also developing algorithms based on a fundamentally new approach using topological and graph properties of the grid. In



Results from an EM model show the Titan target chamber of LLNL's Nova laser. Simulations of EM models are enhanced by the scalable linear solvers developed with parallel algorithms.

the near future, the team will parallelize these solvers and integrate them into the *hypre* library. They will then be available to simulation code using *hypre*.

Significance

High-resolution accuracy is a crucial component in EM modeling at LLNL and elsewhere. However, large linear systems of equations must be solved efficiently to simulate high-resolution models. The new scalable, parallel algorithms and software developed by the SLS project will enable EM simulation codes to solve these systems optimally and much faster than before, sometimes 10 times faster. With the new solvers, the simulation capabilities of LLNL codes that rely on *hypre* libraries will be enhanced for EM applications in MHD, laser optics, radar, antenna, and circuit design.

Contact Information

Barry Lee, lee123@llnl.gov

5.08 — Data Management for the Large Synoptic Survey Telescope

Overview

A key science requirement of the Large Synoptic Survey Telescope (LSST) is that survey data be reduced in real time to identify alerts—scientifically interesting, transient events in the sky. Once identified, such events can be studied in detail by a network of telescopes. LLNL will use these data for research on astrophysical transients, gamma-ray bursts, planetary transients, microlensing, stellar variability, and more. This scientific driver for Livermore's LSST effort is matched by the Computation Directorate's interest in real-time analysis systems and the research challenges they pose, including the computational problems involving large sensor networks and the need for real-time response and analysis. Because of this synergy, the multidisciplinary team at LLNL is investigating the real-time aspects of the LSST data-analysis pipelines. Our activities will help scientists address the challenges and develop needed competencies in this data-intensive computing field.

A current rendering of the 8.4-meter LSST that uses a special three-mirror design, creating an exceptionally wide field of view. The LSST will have the ability to survey the visible sky in only three nights. (Image courtesy of LSST Corporation.)

Progress in 2005

The Laboratory is participating in several areas of the LSST data-management project. In particular, LLNL is leading the western team and three of the six working groups within the national data-management team.

Through several research and development activities, Livermore researchers helped define the LSST vision and the data challenges to be addressed to support the project's science mission. Our team explored innovative software and hardware design concepts, focusing on three main areas.

First, we designed a data-centric middleware in which the proposed framework provides data services, such as provenance, smart reprocessing, and memorization.

Second, we investigated methods for applying high-performance computing to support the real-time requirements of LSST and image-analysis pipelines. We also are collaborating with colleagues from the University of California at Berkeley to explore the suitability of available file systems for data-intensive computing.

Third, we worked on designing the LSST database schema and defined the science use cases. The use cases became part of a unified modeling language (UML) model that describes the design of the LSST data-management system. Working from these use cases, the LLNL team led the development of a capability matrix to identify available middleware technologies, technology gaps, and the software engineering and integration challenges to be addressed during development.

In addition to producing several publications and presentations, our results are feeding directly into the LSST data-management system's design and will help shape the LSST construction proposal.

Significance

The work on parallelizing image-processing pipelines has demonstrated considerable merit. A parallel implementation of the image content engine pipeline is being used to efficiently analyze satellite images. Research on data-centric pipelines is helping LSST to explore data challenges for telescope construction and to identify sources of performance limitations for data-intensive applications. Using the capability matrix allowed the LSST team to uncover the middleware technology gaps and define the research and development needed to complement available technology. Finally, researchers at the Stanford Linear Accelerator Center, Johns Hopkins University, and LLNL are using the proposed LSST schema to test data-ingesting algorithms.

Contact Information

Ghaleb Abdulla, abdulla1@llnl.gov

5.09 — Data Exploration via Morse Analysis

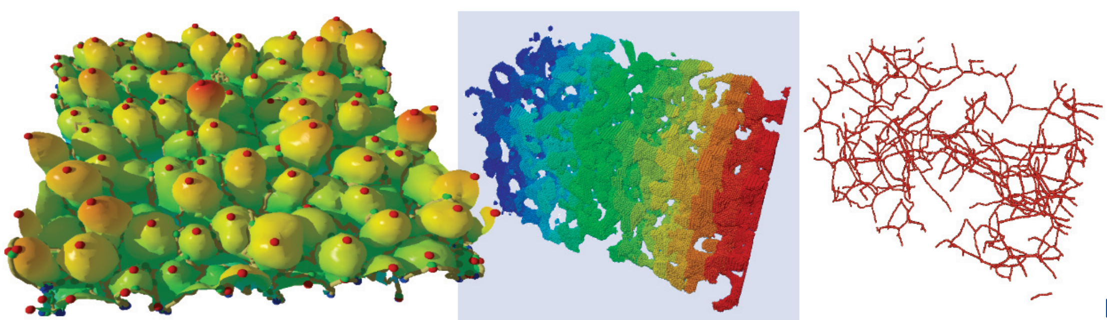
Overview

We are developing a new visualization framework based on general-purpose data-analysis tools coupled with information visualization techniques. The framework will allow fast computation and effective display of metadata road maps guiding the interactive exploration of terascale data sets. We are using Morse analysis to build multiscale topology graphs. (See the figure below.) Because of the graphs' size and complexity, we must develop general multiscale techniques that are suitable for exploring a variety of combinatorial models. (See the figure at right.) The framework will use progressive rendering of multiple linked views and present the graphs with context information that improves the overall data exploration and understanding process.

Progress in 2005

In 2005, we developed core software tools to: (1) construct simplified topologies with tight error bounds, (2) construct robust Jacobi sets for 2D scalar fields, (3) compute integrals on Jacobi sets to correlate 2D fields, and (4) demonstrate the robust computation of critical points in 3D sampled data. We also: (5) implemented a framework with linked views coordinating multiple presentation modalities, (6) scaled the interface for out-of-core graphs, and (7) defined data layouts and file formats to support out-of-core graphs. In addition, we published results in eight refereed papers that appeared in major journals and conference proceedings on visualization.

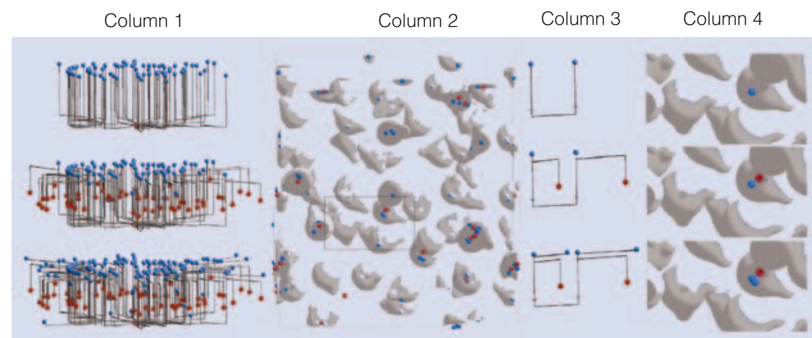
In 2006, we will extend the 2D techniques to 3D so we can test the simplification of the Jacobi sets over time. We will use topological analysis to segment bubble structures in mixing fluid dynamics simulations and time tracking to provide detailed information and global summaries describing turbulent mixing behavior. We will also start integrating the topology components with the visualization framework and extend the tree visualization tools to general graphs with unconstrained connectivity.



Topological analysis was used to define and extract features from two data sets: (left) data from a simulation of Rayleigh–Taylor instability showing Morse complex segmenting into bubbles at the mixing interface, and (middle) complex structures developed in a simulation of a porous medium. The original data modeled as a discrete set of atoms are shown on the right.

Significance

Our research will allow us to develop new tools for data analysis and presentation that will improve the speed at which users can access the information stored in terascale scientific data sets and large semantic graphs. This project will contribute new basic research both in information visualization and in topology-based data analysis.



Multiscale topology graphs showing the electron density of water molecules at high pressure (column 1) at three levels of resolution and (column 2) in semi-transparent level sets of electron density with fine-scale topology. At coarse resolution (top), the graph has one maximum (blue) per molecule. At medium resolution (middle), it has one maximum and one minimum (red) per dipole. At fine resolution (bottom), it has one extremum per atom. Columns 3 and 4 illustrate adaptive refinement used within the rectangle.

Contact Information

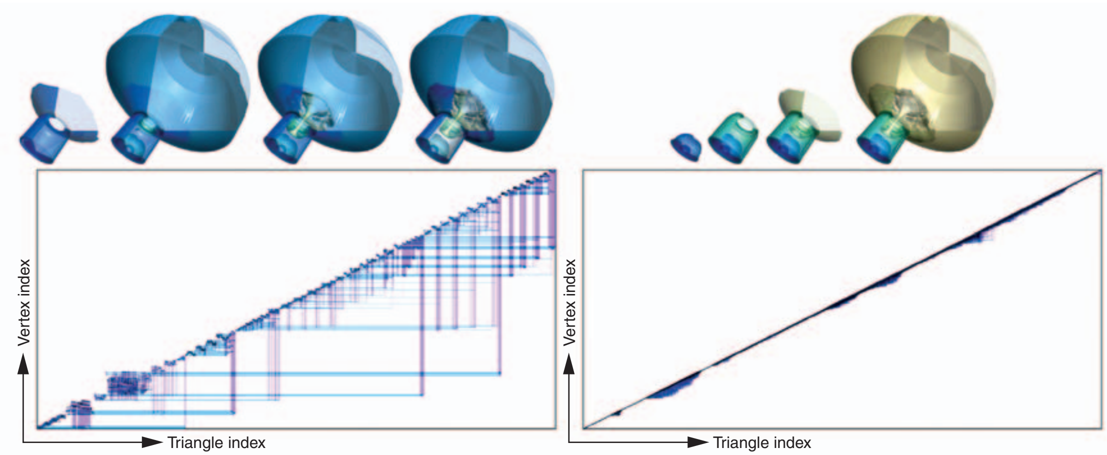
Valerio Pascucci, pascucci1@llnl.gov

5.10 — Enhanced Performance with Streaming and Locality on Unstructured Data

Overview

Mesh processing is an essential part of the simulation cycle and is the foundation of compression, simplification, resampling, filtering, derived-field computation, and other tasks used to set up problems and analyze and visualize data. In scientific simulation and remote sensing, meshes can have up to a billion elements and are outgrowing available memory on today's computers. Their size poses a challenge to efficient mesh processing, especially because the massively parallel supercomputers used to run the simulations are seldom available for subsequent data analysis and visualization. Instead, such tasks are often relegated to smaller clusters or desktop computers with limited memory and input/output (I/O) capabilities.

To address these issues, we are investigating techniques for fast, memory-efficient processing of unstructured meshes, including windowed streaming, cache-coherent data organization, and data compression. These techniques can greatly reduce resource requirements and help prevent today's powerful processors from sitting idly as they wait for data to arrive.



Original layout: footprint = 9.3%, latency = 76% Optimized layout: footprint = 0.83%, latency = 6.1%

Using our metrics for coherence to lay out an unstructured surface mesh for streaming greatly reduces the latency and memory use in a complex simulation. (bottom row) A comparison of mesh vertices and triangles in two layouts of a surface mesh before (left) and after (right) applying our techniques. (top row) Snapshots of the partial mesh are shown as it streams into memory.

Progress in 2005

We designed a framework for sequential, streaming access to arbitrarily large unstructured meshes via a small memory buffer that supports random access to the buffered pieces of a larger data set. This approach allows us to compress a 10-gigabyte mesh in 15 minutes using only 5 megabytes of main memory on a desktop computer, a 50-fold speedup and 3,000-fold reduction in memory and disk use over previous external-memory methods. Furthermore, we demonstrated the ability to string together a sequence of stream modules in a pipeline, much like an assembly line of small specialized tasks. Passing intermediate data directly between modules instead of temporarily storing it on disk results in low latency and superlinear speedup.

Our streaming techniques work best if adjacent elements in the mesh are arranged to appear close together in the stream. We defined metrics for data-stream coherence to estimate resource requirements and direct the organization of a mesh on disk. As shown above, coherence is improved when the metrics are used to design an unstructured surface mesh for streaming.

Cache-oblivious data layout is a complementary approach to streaming when tasks require instant random access to part of a data set (such as for interactive user queries or real-time visualization). Our algorithm for cache-coherent layout results in nearly optimal run-time performance at all levels, from on-chip to external-storage caches, and exhibits 10-fold speedups in interactive visualization and isocontour computation.

Significance

Our resource-efficient techniques help code developers manage huge data streams. Using streaming compression, we are reducing the overall simulation, archival, and analysis time through a more efficient I/O. Our approach also can be applied to structured, semi-structured, and unstructured data. With it, scientists can improve many simulation codes and solve other data-management problems.

Contact Information

Peter Lindstrom, pl@llnl.gov

Project Web site:
http://www.llnl.gov/casc/local/

SECTION

6

Appendices

6.01 — Computation Directorate Industrial Collaborators

Company	Topic	LLNL contact
IBM	Scalable systems, multiple areas	Mark Seager and others
IBM	High-Performance Storage System	Mark Gary
Etnus	Debuggers	Dong Ahn
OpenWorks	Valgrind memory tool development	John Gyllenhaal
SGI	Performance Tool PathForward	Martin Schulz
Tungsten Graphics	Visualization cluster	Dale Southard
R3Vis	Visualization software	Eric Brugger
Kitware	Visualization toolkit	Kathleen Bonnell
CEI–EnSight	Visualization package	Eric Brugger
Intel	Hardware, software, and performance tools	Mark Seager and Chris Chambreau
Hewlett-Packard	Scalable Global File System PathForward	Brent Gorda
Hewlett-Packard	SLURM resource management software	Morris Jette
Cluster File Systems	Lustre file system deployment	Brent Gorda
Red Hat	Operating systems	Mark Seager
OpenMP Consortium	Shared memory programming models	Bronis de Supinski
OpenIB Alliance	Infiniband Linux kernal software	Kim Yates
ZeroFault	Memory correctness tools	John Gyllenhaal
Paratools	Performance tools for BlueGene/L	Scott Futral
Perforce	New source version-control features	Paul Dubois and Carolyn Owens
Iquum	DNA signature target design	Tom Slezak
AeroSoft	Fluid dynamic solver interface	Vic Castillo
Liquid Computing	Hardware alpha testing	Brent Gorda

6.02 — Academic Outreach

University	Faculty	Activity type	Topic	Funding source	LLNL contact
UC Berkeley	Dorian Liepmann	Subcontract	DNA transport in micro- and nanofluidic systems	LDRD	David Trebotich
UC Berkeley	James Demmel	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
UC Berkeley	Jonathan Wurtele	Subcontract	Theoretical and numerical investigations of Raman backscatter	UEPP	Richard Berger
UC Davis	Bernd Hamann	Subcontract	Topological analysis for scientific visualization	LDRD	Valerio Pascucci
UC Davis	Berni Alder	Subcontract	Investigation of flow instabilities at the molecular level	ASC ITS, PSE	Garry Rodrigue
UC Davis	Gerry Puckett	Joint research	An algorithmic and software framework for applied partial differential equations	DOE SciDAC	Rob Falgout
UC Davis	Greg Miller	Subcontract	Algorithm development for interface tracking in a compressible medium	DOE MICS	David Trebotich
UC Davis	Greg Miller	Subcontract	Multiscale modeling of DNA flow and extraction	LDRD	David Trebotich
UC Davis	Greg Miller	Consultant	Computational modeling tools for biodetection microdevices	Overhead PMC	David Trebotich
UC Davis	Ken Joy	Subcontract	Scientific visualization research	ASC DVS	Mark Duchaineau
UC Davis	Kwan-Liu Ma	Subcontract	Large-graph data visualization	OPC	Valerio Pascucci
UC Davis	Kwan-Liu Ma	Subcontract	Interactive visualization of large mixed-resolution volume data	ASC	Steve Louis
UC Davis	Kwan-Liu Ma	Subcontract	Large-graph visualization for homeland security	DOE OCIO	Tony Bartoletti
UC Davis	Mark Rashid	Subcontract	Lagrangian simulation of penetration and other extreme-deformation events: moving beyond meshless methods	UCDRD	Michael Puso
UC Davis	Warren Pickett	Subcontract	Numerical study of coexisting superconductivity and ferromagnetism	UCDRD	Francois Gygi
UC Irvine	Said Elghobashi	Subcontract	Direct numerical simulation and modeling of the dispersion of solid or liquid particles in turbulent flows	UCDRD	Robert Lee
UC Irvine	Said Elghobashi	Subcontract	Direct numerical simulation of turbulent flow laden with solid spheres moving at supersonic speeds	DNT	Bill Bateson
UC Los Angeles	Achiezer Brandt	Consultant	Geometric and algebraic multigrid techniques	DOE MICS	Rob Falgout
UC Los Angeles	Alan Laub	Multilocation appointment	Computational Science and Engineering Consortium	OPC	James McGraw
UC Los Angeles	Heinz-Otto Kreiss	Consultant	Adaptive methods for partial differential equations	DOE MICS	Lori Diachin
UC Los Angeles	Mark Green	Industrial mentor	RIPS: Research in Industrial Projects for Students	ASC ITS	Barna Bihari
UC San Diego	Allan Snavely	Subcontract	Memory access pattern signatures and certificates of relevance for benchmarks	UCDRD	Bronis de Supinski
UC San Diego	Allan Snavely	Joint research	High-end computer system performance: science and engineering	DOE SciDAC	Bronis de Supinski
UC San Diego	Michael Holst	Subcontract	Implicit solvers and preconditioning techniques for simulations of magnetohydrodynamics and core-collapse supernovae	DOE MICS	Carol Woodward
UC San Diego	Randy Bank	Subcontract	Element agglomeration AMGe for contact problems	ASC ITS	Panayot Vassilevski
UC San Diego	Scott Baden	Subcontract	Data-driven execution of communication-tolerant algorithms	UCDRD	Dan Quinlan
UC San Diego	Scott Baden	Subcontract	Data-driven execution of latency-tolerant algorithms	UEPP	Dan Quinlan
UC San Diego	Reagan Moore	Joint research	Scientific Data Management Enabling Technology Center	DOE SciDAC	Terence Critchlow
UC San Diego	Serge Belongie	Subcontract	Robust multiple-object tracking	LDRD	Chandrika Kamath
UC San Diego	Serge Belongie	Subcontract	Computationally efficient example-based image segmentation	UCDRD	Chandrika Kamath
UC San Diego	Reagan Moore, Tim Barnett, Michael Norman, Randy Bank	Subcontract	LLNL–UCSD scientific data management	UCDRD	James McGraw
UC San Francisco	David Saloner	Consultant	Use of computational microfluidics tools to model the hemodynamics of vascular disease	Overhead-PMC	David Trebotich

6.02 — Academic Outreach (cont.)

University	Faculty	Activity type	Topic	Funding source	LLNL contact
UC Santa Barbara	Linda Petzold	Consultant	Numerical methods for sensitivity analysis of DAE systems; algorithmic implementation of methods aimed at large-scale applications	DOE MICS, ASC ITS	Carol Woodward
UC Santa Barbara	Matthew Turk	Subcontract	Visual tracking and recognition for biometrics and interactive visualization	UCDRD	Lenny Tsap
UC Santa Cruz	Darrell Long	Subcontract	Scalable file systems for high-performance computing	ASC	Bill Loewe
Ball State University	Irene Livshits	Summer faculty	Scalable linear solvers	DOE MICS	Rob Falgout
Brown University	Andries van Dam	Subcontract	Developing immersive VR techniques for interactive scientific visualization	ASC	Steve Louis
California Institute of Technology	Daniel I. Meiron	ASC Alliance	Center for Simulating the Dynamic Response of Materials	DOE NNSA	Dick Watson
Carnegie Mellon University	Omar Ghattas	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
Columbia University	David Keyes	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
Columbia University	David Keyes	Subcontract	Acting Director, ISCR	ASC ITS, Overhead	Pete Eltgroth
Cornell University	Sally McKee	Subcontract	Studies in BG/L scalability and reconfigurability	ASC PSE	Bronis de Supinski
Cornell University	Sally McKee	Subcontract	Neural network performance modeling	ASC PERC	Bronis de Supinski
Davidson College	Timothy Chartier	Summer faculty	Scalable linear solvers	ASC ITS	Van Henson
Duke University	Herbert Edelsbrunner	Subcontract	Discrete methods for computing continuous functions	ASC DVS	Valerio Pascucci
Florida Institute of Technology	Jim Jones	Subcontract	Multigrid approaches for Maxwell equations	ASC ITS	Rob Falgout
Front Range Scientific Computations, Inc.	John Ruge	Subcontract	FOSPACK, AMG, and Boomer AMG codes	ASC ITS	Rob Falgout
Front Range Scientific Computations, Inc.	Steve McCormick	Subcontract	FOSPACK-, AMG-, Adaptive AMG–compatible relaxation, parallelization, and application codes	ASC ITS	Rob Falgout
Fusion Numerics	Andrew Knyazev	Subcontract	HYPRE for symmetric eigenvalue problems	ASC ITS	Rob Falgout
Georgia Institute of Technology	Carlton Pu	Joint research	Scientific Data Management Enabling Technology Center	DOE SciDAC	Terence Critchlow
Georgia Institute of Technology	Jarek Rossignac	Subcontract	Streaming isosurfaces for interactive exploration	ASC	Peter Lindstrom
Indiana University	Dennis Gannon	Joint research	Center for Component Technology for Terascale Simulation Software	DOE SciDAC	Tom Epperly
Krell Institute	Lucille Kilmer	Subcontract	Department of Energy High-Performance Computer Science Fellowship Program—LLNL portion	ASC ITS	John May
Krell Institute	Tom Brennan	Subcontract	ASC requirements identification workshops	DHS	David Brown
McMaster University	Nedialko Nedialkov	Sabbatical	Nonlinear solvers and differential equations project	ASC ITS	Radu Serban
New York University	Marsha Berger	Joint research	An algorithmic and software framework for applied partial differential equations	DOE SciDAC	Rob Falgout
New York University	Marsha Berger	Consultant	Adaptive urban-dispersion integrated model project	DHS	Andy Wissink
New York University	Olof Widlund	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
North Carolina State University	Mladen Vouk	Joint research	Scientific Data Management Enabling Technology Center	DOE SciDAC	Terence Critchlow
Northwestern University	Alok Choudhary	Subcontract	Support for enhanced dyninst testing and initial steps toward open binary editing environment	ASC	Bill Loewe
Northwestern University	Alok Choudhary	Joint research	Scientific Data Management Enabling Technology Center	DOE SciDAC	Terence Critchlow
Ohio State University	Umit Catalyurek	Subcontract	Clustering streaming graph data	DHS	Andy Yoo
Ohio Supercomputer Center	Kevin Wohlever	Joint research	Data-intensive computing	ASC	Steve Louis
Pacific Northwest National Laboratory	Eric Stephan	Subcontract	DHS IDS data integration and dissemination workshop	DHS	David Brown
Penn State University	Ludmil Zikatanov	Subcontract	Optimal AMG interpolation and AMG convergence theory	ASC ITS	Rob Falgout
Portland State University	Karen Karavanic	Subcontract	Performance analysis infrastructure for petascale applications	ASC PSE	John May

University	Faculty	Activity type	Topic	Funding source	LLNL contact
Portland State University	Karen Karavanic	Sabbatical	ASC simulation development environments	ASC	John May
Princeton University	Kai Li	Subcontract	R&D scalable display software	ASC	Steve Louis
Purdue University	Zhiqiang Cai	Summer faculty	Scalable linear solvers	ASC ITS	Barry Lee
Rensselaer Polytechnic Institute	Mark Shephard	Joint research	Terascale simulation tools and technologies	DOE SciDAC	Lori Diachin
Rensselaer Polytechnic Institute	Don Schwendeman	Subcontract	Development of numerical methods	DOE MICS	Bill Henshaw
Rice University	John Mellor-Crummey	Subcontract	Open-source software technology for transforming scientific problems	DOE MICS	Dan Quinlan
San Diego State University	Calvin Johnson	Subcontract	Frontier computations in the structure of atomic nuclei	UEPP	W. Erich Ormand
Stanford University	Pat Hanrahan	Subcontract	Scalable visualization of large time-varying data sets	ASC	Holger Jones
Stanford University	Parvis Moin	ASC Alliance	Center for Integrated Turbulence Simulations	DOE NNSA	Dick Watson
Stanford University	Pat Hanrahan	Subcontract	PC-based visualization clusters and software	ASC DVS	Steve Louis
State University of New York, Stony Brook	Jim Glimm	Joint research	Terascale simulation tools and technologies	DOE SciDAC	Lori Diachin
Texas A&M University	Raytcho Lazarov	Subcontract	Construction and preconditioning of finite-element approximations of mixed problems	ASC ITS	Panayot Vassilevski
Texas A&M University	Marvin L. Adams	Subcontract	Efficient massively parallel adaptive algorithms for time-dependent transport on arbitrary spatial grids	ASC	Jim Rathkopf
Texas State University	Anne Ngu	Summer faculty	Working with local and external collaborators to enhance workflow-based infrastructure	DOE MICS	Terence Critchlow
University of Chicago	Don Lamb	ASC Alliance	Center for Astrophysical Thermonuclear Flashes	DOE NNSA	Dick Watson
University of Colorado	Steve McCormick	Consultant	Geometric and algebraic multigrid techniques; FOSLS approach to solving partial differential equations	ASC ITS	Rob Falgout
University of Colorado	Steve McCormick	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
University of Colorado	Tom Manteuffel	Consultant	Algorithms for solving the mathematical equations arising in the simulation of radiation diffusion and transport	ASC ITS	Rob Falgout
University of Illinois, Urbana-Champaign	Michael T. Heath	ASC Alliance	Center for Simulation of Advanced Rockets	DOE NNSA	
University of Illinois, Urbana-Champaign	Mike Folk	Subcontract	HDF5 parallel hierarchical data formats	ASC DVS	Eric Brugger
University of Maryland	Jeff Hollingsworth	Joint research	High-end computer system performance: science and engineering	DOE SciDAC	Bronis de Supinski
University of Maryland	Jeff Hollingsworth	Subcontract	Dynamic instrumentation	ASC	John Gyllenhaal
University of Massachusetts	David Jensen	Subcontract	Relational pathfinding	DOE MICS	Tina Eliassi-Rad
University of Michigan	Peter Honeyman	Subcontract	NFSv4 scalable performance, security, and wide-area optimization	ASC-PSE	Bill Loewe
University of Minnesota	Douglas Arnold	Participating institution	Institute for Mathematics and Its Applications	ASC ITS	James McGraw
University of New Mexico	Thomas Hagstrom	Sabbatical	High-order structure grid methods for wave propagation on complex unbounded domains	DOE MICS	Bill Henshaw
University of North Carolina	Dan Reed	Joint research	High-end computer system performance: science and engineering	DOE SciDAC	Dan Quinlan
University of North Carolina	Dinesh Manocha	Subcontract	Efficient and scalable data structures for topological geometric models	LDRD	Valerio Pascucci
University of North Carolina	Henry Fuchs	Subcontract	Continued group tele-immersion research	ASC	Steve Louis Christine Yang
University of North Carolina	M. Minion	Joint research	An algorithmic and software framework for applied partial differential equations	DOE SciDAC	Rob Falgout

6.02 — Academic Outreach (cont.)

University	Faculty	Activity type	Topic	Funding source	LLNL contact
University of Oregon	Al Malony	Subcontract	TAU instrumentation for high-performance computing	ASC	John May
University of San Francisco	Benjamin (Pete) Wells	Summer faculty	General issues in computer science involving sequential machines, string algorithms, foundations of programming, etc.	ASC ITS	Pat Miller
University of San Francisco	Jeff Buckwalter	Sabbatical	Queuing network models of performance of high-end computing systems	DOE SciDAC	Bronis de Supinski
University of Tennessee	Jack Dongarra	Joint research	Terascale-optimal partial-differential-equation solvers	DOE SciDAC	Lori Diachin
University of Tennessee	Jack Dongarra	Joint research	High-end computer system performance: science and engineering	DOE SciDAC	Dan Quinlan
University of Texas	Kazushige Goto	Subcontract	Optimization and tuning of linear algebra libraries for LLNL IBM systems	ASC	Mark Seager
University of Texas at Austin	Clint Dawson	Sabbatical	Discretizations and splitting methods for radiation-diffusion and compressible flows	ASC ITS	Carol Woodward
University of Utah	Charles Hansen	Consultant	Data exploration, multiresolution scientific data visualization, and algorithm design	ASC DVS	Mark Duchaineau
University of Utah	Charles Hansen	Subcontract	Large data visualization techniques	ASC DVS	Mark Duchaineau
University of Utah	Claudio Silva	Subcontract	Studying the topology of point-set surfaces	LDRD	Valerio Pascucci
University of Utah	David W. Pershing	ASC Alliance	Center for the Simulation of Accidental Fires and Explosions	DOE NNSA	Dick Watson
University of Utah	Ellen Riloff	Subcontract	Bioforensics text extraction	DHS ASC	Terence Critchlow
University of Utah	Steve Parker	Joint research	Center for Component Technology for Terascale Simulation Software	DOE SciDAC	Gary Kumfert
University of Washington	Anne Greenbaum	Consultant	Linear solvers for transport problems; adaptive mesh refinement techniques for transport	ASC ITS	Peter Brown
University of Washington	R. Leveque	Joint research	An algorithmic and software framework for applied partial differential equations	DOE SciDAC	Rob Falgout
University of Waterloo	Hans de Sterck	Subcontract	Improving complexity of parallel AMG	ASC ITS	Rob Falgout
University of Wisconsin	Bart Miller	Subcontract	Support for enhanced dyninst testing and initial steps toward open binary editing environment	ASC	Mary Zosel
University of Wisconsin	Bart Miller	Subcontract	Paradyn performance tools on the AIX platform	ASC	John May
University of Wisconsin	C. Rutland	Joint research	An algorithmic and software framework for applied partial differential equations	DOE SciDAC	Rob Falgout
University of Wisconsin	Jason Kraftcheck	Subcontract	Mesquite software development	DOE SciDAC	Lori Diachin
Worcester Polytechnic Institute	Homer Walker	Consultant	Development and analysis of iterative methods for large-scale linear and nonlinear algebraic systems	ASC ITS	Carol Woodward

6.03 — Journal Papers and Book Chapters

Amendt, P. A., et al. (2005), "Hohlraum-driven ignitionlike double-shell implosions on the Omega laser facility," *Phys. Rev. Lett.* **94** (2), 65004.

Armstrong, R., et al. (2005), "The CCA component model for high-performance scientific computing," *Int. J. Concurrency and Computation: Practice and Experience* **18** (2), (UCRL-JRNL-208616).

Baker, A. H., E. R. Jessup, and T. Manteuffel (2005), "A technique for accelerating the convergence of restarted GMRES," *SIAM J. Matrix Analysis and Applications* **26**, 962–984.

Bala, G., et al. (2005), "Multi-century changes to global climate and carbon cycle: results from a coupled climate and carbon cycle model," *J. Climate* **18**, 4531 (UCRL-JRNL-209851).

Bernholdt, D. E., et al. (2005), "A component architecture for high-performance scientific computing," *Int. J. High-Performance Computing Applications, ACTS Collection Special Issue* (11).

Bertoldo, N. A., et al. (2005), "Development of a real-time radiological area monitoring network for emergency response at Lawrence Livermore National Laboratory," *IEEE Sensors J.* **5** (8), 565–573 (UCRL-JRNL-203376).

Beuning, P. J., et al. (2005), "A non-cleavable umud variant that acts as a umud' mimic," accepted by *J. Biol. Chem.* (UCRL-JRNL-216587).

Bramble, J., Tz. Kolev, and J. Pasciak (2005), "A least-squares approximation method for the time-harmonic Maxwell equations," *J. Numerical Mathematics* **13**, 237–263 (UCRL-JRNL-212459).

Bremer, P.-T., and V. Pascucci, (in press), "A practical approach to two-dimensional scalar topology," in *Topology-Based Methods in Visualization*, H. Hagen, H. Hauser, and H. Theisel, Eds. (Springer-Verlag: New York, NY) (UCRL-BOOK-218178).

Bremer, P.-T., V. Pascucci, and B. Hamann (2005), "Adaptive refinements in hierarchical topological model," in *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, T. Mvller, B. Hamann, and R. Russell, Eds. (Springer-Verlag: New York, NY) (UCRL-BOOK-216201).

Brezina, M., et al. (2005), "Adaptive algebraic multigrid," accepted by *SIAM J. Scientific Computing* (UCRL-JRNL-203501).

Brezina, M., et al. (2005), "Adaptive smoothed aggregation (alpha SA) multigrid," *SIAM Rev.* **47** (6), 317–346 (UCRL-JC-151149).

Brown, P. N., D. E. Shumaker, and C. S. Woodward (2005), "Fully implicit solution of large-scale non-equilibrium radiation diffusion with high order time integration," *J. Computational Physics* **204** (4), 760–783 (UCRL-JRNL-203809).

Budowle, B., et al. (2005), "Toward a system of microbial forensics: from sample collection to interpretation of evidence," *Applied and Environmental Microbiology* **71** (5), 2209–2213.

Buehler, M. J., et al. (2005), "The dynamical complexity of work-hardening: a large-scale molecular dynamics simulation," *Acta Mechanica Sinica* **21** (4), 103–111 (UCRL-JRNL-204002).

Cabot, W. H., et al. (2005), "Large-eddy simulation of Rayleigh–Taylor instability," *Phys. Fluids Gallery of Fluid Motion* **17** (9), 91106.

Cai, X.-C., L. Marcinkowski, and P. S. Vassilevski (2005), "An element agglomeration nonlinear additive Schwarz preconditioned Newton method for unstructured finite element problems," *Applications of Mathematics* **50**, 247–275 (UCRL-JRNL-209688).

Caldeira, K., and M. E. Wickett (2005), "Ocean model predictions of chemistry changes from carbon dioxide emissions to the atmosphere and ocean," *J. Geophys. Res. Oceans* **110** (9), C09S04 (UCRL-JRNL-206209).

Callahan, D. A., et al. (2005), "Heavy-ion target physics and design in the USA," *Nucl. Instrum. Meth. Phys. Res. A Accelerators Spectrometers Detectors and Associated Equipment* **544** (5), 9–15.

Cantz-Paz, E., and C. Kamath (2005), "An empirical comparison of combinations of evolutionary algorithms and neural networks for classification problems," *IEEE Trans. Systems Man Cybernetics Part B* **35** (10), 915–927.

Chand, K. K. (2005), "Component-based hybrid mesh generation," *Int. J. Numerical Meth. Eng.* **62** (2), 747–773.

Chang, B. (2005), "A photon free method to solve radiation transfer equations," accepted by *J. Comput. Phys.* (UCRL-JRNL-215133).

Cheung, S.-C. and C. Kamath (2005), "Robust background subtraction with foreground validation for urban traffic video," *Eurasip J. Applied Signal Processing* **14** (8), 2330–2340.

Chin, H. N. S., et al. (2005), "Evaluation of an urban canopy parameterization in a mesoscale model using VTMX and URBAN 2000 data," *Monthly Weather Rev.* **133** (7), 2043–2068 (UCRL-JRNL-203360).

Chow, E., et al. (2005), "A survey of parallelization techniques for multigrid solvers," accepted by *Frontiers of Parallel Processing for Scientific Computing* (UCRL-BOOK-205864).

Coleman, M. A., et al. (2005), "Low-dose irradiation alters the transcript profiles of human lymphoblastoid cells including genes associated with cytogenetic radioadaptive response," *Radiat. Res.* **163**, 369–382.

Considine, D. B., D. J. Bergmann, and H. Liu (2005), "Sensitivity of Global Modeling Initiative chemistry and transport model simulations of radon-222 and lead-210 to input meteorological data," *Atmos. Chem. Phys.* **5** (12), 3389–3406 (UCRL-JRNL-214794).

Covello, P., and G. Rodrigue (2005), "Solving the Eikonal equation on an adaptive mesh," *Applied Mathematics and Computation* **166** (7), 678–695.

De Sterck, H., U. M. Yang, and J. J. Heys (2005), "Reducing complexity in parallel algebraic multigrid preconditioners," accepted by *SIAM J. Matrix Analysis and Applications* (UCRL-JRNL-206780).

Demmel, J., et al. (2005), "Self-adapting linear algebra algorithms and software," *Proc. IEEE* **93** (2), 293–312.

Dewald, E. L., et al. (2005), "Radiation-driven hydrodynamics of high-Z hohlraums on the National Ignition Facility," *Phys. Rev. Lett.* **95** (11), 215004.

D'haeseleer, P. (2005), "Closing the circle of osmoregulation," *Nature Biotechnol.* **23** (8), 941–942.

D'haeseleer, P. (2005), "How does gene expression clustering work?" *Nature Biotechnol.* **23** (12), 1499–1501.

Diachin, L., et al. (in press), "Parallel adaptive mesh generation," in *Frontiers of Parallel Processing for Scientific Computing, SIAM Book Series on Software, Environment, and Tools*, M. Heroux, P. Raghavan, and H. Simon, Eds. (Society for Industrial and Applied Mathematics: Philadelphia, PA) (UCRL-JRNL-317342).

Dolenc, D., D. Dreger, and S. Larsen (2005), "Basin structure influences on the propagation of teleseismic waves in the Santa Clara Valley, California," *Bull. Seismol. Soc. Am.* **95** (6), 1120–1136.

Eliassi-Rad, T., et al. (2005), "Statistical modeling of large-scale scientific simulation data," in *Next Generation of Data-Mining Applications*, M. Kantardzic and J. Zurada, Eds. (Wiley IEEE Press: New York, NY), pp. 233–262 (UCRL-BOOK-200992).

Espeland, E. K., T. M. Carlsen, and D. Macqueen (2005), "Fire and dynamics of granivory on a California grassland forb," *Biodiv. Conserv.* **14** (2), 267–280.

6.03 — Journal Papers and Book Chapters (cont.)

Falgout, R. D., J. E. Jones, and U. M. Yang (2005), "The design and implementation of hypre, a library of parallel high-performance preconditioners," in *Numerical Solution of Partial Differential Equations on Parallel Computers*, A. M. Bruaset and A. Tveito, Eds. (Springer Verlag: New York, NY), pp. 267–294 (UCRL-JRNL-205459).

Falgout, R. D., J. E. Jones, and U. M. Yang (2005), "Pursuing scalability for hypre's conceptual interfaces," *ACM Trans. Numerical Software* **31** (9), 326–350 (UCRL-JRNL-205407).

Falgout, R. D., J. E. Jones, and U. M. Yang (2005), "Conceptual interfaces in hypre," *Future Generation Computer Systems* **22** (1), 239–251 (UCRL-JC-148957).

Falgout, R. D., P. S. Vassilevski, and L. T. Zikatanov (2005), "On two-grid convergence estimates," *Numerical Linear Algebra with Applications* **12** (6), 471–494 (UCRL-JRNL-203843).

Feil, H., et al. (2005), "Comparison of the complete genome sequences of *Pseudomonas syringae* pv. Syringae B728a and pv. Tomato DC3000," *Proc. Natl. Acad. Sci. U.S.A.* **102** (8), 11064–11069.

Fernandez, J. C., et al. (2005), "Gas-filled hohlraum experiments at the National Ignition Facility," accepted by *Phys. Plasmas.*

Fisher, A., et al. (2005), "A generalized mass lumping technique for vector finite-element solutions of the time-dependent Maxwell equations," *IEEE Trans. Antennas Propagation* **53** (9), 2900–2910.

Fodor, I. K., et al. (2005), "Statistical challenges in the analysis of two-dimensional difference gel electrophoresis experiments using decyder," *Bioinformatics* **21**, 3733–3740.

Gardner, S. N., et al. (2005), "Draft versus finished sequence data for DNA and protein diagnostic signature development," *Nucleic Acids Res.* **33**, 5838–5850 (UCRL-JRNL-208671).

Gardner, S. N., et al. (2005), "System to assess genome sequencing needs for viral protein diagnostics and therapeutics," *J. Clinical Microbiol.* **43** (4), 1807–1817.

Gardner, S. N. and M. C. Wagner (2005), "Software for optimization of SNP and PCR-RFLP genotyping to discriminate many genomes with the fewest assays," *BMC Genomics* **6** (5), 73.

Gardner, S. N. and M. Fernandes (2005), "Prediction of cancer outcome with microarrays," *Lancet* **365** (5), 1685–1685.

Gatti-Bono, C. and P. Colella (2005), "An anelastic allspeed projection method for gravitationally stratified flows," accepted by *J. Comput. Phys.*

Geisbrecht, B. V., et al. (2005), "The crystal structures of EAP domains from *Staphylococcus aureus* reveal an unexpected homology to bacterial superantigens," *J. Biol. Chem.* **280** (4), 17243–17250 (UCRL-JRNL-216376).

Gibbard, S., et al. (2005), "Climate effects of global land cover change," *Geophys. Res. Lett.* **32**, L23705 (UCRL-JRNL-215046).

Giebel, M., D. Oeltz, and P. S. Vassilevski (2005), "Space-time approximation with sparse grids," accepted by *SIAM J. Sci. Computing* (UCRL-JRNL-211686).

Govindasamy, B., et al. (2005), "Increase of carbon cycle feedback with climate sensitivity: results from a coupled climate and carbon cycle model," *Tellus Series B Chem. Phys. Meteorol.* **57** (4), 153–163 (UCRL-JRNL-203401).

Grumet, N. S., et al. (2005), "Intrabasin comparison of surface radiocarbon levels in the Indian Ocean between coral records and three-dimensional global ocean models," *Global Biogeochemical Cycles* **19** (4), GB2010 (UCRL-JRNL-209584).

Guerra, C. and V. Pascucci (2005), "Line-based object recognition using Hausdorff distance: from range images to molecular secondary structures," *Image Vision Computing* **23** (4), 405–415.

Gyulassy, A., et al. (2005), "A topological approach to simplification of three-dimensional scalar functions," accepted by *IEEE Trans. Visualization and Computer Graphics.*

Hale, L. C., T. A. Wulff, and J. C. Sedgewick (2005), "Testing a low-influence spindle drive motor," *Prec. Eng. J. Int. Soc. Prec. Eng. Nanotechnol.* **29** (1), 38727.

Henshaw, W. D. (2005), "On multigrid for overlapping grids," *SIAM J. Sci. Computing* **26**, 1547–1572.

Herrnstein, A., M. Wickett, and G. Rodrigue (2005), "Structured adaptive mesh refinement using leapfrog time integration on a staggered grid for ocean models," *Ocean Modelling* **9**, 283–304 (UCRL-JRNL-202368).

Hillier, L.W., et al. (2005), "Generation and annotation of the DNA sequences of human chromosomes 2 and 4," *Nature* **434** (4), 724–731.

Hindmarsh, A. C., et al. (2005), "SUNDIALS: suite of nonlinear and differential/algebraic equation solvers," *ACM Trans. Mathematical Software* **31** (9), 363–396 (UCRL-JP-200037).

Hinkel, D. E., et al. (2005), "Laser coupling to reduced-scale hohlraum targets at the early light program of the National Ignition Facility," *Phys. Plasmas* **12** (5), 56305.

Hittinger, J. A. F., et al. (2005), "Simulating time-dependent energy transfer between crossed laser beams in an expanding plasma," *J. Comp. Phys.* **209** (11), 695–729 (UCRL-JRNL-207422).

Hornung, R., A. Wissink, and S. Kohn (2005), "Managing complex data and geometry in parallel structured AMR applications," accepted by *Engineering with Computers* (UCRL-JRNL-214559).

Hwa, L. M., M. A. Duchaineau, and K. I. Joy (2005), "Real-time optimal adaptation for planetary geometry and texture: 4–8 tile hierarchies," *IEEE Trans. Visualization and Computer Graphics* **11** (7), 355–368 (UCRL-JRNL-211433).

Isenburg, M., P. Lindstrom, and J. Snoeyink (2005), "Lossless compression of predicted floating-point geometry," *Computer-Aided Design* **37** (7), 869–877 (UCRL-JRNL-208490).

Jones, I. M., et al. (2005), "The genetic basis for variation in radiation sensitivity in the general population," *Radiation Res.* **163** (6), 700–701.

Jonsson, P. R., et al. (2005), "Particle depletion above experimental bivalve beds: In situ measurements and numerical modeling of bivalve filtration in the boundary layer," *Limnol. Oceanog.* **50**, 1989–1998.

Kim, T., J. E. Pasciak, and P. S. Vassilevski (2005), "Mesh independent convergence of the modified inexact Newton method for a second order nonlinear problem," *Numerical Linear Algebra with Applications* **13**, 23–47 (UCRL-JRNL-206774).

Kostova, T. (2005), "Persistence of viral infections on the population level explained by an immunoepidemiological model," accepted by *Mathematical Biosciences* (UCRL-JRNL-209391).

Kostova, T. and T. Carlsen (2005), "The effect of habitat size and predation on the time to extinction of prairie vole populations: simulation studies via SERDYCA," *Ecological Complexity* **2** (3), 35–57.

Kostova, T. and T. Carlsen (2005), "The effect of small-size habitat disturbances on population density and time to extinction of the prairie vole," *Nonlinear Analysis Real World Applications* **6** (9), 731–746.

Kreiss, H.-O. and N. A. Petersson (2005), "A second order accurate embedded boundary method for the wave equation with Dirichlet data," accepted by *SIAM J. Sci. Computing* (UCRL-JRNL-202686).

Kubista, M., et al. (2005), "Trilinear fluorescence spectroscopy," *Spectroscopy Europe* **16** (1), 8–13.

Kulp, K. S., et al. (2005), "Essiac and Flor-Essence herbal tonics stimulate the in vitro growth of human cancer cells," accepted by *Breast Cancer Res. Treat.*

Kyker, A. B., W. E. Pickett, and F. Gygi, (2005), "Fermiology and Fulde-Ferrell-Larkin-Ovchinnikov phase formation," *Phys. Rev. B* **71** (6), 224517.

Lehmann, J., et al. (2005), "Monte Carlo treatment planning for molecular targeted radiotherapy within the MINERVA system," *Phys. Med. Biol.* **50** (3), 947–958.
Lightstone, F. C., et al. (2005), "A first-principles molecular dynamics study of calcium in water," *Chemphyschem* **6** (9), 1745–1749.
Loots, G. G. and I. Ovcharenko (2005), "Dcode. org anthology of comparative genomic tools," *Nucleic Acids Res.* **33** (7), W56–W64.
Ma, J., et al. (2005), "Multiscale simulations using generalized interpolation material point (GIMP) method and SAMRAI parallel processing," *CMES Computer Modeling in Engineering and Sciences* **8** (5), 135–152 (UCRL-JRNL-208241).
Max, N. (2005), "Progress in scientific visualization," *The Visual Computer* **21** (11), 980–985 (UCRL-JRNL-208330).
McEnerney, J., R. Robson, and G. Stengle (2005), "Convex polarities over ordered fields," accepted by *J. Pure Applied Algebra.*
McEnerney, J. and G. Stengle (2005), "Alternative evidence for nonnegativity," accepted by *J. Pure Applied Algebra.*
McInnes, L. C., et al. (in press), "Parallel PDE-based simulations using the common component architecture," in *Numerical Solution of Partial Differential Equations on Parallel Computers*, A. M. Bruaset and A. Tveito, Eds. (Springer-Verlag: New York, NY), pp. 327–384.
Mirin, A. A. and W.B. Sawyer (2005), "A scalable implementation of a finite-volume dynamical core in the Community Atmosphere Model," *Int. J. High-Performance Computing Applications* **19**, 203–212 (UCRL-JRNL-206816).
Nishtala, R., et al. (2005), "When cache block sparse matrix vector multiply works and why," accepted by *Applicable Algebra in Engineering, Communication, and Computing: Special Issue on Computational Linear Algebra and Sparse Matrix Computations* (UCRL-JRNL-217633).
Ovcharenko, I., et al. (2005), "Mulan: Multiple-sequence local alignment and visualization for studying function and evolution," *Genome Res.* **15** (1), 184–194.
Ovcharenko, I., et al. (2005), "Evolution and functional classification of vertebrate gene deserts," *Genome Res.* **15** (1), 137–145.
Ovcharenko, I. and M. A. Nobrega (2005), "Identifying synonymous regulatory elements in vertebrate genomes," *Nucleic Acids Res.* **33** (7), W403–W407.
Pascucci, V., K. Cole-McLaughlin, and G. Scorzelli (2005), "Multi-resolution computation and presentation of contour trees," in *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, T. Mvller, B. Hamann, and R. Russell, Eds. (Springer-Verlag: New York, NY) (UCRL-BOOK-216196).
Pernice, M. and R. Hornung (2005), "Newton-Krylov-FAC methods for problems discretized on locally-refined grids," *Computing and Visualization In Science* **8**, 107–118 (UCRL-JC-151614).
Pierce, T. and G. Rodrigue (2005), "A parallel two-sided contact algorithm in ALE3D," *Computer Methods in Applied Mechanics and Engineering* **194**, 3127–3146.
Ramirez, A. L., et al. (2005), "Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach," *J. Geophys. Res. Solid Earth* **110** (2), B02101.
Raty, J. Y., F. Gygi, and G. Galli (2005), "Growth of carbon nanotubes on metal nanoparticles: A microscopic mechanism from ab initio molecular dynamics simulations," *Phys. Rev. Lett.* **95** (8), 96103.
Rieben, R. N., G. H. Rodrigue, and D.A. White (2005), "A high order mixed vector finite element method for solving the time dependent Maxwell equations on unstructured grids," *J. Comp. Phys.* **204** (4), 490–519.
Santer, B. D., et al. (2005), "Amplification of surface temperature trends and variability in the tropical atmosphere," *Science* **309** (9), 1551–1556.
Scherlis, D. A., et al. (2005), "A unified electrostatic and cavitation model for first-principles molecular dynamics in solution," accepted by *J. Chem. Phys.* (UCRL-JRNL-217141).
Shestakov, A. I., J. A. Greenough, and L. H. Howell (2005), "Solving the radiation diffusion and energy balance equations using pseudo-transient continuation," *J. Quantitative Spectroscopy and Radiative Transfer* **90** (1), 1–28 (UCRL-JC-138727-REV-1).
Sjogreen, B. and N. A. Petersson (2005), "Perfectly matched layers for Maxwell's equations in second order formulation," *J. Comp. Phys.* **209** (10), 19–46.
Tong, C. (2005), "Refinement strategies for stratified sampling methods," accepted by *Reliability Eng. System Safety* (UCRL-JRNL-204408).
Tong, C., M. Brezina, and R. Becker (2005), "Parallel algebraic multigrids for structural mechanics," accepted by *SIAM J. Sci. Computing* (UCRL-JRNL-204167).
Trebotich, D., P. Colella, and G.H. Miller (2005), "A stable and convergent scheme for viscoelastic flow in contraction channels," *J. Comp. Phys.* **205** (5), 315–342.
Trebotich, D., et al. (2005), "Modeling of microfabricated microneedles for minimally invasive drug delivery, sampling and analysis," *Biomed. Microdevices* **5** (9), 245–251.
Tringe S. G., et al. (2005), "Comparative metagenomics of microbial communities," *Science* **308** (4), 554.
Vassilevski, P. S. and L. T. Zikatanov (2005), "Multiple vector preserving interpolation mappings in algebraic multigrid," accepted by *SIAM J. Matrix Analysis and Applications* (UCRL-JRNL-208036).
Vetter, J. S., et al. (2005), "Evaluating high-performance computers," *Concurrency and Computation—Practice and Experience* **17** (8), 1239–1270 (UCRL-JC-151967).
Wang, P., et al. (2005), "Parallel computation of the Regional Ocean Model System (ROMS)," *Int. J. High-Performance Computing Applications* **19**, 375–385, (UCRL-JRNL-211096).
Yang, U. M. (2005), "Parallel algebraic multigrid methods—high-performance preconditioners," in *Numerical Solution of Partial Differential Equations on Parallel Computers*, A. M. Bruaset and A. Tveito, Eds. (Springer-Verlag: New York, NY), pp. 209–236 (UCRL-BOOK-208032).
Yi, Q. and D. Quinlan (2005), "Applying loop optimizations to object-oriented abstractions through general classification of array semantics," *Lecture Notes in Computer Science* **3602**, 253–267.
Yoon, S. E., et al. (2005), "Cache-oblivious mesh layouts," *ACM Trans. Graphics* **24** (7), 886–893 (UCRL-JRNL-211774).
Zemla, A., et al. (2005), "AS2TS system for protein structure modeling and analysis," *Nucleic Acids Res.* **33** (7), W111–W115 (UCRL-JRNL-209684).
Zhou, C. L. E., et al. (2005), "Computational approaches for identification of conserved/unique binding pockets in the A chain of ricin," *Bioinformatics* **21** (7), 3089–3096 (UCRL-JRNL-209388).

6.04 — Refereed Conferences

Almasi, G., et al. (2005), "Scaling physics and material science applications on a massively parallel BlueGene/L system," in *Proc. 19th ACM Int. Conf. Supercomputing* (Association for Computing Machinery) (UCRL-CONF-212958).

Bagherjeiran, A., and C. Kamath (2006), "Graph-based methods for orbit classification," in *Proc. Sixth SIAM Int. Conf. Data Mining.*

Barthelemy, M., E. Chow, and T. Eliassi-Rad (2005), "Knowledge representation issues in semantic graphs for relationship detection," 2005 AAAI Spring Symposium—AI Technologies for Homeland Security (Stanford, CA), in *Papers from the 2005 AAAI Spring Symp.—AI Technologies for Homeland Security* (AAAI Press) (UCRL-CONF-209845).

Barton, N. R., et al. (2005), "Adaptive sampling in computational solid mechanics," in *8th U.S. Natl. Cong. Computational Mechanics* (USACM) (UCRL-PRES-213881).

Bihari, B. L., and P. N. Brown (2004), "High order finite volume nonlinear schemes for the Boltzmann transport equation," in *Lecture Notes in Computational Science and Engineering* (Springer), **48**, pp. 401–421.

Bremer, P.-T., V. Pascucci, and B. Hamann (2005), "Maximizing adaptivity in hierarchical topological models using extrema trees," in *IEEE Proc. Int. Conf. Shape Modeling and Applications 2005* (IEEE), pp. 298–307 (UCRL-PROC-216200).

Carey, R. W., et al. (2005), "Status of the use of large-scale Corba-distributed software framework for NIF controls," in 10th Int. Conf. Large Experimental Physics Control Systems (http://elise.epfl.ch/html/contents.htm) **29J**, pp. TU1.2–50 (UCRL-CONF-215284).

Chartier, T., et al. (2005), "Spectral element agglomerate AMGe," in *Lecture Notes in Computational Science and Engineering* (Springer-Verlag) (UCRL-PROC-212439).

Childs, H., E. Brugger, K. Bonnell, J. Meredith, M. Miller, B. Whitlock, and N. Max (2005), "A contract based system for large data visualization," in *Proc. IEEE Visualization 2005* (IEEE Computer Society), pp. 191–198 (UCRL-CONF-211357).

Cook, A., W. Cabot, M. Welcome, P. Williams, B. Miller, B. de Supinski, and R. Yates (2005), "Terascalable algorithms for variable-density elliptic hydrodynamics with spectral accuracy," in *Proc. ACM/IEEE Supercomputing 2005 Conf.*

Dahlgren, T. L., and P. T. Devanbu (2005), "Improving scientific software component quality through assertions," in *Proc. Second Int. Workshop Software Engineering for High Performance Computing System Applications* (IEEE), pp. 73–77 (UCRL-CONF-211000).

DeSouza, J., et al. (2005), "Automated, scalable debugging of MPI programs with Intel® Message Checker," in *Int. Workshop Software Engineering for High Performance Computing System Applications (*IEEE) (UCRL-CONF-210771).

Eder, D. C., et al. (2005), "Optimization of experimental designs by incorporating NIF facility impacts," in *IFSA 2005*.

Eliassi-Rad, T., and T. Critchlow (2005), "A hybrid approach to multiresolution modeling of large-scale scientific data," in *Proc. 20th Annual ACM Symp. Applied Computing* (ACM), pp. 511–518 (UCRL-CONF-206419).

Fodor, I. K., A. E. Holtz-Morris, and S. L. McCutchen-Maloney (2005), "Growth curve models for the analysis of phenotype arrays for a systems biology overview of *Yersinia pestis*," in *Proc Am. Statistical Assoc. Biopharmaceutical Section* (UCRL-PROC-215338).

Gardner, S.N., and T.R. Slezak (2005), "Bioinformatics for diagnostics, forensics, and virulence characterization and detection," DHS Homeland Security Conference (Boston, MA), available on DOE's Information Bridge at www.osti.gov/servlets/purl/15015889-mpo3Hc/native/ (UCRL-CONF-211202).

Gygi, F., et al. (2005), "Large-scale first-principles molecular dynamics simulations on the BlueGene/L platform using the Qbox code," in *Proc. ACM/IEEE Supercomputing 2005 Conf.* (Association for Computing Machinery) (UCRL-CONF-211703).

Gyulassy, A., et al. (2005), "Topology-based simplification for feature extraction from 3D scalar fields," in *Proc. IEEE Conf. Visualization 2005 (*IEEE) 535–542 (UCRL-PROC-216195).

Hedges, R., et al. (2005), "Parallel file system testing for the lunatic fringe: the care and feeding of restless I/O power users," in *Information Retrieval from Very Large Storage Systems* (IEEE Computer Society), pp. 3–13 (UCRL-315236).

Hittinger, J. A. F., Y. Suzuki, and B. van Leer (2005), "Investigation of the discontinuous Galerkin method for first-order PDE approaches to CFD," in *AIAA Paper* (American Institute of Aeronautics and Astronautics), **4989** (UCRL-CONF-212631).

Hoeflinger, J. P., and B. R. de Supinski (2005), "The OpenMP memory model," in *Int. Workshop on OpenMP* (Springer-Verlag) (UCRL-CONF-212641).

Ipek, E., et al. (2005), "An approach to performance prediction for parallel applications," in *Euro-Par (*ACM) (UCRL-CONF-212365).

Isenburg, M., and P. Lindstrom (2005), "Streaming meshes," in *Proc. IEEE Visualization* (IEEE) 231–238 (UCRL-CONF-211608).

Isenburg, M, P. Lindstrom, and J. Snoeyink (2005), "Streaming compression of triangle meshes," in Proc. *Eurographics Symposium on Geometry Processing 2005*, pp. 111–118 (UCRL-CONF-210481).

Kamath, C., et al. (2005), "Salient points for tracking moving objects in video," in *Proc. SPIE* **5685**, pp. 442–453.

Karavanic, K. L., et al. (2005), "Integrating database technology with comparison-based parallel performance diagnosis: the PerfTrack performance experiment management tool," in *Proc. ACM/IEEE Supercomputing 2005 Conf.* (Association for Computing Machinery) (UCRL-CONF-211654).

Koniges, A. E., et al. (2005), "Modeling NIF experimental designs with adaptive mesh refinement and Lagrangian hydrodynamics," in *Inertial Fusion Sciences and Applications 2005* (American Nuclear Society) (UCRL-CONF-217824).

Laney, D. E., et al. (2005), "Hardware accelerated simulated radiography," in *Proc. IEEE Visualization 2005* (IEEE), pp. 343–350 (UCRL-CONF-214415).

Liu, L., et al. (2005), "XWRAPComposer: a multi-page data extraction service for bio-computing applications," in *Proc. IEEE Service Computing Conf.*, pp. 271–278 (UCRL-CONF-209837).

Marathe, J., F. Mueller, and B. R. de Supinski (2005), "A hybrid hardware/software approach to efficiently determine cache coherence bottlenecks," in *ACM Int. Conf. Supercomputing* (UCRL-CONF-212502).

McCallen, R., et al. (2005), "DOE's effort to reduce truck aerodynamic drag through joint experiments and computations," in *Proc. 2005 SAE Commercial Vehicle Engineering Conf.* **05CV-107** (UCRL-CONF-213551).

Natarajan, V., and V. Pascucci (2005), "Volumetric data analysis using Morse–Smale complexes," in *IEEE Proc. Int. Conf. Shape Modeling and Applications 2005*, pp. 320–325 (UCRL-PROC-216199).

Newsam, S., and C. Kamath (2005), "Comparing shape and texture features for pattern recognition in simulation data," in *Proc. SPIE* **5672**, pp. 106–117.

Ngu, A. H. H., D. Buttler, and T. Critchlow (2005), "Automatic generation of data types for classification of Deep Web sources," in *Lecture Notes in Computer Science* (Springer-Verlag), **3615**, pp. 266–274 (UCRL-CONF-209719).

O'Brien, M. J., J. M. Taylor, and R. J. Procassini (2005), "Dynamic load balancing of parallel Monte Carlo transport calculations," in *Monte Carlo 2005 Conf. Proc.* (UCRL-CONF-208938).

Oliker, L., et al. (2005), "Leading computational methods on scalar and vector HEC platforms," in *Proc. ACM/IEEE Supercomputing 2005 Conf.* (UCRL-CONF-212184).

Ollivier-Gooch, C., et al. (2006), "The TSTT mesh interface," in *Proc. 44th AIAA Aerospace Sciences Meeting and Exhibit*.

Pope, G. (2005), "Risk based regression testing," in *System and Software Technology Conf.* CD-ROM.

Pope, G. (2005), "Test is a four letter word," in *Proc. Better Software Conf. Expo 2005*.

Quinlan, D., et al. (2005), "Toward the automated generation of components from existing source code," in *Proc. Second Workshop on Productivity and Performance in High-End Computing* (IBM Research), pp. 12–19 (UCRL-CONF-208403).

Quinlan, D., S. Ur, and R. Vuduc (2005), "An extensible open-source compiler infrastructure for testing," in *Proc. IBM Verification Conf. (*Springer-Verlag) (UCRL-PROC-217691).

Rocco, D., et al. (2005), "Domain-specific Web service discovery with service class descriptions," *Proc. IEEE Int. Conf. Web Services* (UCRL-CONF-209717).

Schneider, M.B., et al. (2005), "X-ray flux and x-ray burn-through experiments on reduced scale targets at the NIF and OMEGA lasers," in *Inertial Fusion Sciences and Applications: State of the Art 2005*.

Schordan, M., and D. Quinlan (2005), "Specifying transformation sequences as computation on program fragments with an abstract attribute grammar," in *Proc. Fifth IEEE Int. Workshop on Source Code Analysis and Manipulation*, pp. 97–106.

Schulz, M. (2005), "Extracting critical path graphs from MPI applications," in *Proc. IEEE Cluster 2005* (UCRL-CONF-214107).

Schulz, M., et al. (2005), "Owl: next generation system monitoring," in *Proc. ACM Computing Frontiers* (UCRL-CONF-209855).

Schulz, M., J. May, and J. Gyllenhaal (2005), "DynTG: a tool for interactive, dynamic instrumentation," in *Lecture Notes in Computer Science* (Springer-Verlag), **3515**, pp. 140–148 (UCRL-CONF-209840).

Schulz, M., et al. (2005), "Scalable dynamic binary instrumentation for BlueGene/L," in *Workshop on Binary Instrumentation and Applications* (UCRL-CONF-215232).

Streitz, F. H., et al. (2005), "100+ TFlop solidification simulations on BlueGene/L," in *Proc. ACM/IEEE Supercomputing 2005 Conf.* (UCRL-CONF-211681).

Tandon, M., et al. (2005), "Structured adaptive mesh refinement (SAMR) simulation of a buoyant helium plume," in *AIP Conf. Proc.*

Van De Vanter, M., D. Post, and M. Zosel (2005), "HPC needs a tool strategy," in *Proc. Second Int. Workshop Software Engineering for High Performance Computing System Applications*, see http://csdl.ics.hawaii.edu/se-hpcs/schedule.htm (UCRL-CONF-210246).

Vuduc, R., and H.-J. Moon (2005), "Fast sparse matrix-vector multiplication by exploiting variable block structure," High Performance Computing and Communications, in *LNCS* (Springer-Verlag), **3726**, pp. 807–816 (UCRL-TR-213454).

White, B. S., et al. (2005), "Improving the computational intensity of unstructured mesh applications," in *ACM Int. Conf. Supercomputing* (UCRL-CONF-212479).

Wissink, A., et al. (2006), "Adaptive urban dispersion integrated model," in *Sixth Symp. Urban Environment* (American Meteorological Society) (UCRL-PROC-216813).

Yee, H. C., and B. Sjogreen (2005), "Non-linear filtering in compact high order schemes," 19th Int. Conf. Numerical Simulation of Plasmas, in *J. Plasma Physics*.

6.05 — Abbreviations and Acronyms

1D – one-dimensional

2D – two-dimensional

3D – three-dimensional

AD – Active Directory

ADVISE – Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement

AIS – Administrative Information Services

ALC – Advanced Simulation and Computing Linux Cluster

ASAP – Academic Strategic Alliance Program

ASC – Advanced Simulation and Computing

ASCI – Advanced Strategic Computing Initiative

BFS – breadth-first search

BG/L – BlueGene/L

BKC – Biodefense Knowledge Center

C-SAFE – Center for Simulation of Accident and Fire Environments

CASC – Center for Applied Scientific Computing

CDC – Centers for Disease Control and Prevention

CHAOS – Clustered High Availability Operating System

CFS – Cluster File Systems

CIAC – Computer Incident Advisory Capability

CMDB – Configuration Management Database

CO_2 – carbon dioxide

CPU – central processing unit

CSP – Cyber Security Program

ddcMD – domain decomposition molecular dynamics

DHS – Department of Homeland Security

DOE – Department of Energy

E&E – Energy and Environment

EM – electromagnetics

EOS – equation of state

EPICS – Experimental Physics and Industrial Control System

FEL – free-electron laser

FMD – foot-and-mouth disease

FPMD – first-principles molecular dynamics

GB/s – gigabytes per second

GUI – graphical user interface

HPC – high-performance computing

I/O – input/output

INCCA – Integrated Climate and Carbon

IP – Internet Protocol

IT – information technology

KBALAP – Knowledge Base Automated Location Assessment and Prioritization

LC – Livermore Computing

LCLS – Linac Coherent Light Source

LDAP – Lightweight Directory Access Protocol

LDRD – Laboratory Directed Research and Development

LEOS – Livermore Equation of State

LSST – Large Synoptic Survey Telescope

M&IC – Multiprogrammatic and Institutional Computing

MB – megabyte

MCR – Multiprogrammatic and Institutional Computing Capability Resource

MD – molecular dynamics

MESA – Multiscale Epidemiological/Economic Simulation and Analysis

MHD – magnetohydrodynamics

MPI – message-passing interface

NAI – Nonproliferation, Arms Control, and International Security

NAP – Nuclear Assessment Program

NAS – network-attached storage

NIF – National Ignition Facility

NLP – natural language processing

NNSA – National Nuclear Security Administration

NWC – nuclear weapons complex

PAT – Physics and Advanced Technologies

PBG – Pathogen Bioinformatics Group

PDF – portable document file

PF – 1,000 trillion (10^{15}) floating-point operations per second

PKI – public-key infrastructure

PPM – parts per million

RBAP – Regional Body-Wave Amplitude Processor

SARS – severe acute respiratory syndrome

SC|05 – 2005 Supercomputing Conference

SCAT – Secure Communication and Teleconference

SGI – Silicon Graphics, Inc.

SLAC – Stanford Linear Accelerator Center

SLS – scalable linear solvers

SLURM – Simple Livermore Utility for Resource Management

SQE – software quality engineering

TB – terabyte

TF – trillion floating-point operations per second

TSF – Terascale Simulation Facility

UCLA – University of California at Los Angeles

UIMA – Unstructured Information Management Architecture

UML – Unified Modeling Language

V&V – verification and validation

XTOD – x-ray beam transport, optics, and diagnostics