ISCR

Institute for Scientific Computing Research

Annual Report: Fiscal Year 2004

Collaborating with Academia

Lawrence Livermore National Laboratory

P.O. Box 808, L-561, Livermore, CA 94551

The University Relations Program (URP) encourages collaborative research between Lawrence Livermore National Laboratory (LLNL) and the University of California campuses. The Institute for Scientific Computing Research (ISCR) actively participates in such collaborative research, and this report details the Fiscal Year 2004 projects jointly served by URP and ISCR. For a full discussion of all URP projects in FY 2004, please request a copy of the URP FY 2004 Annual Report by contacting

Lawrence Livermore National Laboratory
Edie Rock, University Relations Program
P. O. Box 808, L-413
Livermore, CA 94551

UCRL-TR-133866-04

Work 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

Director's Report

Large-scale scientific computation and all of the disciplines that support and help to validate it have been placed at the focus of Lawrence Livermore National Laboratory (LLNL) by the Advanced Simulation and Computing (ASC) program of the National Nuclear Security Administration (NNSA) and the Scientific Discovery through Advanced Computing (SciDAC) initiative of the Office of Science of the Department of Energy (DOE). The maturation of computational simulation as a tool of scientific and engineering research is underscored in the November 2004 statement of the Secretary of Energy that, "high performance computing is the backbone of the nation's science and technology enterprise."

LLNL operates several of the world's most powerful computers—including today's single most powerful—and has undertaken some of the largest and most compute-intensive simulations ever performed. Ultrascale simulation has been identified as one of the highest priorities in DOE's facilities planning for the next two decades. However, computers at architectural extremes are notoriously difficult to use efficiently. Furthermore, each successful terascale simulation only points out the need for much better ways of interacting with the resulting avalanche of data. Advances in scientific computing research have, therefore, never been more vital to LLNL's core missions than at present. Computational science is evolving so rapidly along every one of its research fronts that to remain on the leading edge, LLNL must engage researchers at many academic centers of excellence. In Fiscal Year 2004, the Institute for Scientific Computing Research (ISCR) served as one of LLNL's main bridges to the academic community with a program of collaborative subcontracts, visiting faculty, student internships, workshops, and an active seminar series.

The ISCR identifies researchers from the academic community for computer science and computational science collaborations with LLNL and hosts them for short- and long-term visits with the aim of encouraging long-term academic research agendas that address LLNL's research priorities. Through such collaborations, ideas and software flow in both directions, and LLNL cultivates its future workforce. The Institute strives to be LLNL's "eyes and ears" in the computer and information sciences, keeping the Laboratory aware of and connected to important external advances. It also attempts to be the "feet and hands" that carry those advances into the Laboratory and incorporates them into practice.

ISCR research participants are integrated into LLNL's Computing and Applied Research (CAR)

Department, especially into its Center for Applied Scientific Computing (CASC). In turn, these organizations address computational challenges arising throughout the rest of the Laboratory. Administratively, the ISCR flourishes under LLNL's University Relations Program (URP). Together with the other five institutes of the URP, it navigates a course that allows LLNL to benefit from academic exchanges while preserving national security. While it is difficult to operate an academic-like research enterprise within the context of a national security laboratory, the results declare the challenges well met and worth the continued effort.

Fiscal year 2004 was the fifth full year under Acting Director David Keyes. Keyes, the Fu Foundation Professor of Applied Mathematics at Columbia University and an ISCR faculty participant since October 1997, dedicated one-third of his time to the technical program of the ISCR. James McGraw continued as the Deputy Director of the ISCR, and Linda Becker as the Institute Administrator. Paula Ashley, Pamela Mears, and Char Paulo logistically supported the large visitor and summer programs of the ISCR.

The ISCR continues to have a small contingent of research staff members within its organization. Three ISCR staff—Nelson Max, Garry Rodrigue, and Rao Vemuri—hold joint appointments as professors at the University of California, Davis and senior researchers at LLNL. In addition, the ISCR hosted eight post-doctoral staff: Alison Baker, David Buttler, Tzanio Kolev, Shawn Newsam, Stefan Nilsson, Dan Reynolds, Megan Thomas, and Qing Yi. Finally, the ISCR served as the host for 10 students (listed in Table 1) on a Student-Employee Graduate Fellowship (SEGRF). This fellowship enables students to work with LLNL researchers half-time while pursuing their PhDs.

The ISCR enables substantial interactions between academia and LLNL staff through consultants and participating guests. Consulting agreements are vehicles for permitting academics to interact with LLNL in a compensated fashion. Consultants may serve on review committees, present short courses, and visit LLNL periodically for

Table 1. Students on a Student-Employee Graduate Fellowship (SEGRF).

Name	University	LLNL Advisor(s)	Time at LLNL
Peer-Timo Bremer	University of California, Davis	Dan Laney	June 12, 2002 – Aug. 6, 2004
Sam Brockington	University of California, Davis	Garry Rodrigue & Dave Hwang	Oct. 8, 2001 – Aug. 31, 2006
Paul Castellucci	Stanford University	Rose McCallen	Oct. 13, 2003 – Jan. 31, 2005
Aaron Fisher	University of California, Davis	Garry Rodrigue	July 1, 2002 – June 30, 2006
Benjamin Gregorski	University of California, Davis	Mark Duchaineau	June 25, 2001 – Aug. 13, 2004
Jeff Hagelberg	Purdue University	Paul Amala	Sept. 22, 2003 – Sept. 25, 2004
Aaron Herrnstein	University of California, Davis	Michael Wickett	Mar. 26, 2004 – Mar. 31, 2004
Andrew Nonaka	University of California, Davis	David Trebotich	Oct. 1, 2003 – Sept. 30, 2007
Rob Rieben	University of California, Davis	Garry Rodrigue	Oct. 2, 2000 – Oct. 1, 2004
Joshua Senecal	University of California, Davis	Mark Duchaineau	Nov. 1, 2001 – Oct. 31, 2005

Table 2. Nine ISCR consultants for FY 2004.

Consultant	Affiliation	LLNL Contact
Randolph Bank	University of California, San Diego	Rob Falgout
Achiezer Brandt	University of California, Los Angeles	Rob Falgout
Gene Golub	Stanford University	Edmond Chow
Anne Greenbaum	University of Washington	Peter Brown
Heinz-Otto Kreiss	University of California, Los Angeles	Lori Diachin & Bill Henshaw
Thomas Manteuffel	University of Colorado	Rob Falgout & Peter Brown
Stephen McCormick	University of Colorado	Rob Falgout & Peter Brown
Linda Petzold	University of California, Santa Barbara	Carol Woodward & Radu Serban
Homer Walker	Worcester Polytech Institute	Peter Brown

technical meetings. All consultants have a specific LLNL technical point of contact for overseeing their interactions. Table 2 lists the nine ISCR consultants for FY 2004.

Participating Guests are researchers from academia or industry that need intermittent access to LLNL staff and facilities. This status permits an appropriate security clearance and the ability to quickly arrange for on-site visits with LLNL staff over a period of one month to two years. Table 3 lists ISCR's 38 participating guests for FY 2004.

The pages of this annual report summarize the activities of the faculty members, postdoctoral researchers, students, and guests from industry and other laboratories who participated in LLNL's computational mission under the auspices of the ISCR during FY 2004. These activities, which are further detailed in the accompanying CD-ROM, fall

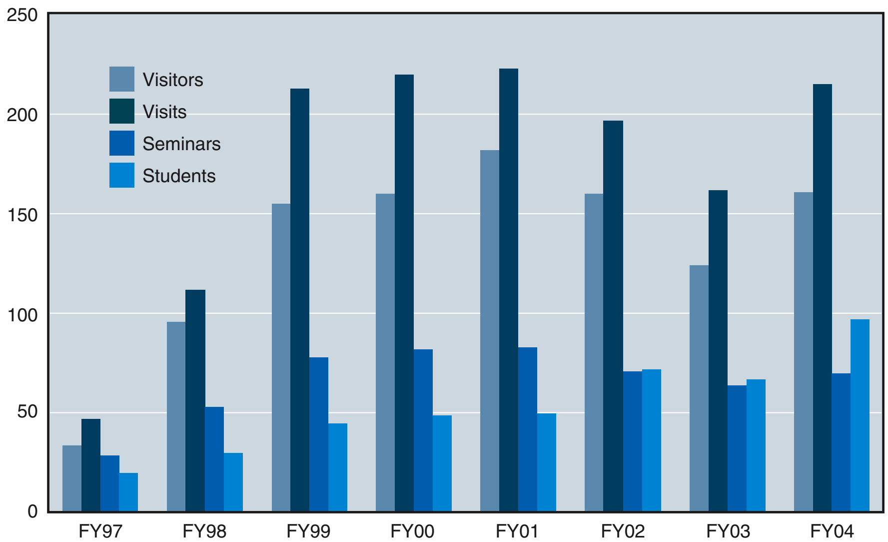
Table 3. ISCR's 38 participating guests for FY 2004.

Guest	Affiliation	LLNL Contact	Dates
Marian Brezina	University of Colorado	Rob Falgout	Dec. 1, 2000 – Nov. 30, 2004
Zhiqiang Cai	Purdue University	Charles Tong	Sept. 8, 2003 – Sept. 7, 2005
Praveen Chandramohan	Oak Ridge National Laboratory	Terence Critchlow	Nov. 7, 2003 – May. 6, 2004
Alok Choudhary	Northwestern University	Terence Critchlow	Sept. 15, 2003 – Sept. 13, 2005
Jennifer Dacles-Mariani	University of California, Davis	Garry Rodrigue	Sept. 1, 2003 – Aug. 30, 2005
Hans de Sterck	University of Colorado, Boulder	Rob Falgout	Oct. 1, 2003 – Sept. 30, 2004
Branden E. Fitelson	University of California, Berkeley	Terence Critchlow	Aug. 9, 2004 – Aug. 8, 2005
Franz Franchetti	Carnegie Mellon University	Kim Yates	June 4, 2004 – June 3, 2005
Alejandro Garcia	San Jose State University	Richard Hornung	Oct. 12, 2001 – Oct. 10, 2004
William Charles Gear	Princeton University	Steve Lee	Sept. 21, 2004 – Sept. 20, 2005
Matthew R. Gibbons	U.S. Airforce Academy	Bill Bateson	Aug. 9, 2004 – Aug. 8, 2005
Bernd Hamann	University of California, Davis	Mark Duchaineau	Aug. 9, 2004 – Aug. 8, 2005
Alan Hindmarsh	LLNL (retired)	Carol Woodward	Oct. 1, 2002 – Sept. 29, 2004
Martin Isenburg	University of North Carolina	Terence Critchlow	Nov. 4, 2003 – Nov. 3, 2004
Ken Joy	University of California, Davis	Mark Duchaineau	June 4, 2003 – June 3, 2005
Ramya Krishnamurthy	Oak Ridge National Laboratory	Terence Critchlow	Nov. 7, 2003 – Nov. 6, 2004
Johannes K. Kraus	University of Loeben	Van Henson	Nov. 30, 2002 – Sept. 28, 2004
Raytcho Lazarov	Texas A&M University	Panayot Vassilevski	Aug. 31, 2002 – Aug. 28, 2005
Oren Livne	Stanford University	Van Henson	Sept. 15, 2003 – Sept. 14, 2004
Bertram Ludaescher	San Diego Supercomputer Center	Terence Critchlow	Sept. 1, 2001 – Oct. 26, 2004
Jeannee Martin	LLNL (retired)	Bronis de Supinski	March 20, 2004 – March 19, 2005
Sally McKee	Cornell University	Bronis de Supinski	Oct. 14, 2001 – Oct. 13, 2004
Esmond Ng	Lawrence Berkeley National Laboratory	Edmond Chow	May 12, 2003 – May 11, 2004
Beth Ong	LLNL	Van Henson	Aug. 1, 2001 – Dec. 1, 2004
Peter Pacheco	University of San Francisco	Pat Miller	Aug. 9, 2004 – Aug. 8, 2005
Joseph E. Pasciak	Texas A&M University	Panayot Vassilevski	July 1, 2002 – June 29, 2005
Christoph Pflaum	Universität Erlangen-Nürnberg	Rob Falgout	Aug. 25, 2003 – Aug. 24, 2004
Elbridge Gerry Puckett	University of California, Davis	Louis Howell	June 30, 2003 – June 29, 2004
Markus Pueschel	Carnegie Mellon University	Kim Yates	April 14, 2003 – May 18, 2005
Ulrich Ruede	Universität Erlangen-Nürnberg	Rob Falgout	Aug. 15, 2000 – July 29, 2004
Paul E. Saylor	University of Illinois	Steve Lee	Aug. 22, 2004 – Aug. 21, 2005
Martin Schultz	Cornell University	Bronis de Supinski	March 1, 2003 – Dec. 30, 2004
M. Alex Schweitzer	Universität Bonn	Rob Falgout	Nov. 17, 2003 – July 22, 2005
Claudio Silva	University of Utah	Randy Frank	May 19, 2003 – June 3, 2004
Lansing Sloan	LLNL (retired)	Pete Eltgroth	June 16, 2004 – June 15, 2005
Christoph W. Ueberhuber	Techische Universität Wien	Kim Yates	March 1, 2003 – Feb. 28, 2004
Beata Winnicka	Argonne National Laboratory	Dan Quinlan	Aug. 15, 2003 – Aug. 14, 2004
Gabriel Wittum	Universität Heidelberg	Rob Falgout	Nov. 17, 2003 – Nov. 16, 2004

under two main themes: sponsored-research activities that stimulate interactions between academia and LLNL staff, and a diverse visitor program that enables both short- and long-term residential stays at LLNL.

ISCR oversees three different types of sponsored-research activities. The University Collaborative Research Program (UCRP), through the ISCR, funded seven research projects during FY 2004 at University of California campuses. These projects supported graduate students working on doctoral thesis research. The faculty principal investigators and students worked closely with an LLNL collaborator. The ISCR also coordinated the funding of 19 research subcontracts to various academic institutions throughout the United States. These contracts are normally funded by programs at LLNL to help address long-term Laboratory requirements. This type of vehicle is also used to fund sabbatical visits to LLNL for three to six months. Eleven faculty members spent at least a portion of their sabbatical leave here during FY 2004. With Laboratory Directed Research and Development (LDRD) funds, the ISCR also funds Exploratory Research in the Institutes (ERI). These research grants are awarded to LLNL staff with the goal of developing ties to academia through co-funded research projects. The ISCR oversaw three such projects in FY 2004. Annual summaries for LDRD projects, UCRP projects, and subcontracts can be found in the next three sections of this document.

Figure 1. ISCR Visitor Program FY1997–FY2004



In FY 2004, the ISCR continued its tradition of an extensive and diverse Visitor Program. This program includes sabbatical visitors, sponsored workshops, summer students, and various seminars featuring external speakers. Altogether, the ISCR hosted 215 visits from 161 different visitors, an average of more than 4 visits per week. The vast majority (67%) of the visitors were from academia, with 9% from industry, 20% from other federal laboratories, and 4% from non-laboratory-based government institutions. Visitors from outside the United States made up 25% of the total.

The ASC Institute for Terascale Simulation Lecture Series was established in 2000 to enrich the intellectual atmosphere of LLNL's large simulation community through the visits of leaders representing the diverse areas of computation. In FY 2004, we hosted five speakers in this series. The general ISCR seminar series included an additional 65 talks covering a wide spectrum of research areas. Titles of all of these talks can be found in the Seminar Series section of this report, and associated abstracts can be found on the accompanying CD-ROM.

During the summer, ISCR hosted 94 visiting students for a total summer student population of 104, including the SEGRF students in residence year-round. The summer program exposes students to the stimulating and challenging work environment of a national laboratory. Successful candidates are hired as summer employees, assigned individual LLNL mentors, and given specific projects to which they will contribute. The project is appropriate to the student's background and skills, and ranges from programming tasks to original research. The topical coverage of the summer research program broadens each year as computation expands into new scientific areas and as computational tools become more powerful and diverse. Scalable algorithms, radiation transport, genomics, terascale visualization, and computer security are just a handful of topics from last summer's lively hallway conversations at the ISCR. The summer program runs from May to September, with most participants spending 10–12 weeks on site. Project reports for most of the students can be found on the accompanying CD-ROM.

In June, with the advent of our large student summer program and sponsorship from the Defense Programs office of DOE Headquarters, we ramped up our fifth annual Summer Student Lecture Series—three different series on Computational Modeling at the Terascale, Computer Science at the Terascale, and Computer Security. Though the lectures were intended for students, permanent CASC researchers also attended. The 35 lecturers are listed in the Seminar section of this annual report.

Poster presentations were made by 41 ISCR summer students at the LLNL Student Research Symposium in August 2004. The event, held at LLNL's Central Café, attracted local media and scientific staff from across the Laboratory, as well as other students and summer research mentors. Overall, 139 students presented posters, making ISCR's share about 30% of LLNL's total. Topics of the posters ranged from Internet routing patterns to advanced gridding techniques for estuarine flow modeling; from performance modeling tools to wavelet-based compression of radiation opacities; from language interoperability to management of data in petabyte-scale file systems. Students ranged in seniority from community-college first-years to graduate students about to complete their doctorates.

Figure 1 charts the numbers of visitors and seminars over the past eight years. The number of students in residence in FY 2004 increased substantially due to the expansion of the ISCR's responsibility in LLNL's summer programs. CASC scientists mentored 55% of these students. Other LLNL organizations mentoring ISCR summer students were: CADSE, DCOM, EEBI, ICCD, NAIC, NIFE, and PAT. Some of these students elected to spend internships prescribed by their national fellowships at the ISCR, at no direct cost to the Laboratory, including DOE Computational Science Graduate Fellowship (CSGF) holders and Department of Homeland Security (DHS) fellows.

Finally, the ISCR sponsored or co-sponsored 10 scientific workshops in FY 2004. Two of these were hosted locally and exclusively by the ISCR; the rest were in cooperation with other organizations and held off-site. In each case, there is a vital LLNL interest and typically, several LLNL researchers participate. Reports on these workshops appear in a later section of this report.

Most of the raw material of this document comes directly from the visitors and principal investigators of the projects being reported upon. We thank Arnold Gatilao and Deanna Midtaune for their editorial work and Dan Moore of the Technical Information Department of LLNL for his graphic artistry in producing an easily navigable and visually pleasing document.

We hope that you enjoy examining this report on the ISCR's diverse activities in FY 2004. For further information about the Institute, please contact us at the address below. Inquiries about how you might enhance the ISCR program for FY 2005 and beyond are always welcome.

David E. Keyes

Institute for Scientific Computing Research
Lawrence Livermore National Laboratory
P. O. Box 808, L-561
Livermore, CA 94551
http://www.llnl.gov/iscr

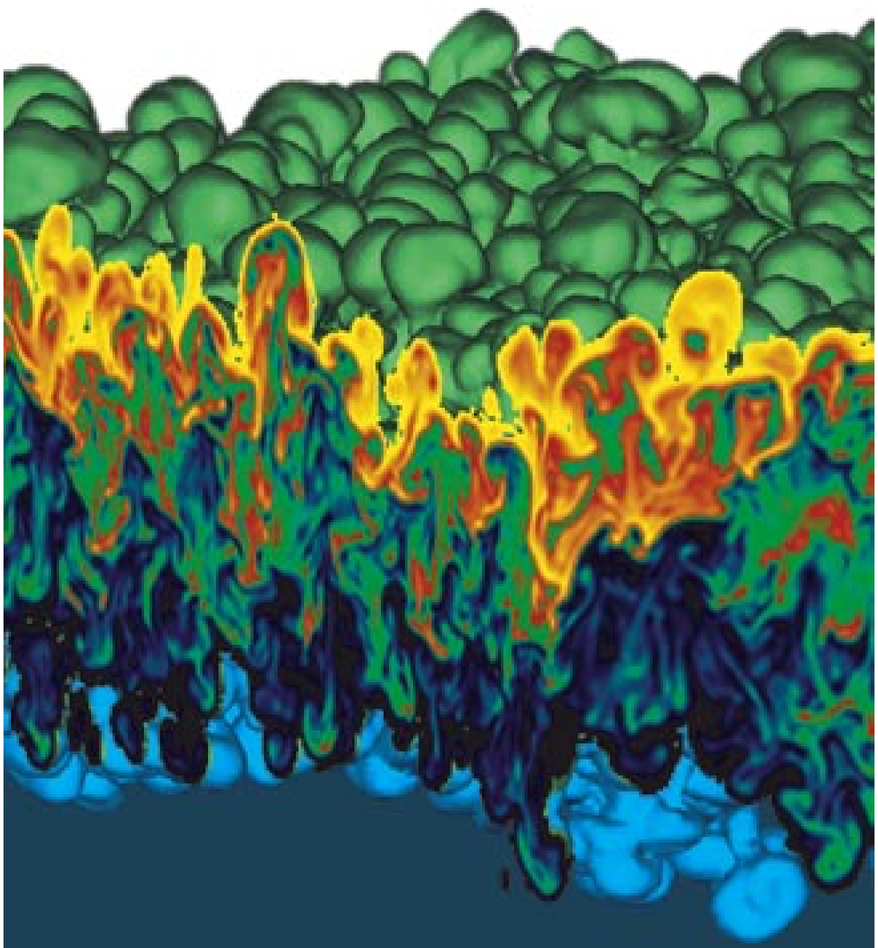
David E. Keyes, Acting Director
http://www.llnl.gov/iscr/people/keyes
dekeyes@llnl.gov
925-422-1325

James R. McGraw, Deputy Director
http://www.llnl.gov/iscr/people/mcgraw
jmcgraw@llnl.gov
925-422-0541

Linda Becker, Administrator
http://www.llnl.gov/iscr/people/becker
becker22@llnl.gov
925-423-0421

Laboratory Directed Research and Development (LDRD)

The ISCR supported three Exploratory Research in the Institutes (ERI) projects with Laboratory Directed Research and Development (LDRD) funds during FY 2004. These research grants are awarded to LLNL staff with the goal of developing ties to academia through co-funded research projects. Anticipating the emergence of data science as a cross-disciplinary theme, the ISCR has concentrated its efforts in developing technologies for large-scale and distributed data sets for the past several years. The current portfolio contains a project on providing a single interface to diverse data sources on the Web, a project on tracking objects in a succession of images, and a project in multiresolution representation of scientific datasets for scalability across a range of computer architectures. This portfolio, originally motivated by purely scientific applications, has already paid dividends in some of the Laboratory's new homeland security applications.



Enabling Large-Scale Data Access

Principal Investigators

Terence Critchlow and David Buttler, CASC

This project's goal was to develop an infrastructure capable of providing scientists with access to large numbers of data sources. To that end, we have developed an infrastructure that employs a user-specified description of a service of interest to crawl the Web. When an interface in the service class is identified, a wrapper that supports an XML-based query interface is automatically created. This year, we successfully demonstrated the application of our infrastructure on service classes from two vastly different domains — Basic Local Alignment Search Tool (BLAST) sequence similarity interfaces and interfaces to publication sites.

We have demonstrated the ability to automatically identify new data sources of interest while crawling the Web. We performed extensive testing of the infrastructure on hundreds of BLAST data sources and half a dozen publication sources. Next, we performed a minimal Web crawl, during which we identified and wrapped nine previously unknown BLAST sites and one new publication site. In addition, we identified approximately 100 Web pages that contained citations but were not queryable.

This work supports national security and other LLNL missions by benefiting ongoing programs at LLNL, such as Department of Homeland Security nonproliferation and detection efforts, that must utilize information from a wide variety of sources, including some that cannot be easily integrated using traditional techniques. Our infrastructure simplifies the process of creating an interface that combines local data with related information publicly available over the Internet, such as scientific publications.

We met and exceeded our FY 2004 milestones. In addition to demonstrating end-to-end automatic wrapper generation for BLAST interfaces, we extended the service class description to handle more complex data types, such as citations that cannot be easily represented by regular expressions. We performed several short Web crawls that successfully demonstrated the application of our infrastructure on both well-structured BLAST interfaces and poorly structured publication interfaces by identifying previously unknown sites in both domains. The expertise gained in this project will be applied to the Biodefense Knowledge Center Bio-Encyclopedia effort.

Publications

(1) D. Rocco, T. Critchlow. "Automatic Discovery and Classification of Bioinformatics Web Sources." *In Proceedings of the Georgia Tech Conference on Bioinformatics.* UCRL-JC-152980. 2003

(2) J. Caverlee, L. Liu, D. Buttler. "Probe, Cluster, and Discover:focused extraction of QA-Pagelets from the deep web." *In proceedings of the IEEE conference on Data Engineering.* 2004

(3) Wei Han. Ph.D. dissertation Georgia Institute of Technology 2003

(4) Dan Rocco. Ph.D. Dissertation, Georgia Institute of Technology 2004

Detection and Tracking in Video

Principal Investigator
Chandrika Kamath, CASC

Video cameras are used for monitoring and surveillance in several applications. We are developing robust, accurate, and near-real-time techniques for detecting and tracking moving objects in video from a stationary camera. This allows us to model the interactions among the objects, thereby enabling us to identify normal patterns and detect unusual events. Our algorithms and software include techniques to separate the moving foreground from the background, extract features representing the foreground objects, track these objects from frame to frame, and post-process the tracks for display. We focus on video taken under less-than-ideal conditions, with objects of different sizes moving at different speeds, occlusions, changing illumination, low resolution, and low frame rates.

The capability to detect and track in video supports the national security mission of LLNL by enabling new monitoring and surveillance applications for counterterrorism and counter-proliferation. This project will produce robust and accurate technology for video detection and tracking under less-than-ideal conditions with occlusions, fog or changing illumination, or at a low resolution or frame rate. This project will enhance existing algorithms to address these situations, allowing us to better understand their limitations, which in turn, will determine the conditions under which successful surveillance is possible. The algorithms and software are being applied to surveillance video, as well as spatiotemporal data from computer simulations.

During FY 2004, we

(1) Created a software infrastructure to handle streaming video data.
(2) Implemented several background subtraction algorithms and evaluated them on videos taken under different conditions.
(3) Proposed a new background subtraction method that outperforms other methods, especially on low-resolution, low-frame-rate video.
(4) Extracted features and used them in simple tracking algorithms.

We also filed a provisional patent on the new method and summarized our work in two papers. We are currently adapting the tracking algorithms to work under adverse conditions. We collaborated with the University of Colorado, Boulder, on tracking people, the University of California, San Diego, on tracking under occlusions, and a summer student on object representations for tracking.

Publications

Cheung, S.C., C. Kamath, "Robust background subtraction with foreground validation for urban traffic video," *Eurasip Journal on Applied Signal Processing,* UCRL-JRNL-201916.

Moelich, M., "Autonomous motion segmentation of multiple objects in low resolution video using variational level sets," UCRL-TR-201054.

Gyaourova, A., C. Kamath, S.C. Cheung, "Block matching for object tracking," UCRL-TR-200271.

Cheung, S.C., C. Kamath, "Robust techniques for background subtraction in urban traffic video," *Video Communications and Image Processing Conference,* **Vol. 5308**, pp. 881-892, UCRL-CONF-200706.

Cheung, S.C., "Robust techniques for background subtraction," UCRL-ABS-200371.

Our new technique for background subtraction (bottom) is less sensitive to changes in illumination in comparison with current techniques (top). The pixels highlighted in purple indicate the moving objects in the frame.

ViSUS: Visualization Streams for Ultimate Scalability

Principal Investigator

Valerio Pascucci, CASC

We are developing a suite of progressive visualization algorithms and a streaming infrastructure to enable interactive exploration of large scientific data sets. The methodology optimizes the data flow in a pipeline of processing modules. Each module reads and writes a multi-resolution representation of a geometric model, providing the flexibility to trade speed for accuracy, as needed. The data flow is streamlined with progressive algorithms that map local geometric updates of the input into immediate updates of the output. A prototype streaming infrastructure will demonstrate the flexibility and scalability of this approach for visualizing large data sets on a single desktop computer, a cluster of personal computers, and heterogeneous computing resources.

In FY 2004, we brought ViSUS to a level of maturity and robustness allowing direct deployment for a number of targeted users. The main milestones achieved include developing new techniques that accelerate isosurface extraction with occlusion culling, graphics hardware, and view-dependent refinements. A test viewer has been developed for datasets from the HYDRA simulation. We provided a stable library with full implementation of our streaming technology that can be used by simulation codes for saving rectilinear grids in ViSUS IDX format. We released a new version of the Progressive Viewer with full slicing, isocontouring, and volume-rendering capabilities. We are working in collaboration with the MIRANDA team to start using the IDX format as output of choice for the Blue Gene/L runs.

The ViSUS project benefits the Laboratory at least at two levels. At the deployment level, the improved efficiency in the use of hardware resources reduces the cost of visualization-hardware infrastructures. At the scientific level, the developed technology reduces the overall time required for the design, simulation, and visualization cycle. The ability to remotely monitor large and expensive simulations saves computing resources through early termination and restart of erroneous test simulations. Runtime steering will be possible for simulation codes with mechanisms for dynamic modification of running conditions.

Use of our innovative, high-performance visualization techniques allows interactive display of very large data sets on simple desktop workstations and the monitoring (or steering) of large parallel simulations. This will have valuable applications to several LLNL missions, including stockpile stewardship, nonproliferation, energy security, and environmental management, that use large-scale modeling and simulations.

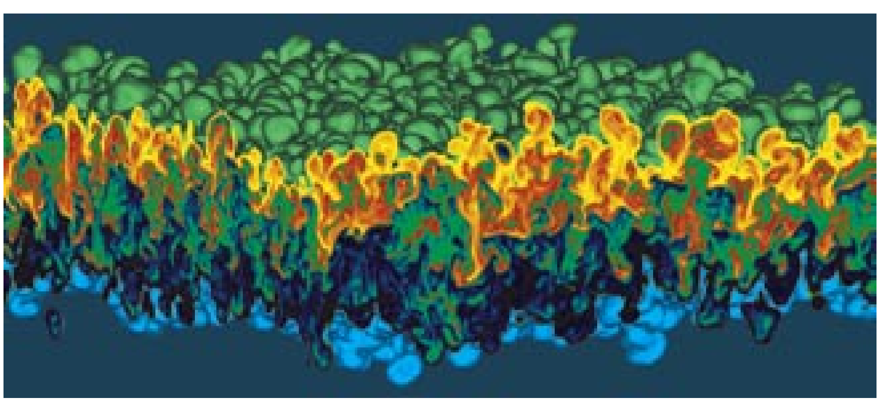
Publications

Laney, D. and V. Pascucci, "Progressive Compression of Volumetric Subdivision Meshes," *Proceedings of the International Symposium on 3D Data Processing, Visualization, and Transmission 2004*, pp. 293–300, UCRL-CONF-203679.

Pascucci, V. "Topology Diagrams in Scientific Visualization" Chapter in: *Surface Topological Data Structures: An Introduction for Geographical Information Science*, pp. 121–130, UCRL-200013-BOOK.

Pascucci, V., "Isosurface computation made simple: Hardware acceleration, adaptive refinement and tetrahedral stripping," *Proceedings of the Joint Eurographics - IEEE TVCG Symposium on Visualization*, pp. 293–300, UCRL-CONF-202459.

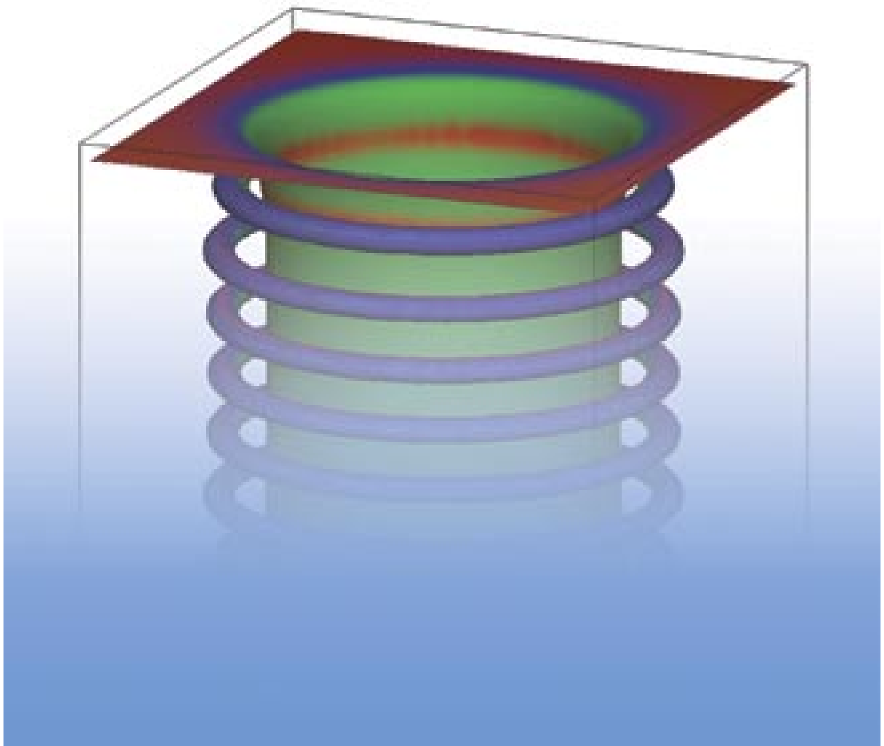
Van Kreveld, M.J., R.van Oostrum, C.L. Bajaj, V. Pascucci, and D.R. Schikore, "Efficient contour tree and minimum seed set construction" Chapter in: *Surface Topological Data Structures: An Introduction for Geographical Information Science*, pp. 71–86, UCRL-200018-BOOK.



The ViSUS Progressive Viewer generates a snapshot of the MIRANDA hydrodynamics code computing a Rayleigh–Taylor turbulent mixing of fluids.

University Collaborative Research Program

The ISCR supported seven University Collaborative Research Program (UCRP) projects during FY 2004 at the University of California campuses. These projects support graduate students working on doctoral thesis research, and the faculty principal investigators and students work closely with an LLNL collaborator. The projects in FY 2004 spanned four different UC campuses, as well as the departments of civil engineering, computer science, mechanical engineering, and physics. They ranged from understanding and optimizing the performance of advanced architecture computers to improving engineering simulation models; from the development of tracking and recognizing objects to basic computational science at the frontier.



Data-Driven Execution of Communication-Tolerant Algorithms

Principal Investigator **Scott B. Baden,** University of California, San Diego

Collaborator **Daniel Quinlan,** CASC

Communication-tolerant algorithms are expected to play an important role in achieving scalability on large-scale platforms with many thousands of processors. However, communication-tolerant application development is time-consuming and prone to error, even for the expert programmer. A run-time substrate that supports non-bulk, synchronous, data-driven execution is under investigation. Because it treats communication and computation as coupled activities, rather than as distinct phases of execution, the data-driven execution model naturally supports irregular or unpredictable communication delays that are expected on large-scale platforms, such as BlueGene/L.

The data-driven substrate programmer expresses an application in terms of partially ordered communication and computation operations, as constrained by the data dependences inherent to the application. The substrate invokes an externally specified scheduler to determine the precise task-execution ordering. The user may further constrain the scheduling by means of *performance metadata* decorating the graph. The user expresses partially ordered operations in terms of a directed graph called a *task graph*. Vertices in the task graph correspond to tasks and edges to data dependences between the tasks.

The substrate is implemented as a C++ library and runs as a background thread (i.e., a proxy) to manage dependence information and task scheduling. This background thread, called the Mover–Dispatcher, routes data from completed tasks to dependent tasks and determines when such tasks are enabled for execution.

When several possible tasks are ready for execution, their exact order is determined by the scheduler in conjunction with performance metadata. We built the communication proxy part of the Mover–Dispatcher, which supports non-blocking, asynchronous communication by means of multithreading. We are currently testing this capability on a latency-tolerant variant of a Gauss Seidel red-black Poisson solver in three dimensions using a Beowulf cluster in our research lab.

The interpretation of performance metadata is up to the application. For example, it might express priorities or affinities. We have tested out this capability on blocked LU running on a single processor. Our goal was to reorder computations in order to improve memory locality. The blocked LU application demonstrated the importance of an economical run-time representation for the task graph structure. The overhead of dynamically reordering task execution is 25% of the overall running time and an improved run-time representation for tasks graph representation is under investigation. We are exploring two optimizations. The first takes advantage of graph sparsity—most nodes depend on only a few other nodes. The second takes advantage of temporal locality. Rather than instantiate an entire task graph at initialization time, another approach is to create task graph nodes on demand and to recycle disused nodes. This strategy effectively throttles task generation and avoids tying up memory with task information that won't be needed for a long time.

Publications

S. B. Baden, "Moving forward in large scale computation" in J Dongarra, K. Madsen, J. Wasniewski (Eds.), 7th International Conference on Applied Parallel Computing (PARA '04), Lyngby Denmark, June 2004, *Lecture Notes in Computer Science,* Springer, 2004.

S. B. Baden, "Masking Latency with Data Driven Program Variants" in J Dongarra, K. Madsen, J. Wasniewski (Eds.), 7th International Conference on Applied Parallel Computing (PARA '04), Lyngby Denmark, June 2004, *Lecture Notes in Computer Science,* Springer, 2004.

Feature-Based Approaches for Long-Range Motion Segmentation and Object Tracking

Principal Investigator **Serge Belongie,** University of California, Davis

Collaborator **Chandrika Kamath,** CASC

Image segmentation is the problem of partitioning the pixels in an image into a relatively small number of regions that correspond to objects or parts of objects. It is one of the hardest (and oldest) open problems in computer vision, and it plays an important role in the process of object detection and recognition. As challenging and computationally intensive as image segmentation is, it also happens to be a problem that the human visual system solves effortlessly. The goal of this project is to develop methods for image and video segmentation with an emphasis on motion-based processing.

Our proposed work divides into two main areas: "what went where" and "who went where." The first area, on which we have already begun working, addresses the problem of motion segmentation for image sequences with large inter-frame displacement, e.g., more than 10% of the image width. The second area, for which the principal application area will be people tracking, deals with the problem of detecting and tracking moving objects that exhibit temporally periodic variation in appearance.

We first describe our solution to the "what went where" problem. The algorithm operates in two stages, starting with robust estimation of the underlying motion fields and concluding with dense assignment of pixels to motion fields. The first stage of this process is the first dense-motion segmentation method to operationalize the layer-based formulation for multiple discrete motions. We detect the motion layers using a variant of RANSAC (Random Sample Consensus) on detected interest points using a planar projective (2D homography) motion model. We perform the layer assignment using a fast, graph-cut-based MRF (Markov Random Field) formulation, which enforces spatially piecewise smooth-pixel assignments.

The case of moving objects that do not fit the planar projective model is addressed in the second stage. A poor fit to the planar projective model can arise from two main causes: non-planarity and non-rigidity. For this purpose, we refine the fit via iterative regression using a regularized thin-plate spline (TPS).

Our approach to the "who went where" problem builds on our approach to the first problem by interfacing it in a novel way with the classical tools of multiview geometry. In particular, we apply the knowledge that sets of frames depicting a periodically moving object in a given shape but with varying poses will (approximately) satisfy the multiview constraints for a rigid object. After automatically detecting the period of a given object in the field of view, we apply the "what went where" framework to segment the object from the background. The segmented foreground object across the period-separated frames is then treated as a conventional stereo- or multiview-reconstruction problem. Using this framework, we demonstrated a 3D reconstruction of sparse interest points on a video sequence of a pedestrian and showed preliminary results computing dense disparity using a graph-cut-based stereo correspondence engine.

Illustration of video object deletion. Original frames (top). Segmented layer corresponding to hand motion (middle). Reconstruction without the hand layer using the recovered motion of the keyboard (bottom). Note that no additional frames beyond the three shown were used as input.

Publications

Sameer Agarwal, Satya Mallick, David Kriegman and Serge Belongie, "On Refractive Optical Flow," *European Conference on Computer Vision*, Prague, Czech Republic, pp. 483–494, **vol. 2**, 2004.

Josh Wills and Serge Belongie, "A feature based method for determining dense long range correspondences," *European Conference on Computer Vision*, Prague, Czech Republic, pp. 170–182, **vol. 3**, 2004.

Serge Belongie and Josh Wills, "Structure from Periodic Motion," *International Workshop on Spatial Coherence for Visual Motion Analysis*, Prague, Czech Republic, 2004.

DNS and Modeling of Dispersion of Solid or Liquid Particles in Turbulent Flows

Principal Investigator **Said Elghobashi,** University of California, Irvine

Collaborator **Robert Lee,** Atmospheric Science Division

The release and subsequent transport of toxic materials into the atmosphere or inside and around buildings are of major concern to national security. These toxic materials can be in the form of gas, liquid droplets or solid particles and often undergo chemical reactions during the transport process. An important distinguishing feature of the transport of toxic materials into the atmosphere or inside and around buildings is the wide spectra of length scales and time scales involved. Length scales range from microns (particle size) to kilometers (turbulent eddy size) and the corresponding time scales range from microseconds to days. The approach of the proposed research is to employ direct numerical simulation (DNS) for prototypical particle-laden turbulent flows (e.g., flow over a backward-facing step) to evaluate the turbulent *correlations directly* and provide accurate models for them that can be introduced in LES (Large Eddy Simulation) or RANS (Reynolds-Averaged Navier–Stokes equations) codes.

The proposed work encompasses the numerical solution of the three-dimensional, time-dependent Navier–Stokes and continuity equations of the turbulent flow in addition to solving the equations of motion of the dispersed particles. The time-averaged correlations needed for RANS equations will be evaluated from their instantaneous values in DNS via ensemble and time averaging, and then models will be developed to relate these correlations to the dependent variables used in RANS equations.

Our work on the project during the past six months concentrated on the mathematical development of the necessary boundary conditions for the DNS code for the backward-facing step and performing short runs on a small parallel computer. We also spent considerable effort in converting our DNS parallel code that was running on a Cray-T3E so that it can run on a MCR computer (LLNL). A new machine-independent Fast Fourier Transform (FFT) was written and incorporated into the code. We are starting the DNS of the flow over the backward-facing step.

Numerical Study of Coexisting Superconductivity and Ferromagnetism: Applications to Real Materials

Principal Investigator **Warren E. Pickett,** University of California, Davis

Collaborator **Francois Gygi,** CASC

The discovery of ferromagnetic metals that become superconducting at lower temperatures (UGe_2, $ZrZn_2$, URhGe) early in this young century provided one of the big surprises in materials physics in recent years. For more than four decades, such coexistence was believed to be all but impossible. This coexistence of two competing phases, each a macroscopic manifestation of the quantum behavior of electrons, fully qualifies this as a new state of matter. The goal of this project is to include material-specific characteristics of the metallic states in the underlying formalism and to perform numerical studies to illuminate the microscopic driving mechanisms.

Our approach consists of three prongs:

(1) Reformulating the theory at its more basic levels to gain insight

(2) Adapting the formalism, developing algorithms and making the resulting codes applicable to real materials, such as those mentioned above, as opposed to the earlier study of model systems

(3) Computing solutions to the superconducting gap equations to map out phase diagrams.

Our work has moved to simulating the novel type of new coexisting ferromagnetic-superconducting state called FFLO, after its discoverers Fulde, Farrell, Larkin and Ovchinnikov. This state, which can occur in very weak ferromagnets or for paramagnets near the critical value of magnetic field, consists of the coalescence of superconducting pairs of electrons with a non-vanishing pair momentum. The resulting state and its order parameter are inhomogeneous in space. Our primary results are the identification of phase boundaries that specify the phase diagram.

The results of this year's research are:

(1) Identification of a characteristic velocity spectral function that is a fundamental part of the integral of the integral equation

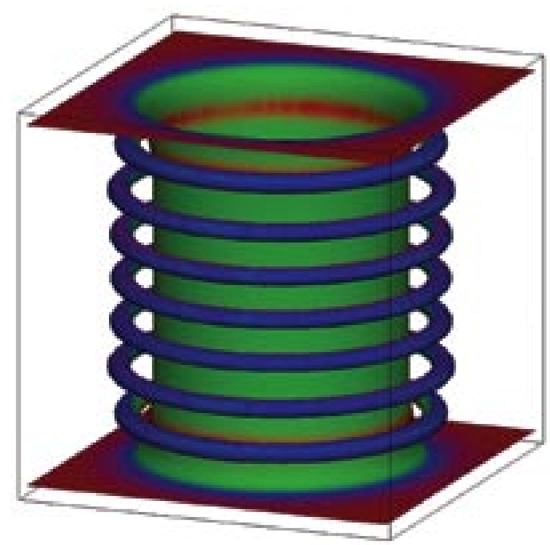
(2) Development and implementation of an algorithm to evaluate this function

(3) Showing how its spectrum affects the phase diagram and obtain predictions for a specific candidate, $ZrZn_2$.

During the past year, two new materials have been found to display some of the characteristics of the FFLO phase. These two systems are the peculiar rare-earth-based superconductor $CeCoIn_5$ and one of the first "heavy fermion" superconductors, UBe_{13}. If confirmed to be FFLO, these will add to the single system (the two-dimensional organic metal k-$(BEDT\text{-}TTF)_2Cu(NCS)_2$) that is relatively well established to enter an FFLO phase when a magnetic field is applied. One strong focus of the coming year will be to apply our new methods to these novel superconductors and help assess whether an FFLO phase really does arise. The computational facilities of ISCR are anticipated to be important for the success of the latter stages of this project.

Publications

A. B. Kyker, W. E. Pickett, and F. Gygi, "Fermiology and Fulde-Farrell-Larkin-Ovchinnikov Phase Formation," *Physical Review B* (submitted), 2004.



A current-carrying wire penetrates a slab of superconductor where it is electrically insulated, producing a radial magnetic field. Such a field, penetrating a type-II superconductor, produces regions (flux lines) of depressed superconducting strength, each containing one quantum of magnetic flux. For this geometry, the flux lines align regularly in rings, a generalization of the Abrikosov vortex lattice state in superconducting films. The color here gives the rate of change of the superconducting order parameter on the various isosurfaces (red is slow, blue is rapid).

Lagrangian Simulation of Penetration and Other Extreme-Deformation Events: Moving Beyond Meshless Methods

Principal Investigator **Mark Rashid,** University of California, Davis

Collaborator **Mike Puso,** Engineering

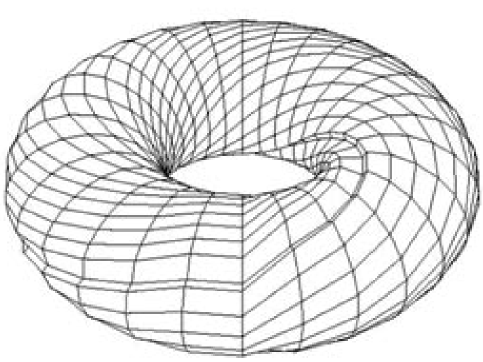
Solid mechanics problems involving extreme-deformation events, such as projectile/target interaction and fragmentation of cased explosives, present a considerable challenge to the conventional Lagrangian finite element method (FEM). Accordingly, beginning in the mid-1990s, various so-called meshless approximation methods have been proposed as possible alternative approaches. This research project involves the development of a variational approximation method that attempts to combine the best features of both the conventional FEM and meshless methods.

The technical approach involves a synergy of two distinct and innovative elements. In the first of these, the solid-mechanics problem is discretized using a flexible geometric subdivision scheme in which "elements" take the form not of hexahedra or tetrahedra but of arbitrary polyhedra. This flexibility vastly simplifies the task of mesh generation and thereby facilitates periodic remeshing, as is often required by extreme mesh distortion. The second element of the project involves a novel method for transferring information relating to the current material state from the old, distorted mesh to the new mesh following a remeshing cycle. This *remapping* step is required before the solution can be continued with the new mesh and can be a significant source of errors.

VETFEM is a finite-element-like Galerkin approximation method in which the basis functions and numerical quadrature are facilitated by a polyhedral partition of the material domain. In contrast to the conventional FEM, wherein all elements must conform to strict geometric and topological requirements, VET elements can take arbitrary polyhedral forms. The VET element formulation has been successfully extended to 3D, coded, and integrated into an existing research code. This code offers flexible and expandable capabilities, including explicit dynamics, implicit dynamics, quasistatics, strongly-objective (Rashid, 1993) finite-deformation one.

As with the VET element formulation, the extension of the existing 2D state remapping methodology to 3D turned out to be challenging. The fundamental algorithmic problem associated with the variational remapping method is that of partitioning the volume of a "new" tributary region among a tiling of "old" tributary regions with which it overlaps. The key innovation involves a highly efficient and robust optimization procedure that closely approximates the required volume partition without resorting to any cumbersome geometric calculations. This has been accomplished and the result is a highly efficient and robust piece of code.

A preliminary result is illustrated in figure, which shows a twisted toroidal mesh consisting of eight-node hex elements. A fairly complex function was defined on this mesh, and then it was rotated about its axis through 90 degrees in a series of steps. After each increment of rotation, the integration-point values of the function were remapped from the "old" mesh to the new one. Because the function itself possessed a periodicity of 90 degrees, the final, succession can be directly compared to the original ones. The error for different numbers of rotation increments is shown in the table. The performance of the new remapping method is being studied with this and other types of analyses.



No. of rotation increments	RMS error	Max error
1	0.013404	0.027977
2	0.0132459	0.0441313
3	0.0148609	0.0547436
4	0.0174637	0.0616309
5	0.0193817	0.0616309
6	0.0212833	0.07386956
7	0.0228581	0.0830038
8	0.0244043	0.087037
9	0.02591708	0.1009022
10	0.027899	0.166627

Twisted toroidal mesh is rotated 90 degrees in a series of steps, with remapping occurring after each increment of rotation. The error is shown for a range of the number of steps.

Publications

M.M. Rashid and M. Selimotic "A Three-Dimensional Finite Element Method With Arbitrary Polyhedral Elements," submitted, 2004

Memory Access Pattern Signatures and Certificates of Relevance for Benchmarks

Principal Investigator **Allan Snavely,** San Diego Supercomputer Center and University of California, San Diego

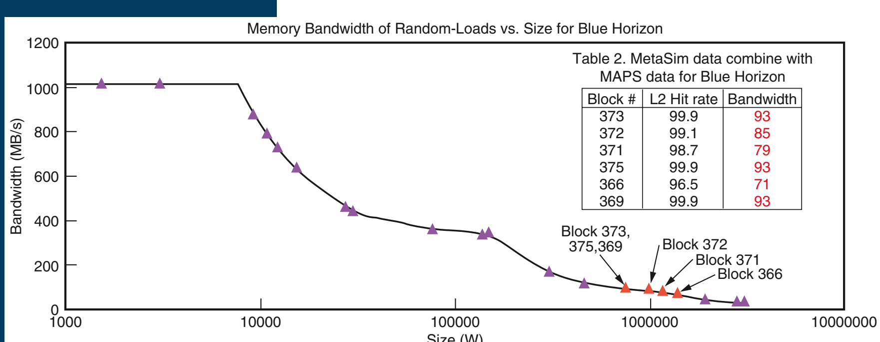
Collaborator **Bronis de Supinski,** CASC

This award has supported Michael O. McCracken in his fundamental research investigations into performance modeling and prediction as part of a collaboration involving the University of California, San Diego's Computer Science Department (where McCracken is a student and Allan Snavely an adjunct assistant professor), the San Diego Supercomputer Center's (SDSC's) Performance Modeling and Characterization (PMaC) laboratory led by Snavely, and LLNL's Bronis de Supinski. Michael has thus been able to participate in research relevant to both organizations and also to PERC (the Performance Evaluation Research Center), a DOE Office of Science ISIC.

So far our research has investigated the performance implications of memory access patterns and useful definitions of "signature distance" between the memory access patterns of different basic blocks from the same or different programs. The goal is to improve the accuracy and speed of the Convolution method for performance prediction. We have enhanced the functionality of the MetaSim tool for gathering memory access pattern signatures and have made this tool platform independent. We have been investigating what kinds of memory access patterns exist "in nature" and exploring the performance implications of memory access patterns. We have developed a nomenclature and symbolic representation of memory access patterns leveraging previous work by Nick Mitchell.

We have been examining definitions of "signature distance" between basic blocks—our expectation is that basic blocks with similar memory access patterns will perform similarly on a given machine. In work defining a meaningful metric for "signature distance," he established orthogonal properties of loops including memory footprint, memory access pattern, type and intensity of floating-point, and ILP (instruction level parallelism) operations that could provide "certificates of relevance" for benchmarks. This work has demonstrated that reasonably accurate estimates of an application's performance can be derived from its MetaSim profile and mapping its basic blocks to a similar set of benchmark basic blocks with an established performance profile. McCracken has published a paper showing how these performance model predictions can guide dynamic algorithm selection.

Currently McCracken is developing a study of the HPC Challenge Benchmarks that will position these, and several strategic DOE applications, in dimensions of spatial and temporal locality. This work was presented at the HPCS meeting at SC04 in Pittsburgh. Further direct collaboration with LLNL is resulting in the means to acquire the application profiles via static analysis using the ROSE compiler infrastructure. This approach can result in significant time-savings and is leveraging de Supinski's expertise in compiler technology. The idea is that much of the information currently acquired via tracing, such as operation types and counts, communication and memory access patterns, can be determined or at least reasonably estimated more rapidly via static analysis. Tracing will then be used simply to confirm compile-time information via sampling and to fill in information (such as loop bounds) that may in some cases be unknown at compile time. The goal is to speed up code profiling by an order of magnitude.



Horizon Machine Profile with MetaSim data.

Publications

M.O. McCracken, A. Snavely, A. Malony, "Performance Modeling for Dynamic Algorithm Selection," *ICCS Workshop on Performance Modeling and Analysis (PMA03)*, June 2003, Melbourne, Australia.

Visual Tracking and Recognition for Biometrics and Interactive Visualization

Principal Investigator **Matthew Turk,** University of California, Santa Barbara

Collaborator **Lenny Tsap**, Electronics Engineering Technologies

At the Four Eyes Lab in the Computer Science Department at the University of California, Santa Barbara, we have pursued research on visual tracking and recognition for biometrics and interactive visualization with notable progress and success. Computer vision is a promising and powerful sensing modality that can be used to unobtrusively track, model, and classify human appearance and behavior.

Our primary goals for the year were to make progress in

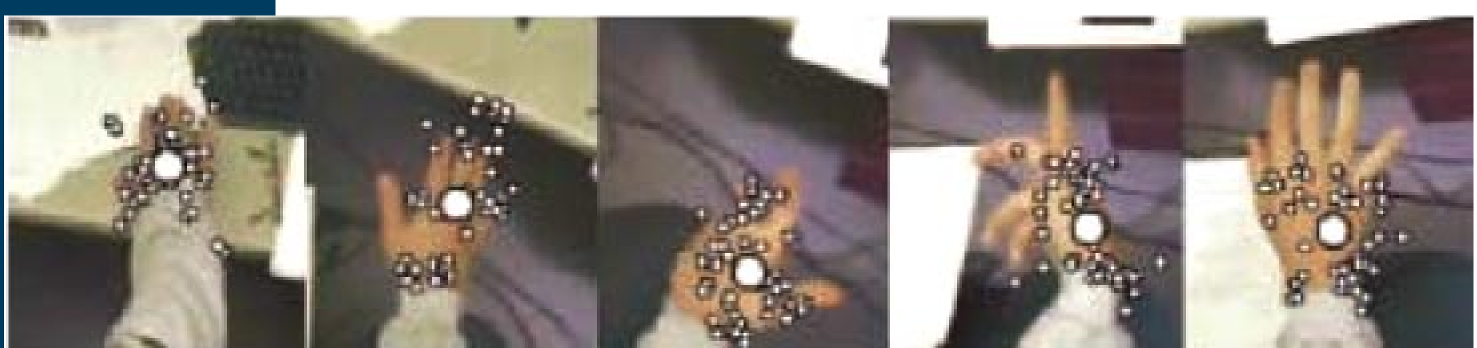
1. Face tracking and facial expression analysis
2. Hand detection, tracking, and gesture recognition
3. Extracting and using reliable depth edges in images and video.

These areas contribute to the overall goals of the project and give us a solid start in addressing the technical needs of biometrics and visualization problems. In face-tracking and facial-expression analysis, we extended our wavelet-based method to enable real-time tracking, and we used embedding techniques to represent high-dimensional face information in low-dimensional manifolds to investigate the representation and recognition of dynamic facial expressions [1, 2]. We implemented and tested real-time hand tracking and recognition and successfully applied it to mobile computing applications [3,4], winning a workshop Best Paper award. We also developed a technique for reliably extracting depth edges using a digital camera with multiple flashes and developed methods to produce non-photorealistic renderings of scenes and recognize "fingerspelling" letters (used in sign language) from images, which resulted in additional publications [5,6].

In the past decade, there has been an increasing interest in moving beyond the traditional applications of computer vision, such as robot navigation, object recognition, and industrial inspection, to using vision technology as an effective input modality in human-computer interaction (HCI) often pursued in the context of multimodal or perceptual interfaces. The general focus of these efforts is to integrate multiple perceptual modalities, such as computer vision, speech and sound processing, and haptic I/O, into the user interface. Such video-based sensing is passive and non-intrusive as it does not require contact with the user or any special-purpose devices.

The primary tasks of computer vision in these scenarios are to detect, model, recognize, and interpret various visual aspects of human behavior. If delivered reliably and robustly, such vision technology can support a range of functionality in

Hand tracking using the "Flocks of Features" method.



interactive systems by conveying relevant visual information about the user, such as identity, location, and movement, thus providing key contextual information. In order to fully support visual aspects of interaction, several tasks must be addressed, such as face detection and recognition, head and face tracking, facial expression analysis, eye-gaze tracking, body tracking, hand tracking, gait recognition, and recognition of postures, gestures, and overall activity.

Current biometric systems are based on processing images of a user's face, iris pattern, hand geometry, fingerprint, and other physical characteristics. As dynamic tracking and recognition technologies mature, behavior may well become an important aspect of biometrics – e.g., how people make facial expressions, how they move, their gaze and blinking patterns, and their postures. Leveraging these technologies will require a different approach to biometrics, with integration across modalities and across time (temporal integration) becoming central to the problem of verification.

Publications

1. C. Hu, Y. Chang, R. Feris, M. Turk, "Manifold based analysis of facial expression," *IEEE Workshop on Face Processing in Video*, Washington, D.C., June 2004.

2. Y. Chang, C. Hu, M. Turk, "Probabilistic expression analysis on manifolds," *International Conference on Computer Vision and Pattern Recognition*, Washington DC, June 2004.

3. M. Kölsch and M. Turk, "Fast 2D hand tracking with flocks of features and multi-cue integration," *IEEE Workshop on Real-Time Vision for Human-Computer Interaction*, Washington DC, USA, June 2004 (Best Paper Award).

4. M. Kölsch and M. Turk, "Analysis of rotational robustness of hand detection with rectangle features," *International Conference on Pattern Recognition*, Cambridge, U.K., August 2004.

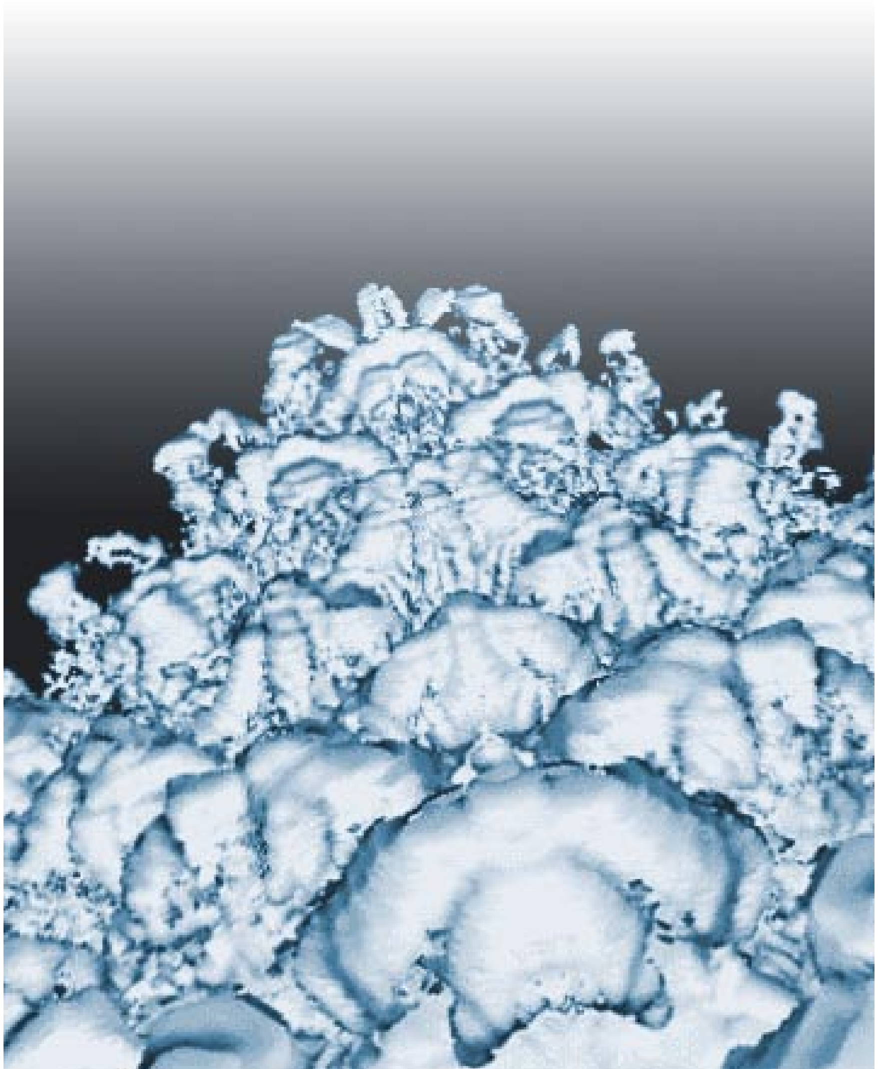
5. R. Feris, M. Turk, R. Raskar, K. Tan and G. Ohashi, "Exploiting depth discontinuities for vision-based fingerspelling recognition," *IEEE Workshop on Real-Time Vision for Human-Computer Interaction*, Washington DC, USA, June 2004.

6. R. Raskar, K. Tan, R. Feris, J. Yu, M. Turk, "A non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging," *ACM SIGGRAPH*, Los Angeles, August 2004.

Figure 2. A user interface employing hand gestures.

Research Subcontracts

The ISCR supported 19 research subcontracts to various institutions throughout the United States. These contracts are normally funded by programs at LLNL to help address long-term Laboratory requirements. These subcontracts typically fund residential visits by university faculty for close collaborations with scientists in the Computation Directorate. Brief reports follow detailing ongoing work enabled by these subcontracts.

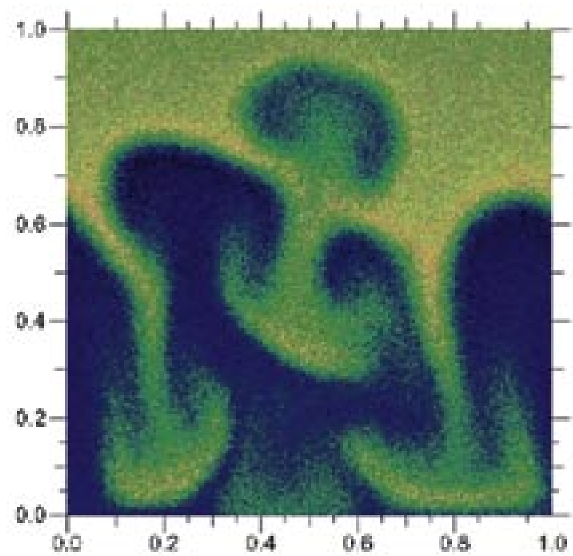


Nanohydrodynamic Simulation of the Rayleigh–Taylor Instability

Principal Investigator **Berni Alder,** University of California, Davis

Collaborator **Garry Rodrigue,** CASC

During the past year, we collaborated with a group at Los Alamos National Laboratory (LANL) and successfully carried out molecular dynamics simulations of the Rayleigh–Taylor instability that quantitatively predicts both the early-time behavior given by linear stability analyses and the long-time behavior given by experiment, as well as some of the numerical solutions of the continuum Navier–Stokes equation. This was published in *Proceedings of the National Academy of Sciences* and helped establish why different long-time behavior was observed in various continuum calculations.



Demonstration of the Rayleigh–Taylor instability. The color variation goes from yellow at the top for the highest density to dark blue for the lowest density.

This is the first time that three-dimensional particle methods quantitatively simulated turbulent mixing. In the biggest and only molecular dynamics run, a qualitatively different asymptotic behavior was found, namely the heavy fluid formed drops instead of coalescing into a single slug. This could be real and appear only in particle simulations due to the presence of fluctuations, which could cut the thin thread by which the mushrooms are connected to the tip of the spike.

In order to firmly establish this phenomenon, which is of great importance to groups such as the NIF and because the mixing region advances more slowly (i.e., time to the first instead of the second power), we have developed an alternative particle algorithm that is 100 times faster and allows us to study even bigger systems for longer times. This was accepted for publication in *Molecular Physics*. To show the power of this scheme, we have attached a graph of the Rayleigh–Taylor instability at fairly long times with a few million particles that were produced on a single–processor computer in a few hours.

What these preliminary results indicate is the necessity to confirm them by carrying out the biggest Rayleigh–Taylor calculation possible on the IBM BlueGene/L machine for which the new particle algorithm is particularly well suited. We also need to carry out a comparison to the best continuum solution to see the quantitative effects of the nonlinear terms and fluctuations left out in the Navier–Stokes equations.

Multiple Animal Tracking in the Smart Vivarium

Principal Investigator **Serge Belongie,** University of California, San Diego

Collaborator **Chandrika Kamath,** CASC

A common trend in 2D object recognition is to detect and leverage the use of sparse, informative feature points. The use of such features makes the problem more manageable while providing increased robustness to noise and pose variation. In this work, we extend these ideas to the spatiotemporal case. For this purpose, we show that the direct 3D counterparts to commonly used 2D interest-point detectors are inadequate and propose an alternative. Anchoring off of these interest points, we devise a recognition algorithm based on spatio-temporally windowed data. We present promising recognition results on a challenging real-world database of mouse behaviors.

We attempt to extend the above approaches developed for object recognition to the problem of behavior recognition, i.e., from the spatial to the spatiotemporal domain. These extensions are not always direct, but rather follow the general spirit of using sparsely detected features for object recognition. We propose to characterize behavior through the use of spatiotemporal feature points.

For the above purpose, we show that the direct 3D counterparts to commonly used 2D interest-point detectors are inadequate and propose an alternative. We extend descriptors for spatial interest points to cuboids of spatiotemporally windowed data. Cuboids extracted from a number of sample behaviors are clustered and the resulting cluster centers serve as a dictionary of atomic units of behavior we call "actons." The only information kept from all subsequent video data is the location and type of the actons present. We argue that such a representation is sufficient for recognition and robust with respect to variations in the data. We show an application of this framework, utilizing a simple classification scheme, to a challenging real-world database of mouse behaviors.

Given these feature vectors and labelled training data, we trained two different classifiers: linear discriminant analysis and a DAG support vector machine using a radial basis kernel. Both were applied after dimensionality reduction using principal components analysis. As mentioned in each case,six clips were used for training and the seventh for testing. We limited the amount of training data for the over-represented categories, such as exploring, otherwise categories such as drinking would be drowned out. However, we used all the testing data. The SVM classifier outperformed LDA by a few percentage points. The classifiers achieved an accuracy ranging from about 75% to 95%.

Robust Trajectory and Appearance-Based Data Association for Multi-Object Tracking

Principal Investigator **Michael C. Burl,** University of Colorado, Boulder

Collaborator **Chandrika Kamath,** CASC

Reliable tracking is clearly a prerequisite for more advanced video-mining operations. This study builds upon our previous work in multi-object tracking from surveillance-style video sequences. The primary goal was to improve upon the tracker's ability to assign and maintain an object's unique track ID for the duration of time that the object appears in the scene.

The general structure of our previous tracker consisted of six steps.

1. Background estimation and subtraction
2. Thresholding and spatial grouping of detected pixels into blobs
3. Gating and data association to link blobs with tracks
4. Birth of new tracks if there are unassociated blobs
5. Measurement update (incorporating new observations into the state estimates for each of the tracks)
6. Time update (predicting the position and appearance of the tracked objects at the next observation time).

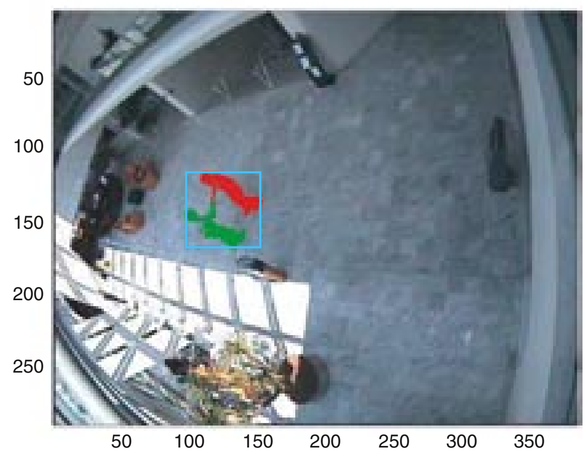
A key problem with this type of tracker is that the blobbing process is not perfect and leads to over-segmentation (a single object split into multiple blobs) and under-segmentation (multiple objects lumped into a single blob). Subsequent steps in our previous tracker could not properly deal with these fundamental grouping errors.

In this study, we investigated a number of ideas that we believed could improve tracking performance. A two-pronged approach was pursued with part of the effort devoted to getting more information from an object's trajectory and part of the effort devoted toward getting more information from an object's appearance.

The following component ideas were considered.

- Compensating for perspective effects
- Using more refined dynamical models
- Using decision-theoretic local search (structured change detection)
- Establishing precise correspondence between parts of a tracked object from one frame to the next, splitting and merging blobs
- Using particle filters to overcome some of the limitations caused by hard data association decisions.

Although we explored these ideas in some detail, we were not able to integrate these disparate components together into a unified tracker within the allotted time, nor were we able to conduct systematic evaluations over benchmark data sets to characterize the overall impact on performance. However, we plan to continue working toward these two goals.



Two people meet and shake hands. The object detection component initially groups the two people together into a single blob (cyan bounding box), but the tracker is able to use the appearance and trajectory information from the previous frames to correctly split the blob into two individuals.

Multi-Constraint and Multi-Objective Partitioning-Complex Networks

Principal Investigator **Umit Catalyurek,** Ohio State University

Collaborator **Edmond Chow,** CASC

The aim of this project was to get better insight into partitioning properties of complex networks. With the advances in high-performance computing and high-resolution sensors in many fields of the computational sciences, data sizes are getting bigger and bigger. The sizes of complex networks that arise in homeland security applications are no different than any the other applications. Both the size and on-demand query and analysis requirements of applications that use complex networks necessitate both parallel and distributed computing. To achieve good performance, data needs to be partitioned in a load-balancing manner and communication requirements needs to be minimized.

Starting with random networks, we investigated the performance of 1D and 2D coarse-grain hypergraph partitioning of various networks provided to us by LLNL. Using the Performing Tools for Hypergraph (PaToH) partitioning tool, we analyzed the change in two cost metrics—graph edge cut and hypergraph connectivity-1—with respect to several graph parameters and the number of partitions. We also derived theoretical relationships between those parameters and the cost metrics.

In the experimentation phase for the type of the networks that have a generator, we generated multiple graphs with different parameter values, such as average degree and number of vertices. For each graph, we first used a random partitioner to generate a partition that would be used as a baseline to evaluate the performance of PaToH. Next, we partitioned these graphs using 1D and 2D coarse-grain hypergraph models using PaToH as our partitioner. Finally, we compared the performances of those partitioning schemes and explained the results in accordance with the theoretical predictions.

Our results showed that both 1D and 2D partitioning achieve much better results than the random partitioning, even though both approaches produce similar communication volume requirements. However, in applications that are sensitive to the number of communications, partitioning via 2D coarse-grain decomposition achieves much better results, providing better performance. Searching through paths in complex networks, especially for the homeland security applications that motivate our work, provides a practical real-world scenario where partitioning using 2D coarse-grain decomposition would be advantageous.

Numerical Simulation of Dispersion of Solid Particles Moving at Supersonic Speeds in Turbulent Flows

Principal Investigator **Said Elghobashi,** University of California, Irvine

Collaborator **Bill Bateson,** CASC

This work aims at developing a mathematical model that predicts the trajectories of spherical solid particles moving at supersonic speeds in a turbulent flow. The mathematical model involves the numerical solution of the three-dimensional, time-dependent, compressible, Navier–Stokes and energy equations. The project comprises two main tasks to be performed in parallel. Each task will require four years for completion.

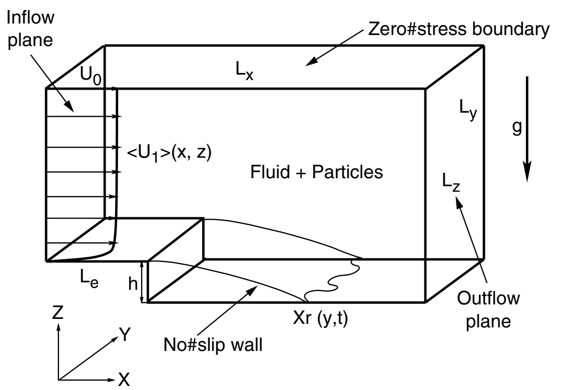


Figure 1. Schematic of the computational domain.

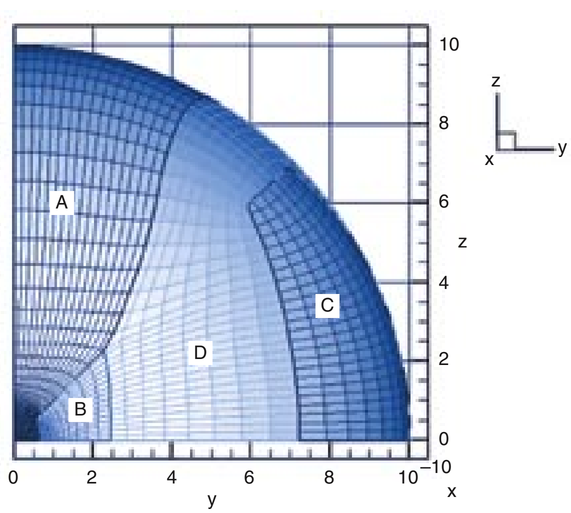


Figure 2. Grid surface ZY plane view

Task 1. The objective of this task is to compute the unsteady 3D compressible flow around a single-fixed sphere subjected to an air flow of a speed about 5000 m/s. An existing code at University of California, Irvine has been modified for a laminar compressible flow to account for the presence of a fixed sphere. Integrating the forces imparted on the sphere surface provides the values of the drag and lift on the sphere. The effects of varying the Mach number and particle Reynolds number will be investigated. A differential equation of motion of a sphere at supersonic speeds ($M > 1$) will be developed based on the computed drag and lift forces.

Task 2. The objective of this task is to compute the 3D turbulent isotropic compressible flow laden with many spherical solid particles moving at supersonic speeds in a cubical domain. The motion of each particle will be governed by the differential equation mentioned above in Task 1. A mathematical model for the collision between the particles will be included in the particle motion equation.

The detailed report describes the mathematical approach and the results of an incompressible flow over a single sphere. The computed values of the drag coefficient agree well with those of Kim, et.al. for the range of Reynolds numbers $Re = 20$ to 100.

Analysis and Visualization of Scientific Data Sets Using Generalized Segmentation Methods

Principal Investigator **Bernd Hamann,** University of California, Davis

Collaborator **Valerio Pascucci,** CASC

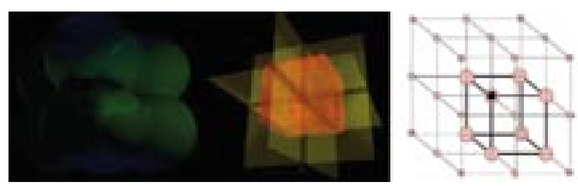
Segmentation of scientific data sets is becoming increasingly important in the context of understanding massive numerically simulated data sets at higher levels of qualitative behavior. Topology is one area in mathematics that allows us to characterize the behavior of scalar-valued data sets based on analyzing critical points and their relationships. This effort developed new techniques and program modules that support the automatic and mathematically sound segmentation of scalar-valued data sets into subsets (subregions of the physical domain) of topologically (and thus qualitatively) distinct behavior.

Traditional volume rendering of a data set involves visualizing surface properties or utilizing a transfer function. An alternative approach visualizes topological information, supporting a more qualitative understanding. The Morse–Smale complex is a structure that represents the topology of a scalar data set. The Morse–Smale complex can be simplified for use in applications, such as topological smoothing or hierarchical data visualization.

We have developed algorithms for extracting and simplifying Morse–Smale complexes that rely solely on combinatorial decisions, therefore avoiding numerical instability. We extract a Morse–Smale complex from a scalar field defined over a tetrahedral mesh using a multi-stage, region-growing approach. We also simplify a Morse–Smale complex through the cancellation of critical point (zero-gradient-point) pairs. There are two types of cancellations—saddle point–extrema cancellations and 1-saddle–2-saddle cancellations. We have developed rules for determining when a cancellation is valid and rules for reconnecting a simplified complex.

A multi-resolution representation of a Morse–Smale complex can be obtained through a simplification hierarchy. We have defined rules for valid simplifications of a complex and developed a hierarchy based on independence of cancellations. We declare two cancellations to be independent when their affected areas in the Morse–Smale complex do not overlap. We also represent a hierarchy as a directed acyclic graph that encodes the independence of cancellations. A multi-resolution reconstruction of the complex can be obtained by cutting the hierarchy graph.

Additional work has been done under this subcontract for error estimation in the context of rendering large multi-resolution data sets. We have developed an algorithm to estimate the error associated with changing resolution levels in a data set stored as a multi-resolution hierarchy. Our method utilizes a pre-processing step that calculates intermediate resolution levels based on error to allow for interactive modification of transfer functions used in a volume rendering application. Our approach balances storage overhead cost and quality of error estimates in support of an efficient method for choosing the subset of a volumetric data set that will be rendered in a near-optimal way.



Iso-surface of test data set, stable and unstable manifolds, and associated Morse–Smale complex.

A saddle-maximum cancellation (left) and a 1-saddle–2-saddle cancellation.

Element Agglomeration AMGe for Contact Problems

Principal Investigators **Ana H. Iontcheva and Randolph E. Bank,** University of California, San Diego

Collaborator **Panayot S. Vassilevski,** CASC

The goal of this project was to develop algebraic multigrid methods for solving constrained-minimization problems, mainly contact problems in linear elasticity, discretized on general unstructured meshes, using the finite element method, in particular

- Signorini's problem
- Two-body contact problem
- Obstacle problem

For the solution of the Signorini's problem—contact of a linearly elastic body with a rigid frictionless foundation—we have developed two multilevel algorithms

- Multilevel subspace minimization algorithm
- FAS-constrained optimization algorithm

These schemes utilize element agglomeration coarsening away from the contact boundary, which allows for a straightforward construction of coarse-level approximations that automatically satisfy the fine-grid constraints. The two algorithms provide monotone reduction of the energy functional throughout the multilevel cycle. A code (in C++) has been developed and a paper has been written for the solution of the Signorini's problem. The results from this paper have been extended to 3D.

For the solution of the two-body contact problem, contact of two linearly elastic bodies in a mortar-method-based algorithm is the target of our future research and development. It will utilize the already developed methods for a single-body Signorini's contact problem.

Hypre for Symmetric Generalized Eigenvalue Problems

Principal Investigator **Andrew Knyazev,** University of Colorado, Denver

Collaborator **Charles Tong,** CASC

Symmetric eigenvalue problems are crucially important in structured mechanics and electronic structure calculations. The goal of the project is to implement a modern eigenvalue solver into Hypre in order to take advantage of Hypre high-quality preconditioners for parallel clusters with a user-friendly interface. The Locally Optimal Block Preconditioned Conjugate Gradient (LOBPCG) method for symmetric eigenvalue problems has been chosen for implementation as one of the most promising eigensolvers in the current literature. The LOBPCG apparently combines the fast convergence of the block's Lanczos method with the simplicity and robustness of the block's steepest descent and allows the preconditioner to be plugged directly into the eigensolver without any inner-outer iteration schemes. The result is a well-written native Hypre C implementation of LOBPCG that works in three Hypre interfaces and is almost as efficient as the Hypre PCG linear solver.

The basic functionality of the code has been tested on University of Colorado, Denver and LLNL clusters on a number of test problems, including non-self-consistent electronic structure calculations eigenproblems, and the code is checked into the Hypre Alpha. Presently, this is the first and only preconditioned eigenvalue solver for parallel clusters.

An entirely new implementation of the LOBPCG for eigenvalue problems has been written as a part of the Hypre, consistent with Hypre Krylov-based linear solvers. The distinctive features of the new LOBPCG code are as follows.

1) The code is rewritten from scratch in order to implement the **lopbcg.c** code at the same level of abstraction as the Hypre **pcg.c** code.
2) The LOBPCG now works for three Hypre interfaces—IJ, Struct and SStruct—and includes test drivers for every interface. As of August 31, 2004, the LOBPCG drivers reflect the latest changes in the corresponding Hypre drivers and are prepared to serve as combined drivers for both linear systems and eigenvalue problems.
3) A new Hypre type, hypre_MultiVector, is introduced, and a preliminary implementation of Hypre MultiVector functions based on the existing parallel vector types (hypre_ParVector, hypre_StructVector and hypre_SStructVector) is used in the LOBPCG code. A significant future effort by the Hypre team is necessary to turn it into an actual MultiVector with efficient implementation of Hypre MultiVector functions by eliminating redundant MPI calls and by using Basic Linear Algebra Subprograms (BLAS).
4) The code solves generalized eigenvalue problems, as well as regular eigenvalue problems.
5) The code allows the use of constraints.

Preconditioning of Finite-Element Saddle-Point Problems

Principal Investigators **Raytcho D. Lazarov and Joseph E. Pasciak,** Texas A&M University

Collaborator **Panayot Vassilevski,** CASC

The goal of this research is the efficient solution of the discrete equations that result from finite-element or finite-difference approximation of partial-differential equations of mathematical physics. This involves the development and analysis of algorithms especially tailored for execution on medium- to large-scale parallel computing platforms. Our approach involves the application of theoretical tools from the analysis of partial differential equations to motivate and analyze new computational algorithms.

For implementation of highly accurate methods and efficient algorithms, we need approximation methods that provide greater flexibility in the grid generation process, increase the portability of various approximation methods and computer implementations, enhance the capabilities of coarsening strategy in parallel algebraic multi-grid methods, and provide a natural and practical way for parallel domain decomposition methods and parallel adaptive methods based on a posteriori error analysis. Our investigations produce competitive algorithms that can be used in various codes for complex applications in physics and engineering.

Main results of the research

1. A new, inexact Newton algorithm for the solution of second-order problems with higher-order nonlinearities was proposed and analyzed. This approach was based on a stability analysis in Sobolev spaces of order greater than one.
2. A new multigrid algorithm was proposed and analyzed for an electromagnetic problem. This method involves strengthening a curl-curl term by adding a discrete "grad-div" term. The resulting two-level multigrid algorithm was analyzed.
3. A stabilization framework for the discontinuous Galerkin finite-element method was developed. It was implemented for Raviart–Crouzeix non-conforming finite elements and tested on three-dimensional problems of elasticity.
4. A new stable scheme of exponential fitting type was proposed and studied for convection-diffusion problems.

Biological Fluid Flow in Micro-Electro-Mechanical Systems

Principal Investigator **Dorian Liepmann,** University of California, Berkeley

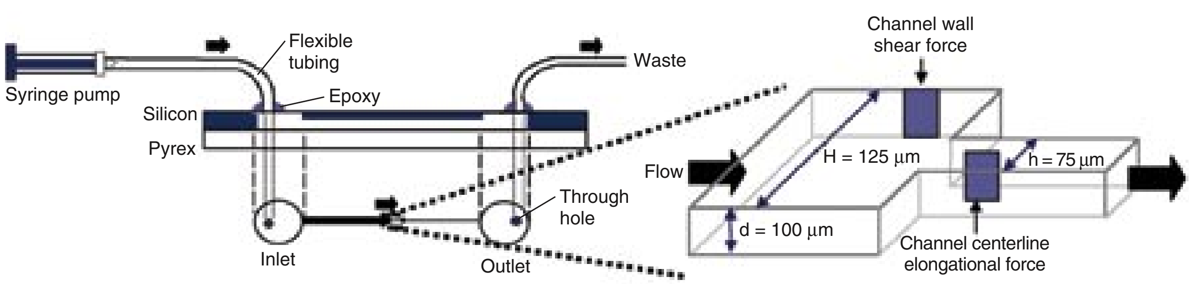
Collaborator **David Trebotich,** CASC

Our investigation aims to characterize the flow behavior of biological macromolecule solutions in silicon microfluidic devices. Digital Particle Image Velocimetry (DPIV) is used to quantify the velocity fields of DNA-laden solution flow under microfluidic conditions. Deviations from Newtonian flow fields brought about by the viscoelastic fluid rheology, concentration effects and conformational changes of the molecules will be assessed for flows through a variety of microfluidic geometries, such as straight channels, contractions, and expansions. Experimental results will be used to validate a computational design tool for bio-detection microdevices being developed at LLNL.

The characterization of flows containing high concentrations of macromolecules is critical for the optimal design of microfluidic systems for biochemical analyses. Since the fluid flow behavior at the microscale differs from the macroscale, a simple scaling down of processes may be insufficient to elicit the same performance. This thereby motivates research directed toward the characterization of the microflow behavior of the biological macromolecule DNA, commonly used in biochemical assays. Studies have shown that λ-DNA can undergo dramatic stretching in microfluidic flow with conformational changes occurring in both elongational and shear flows (Shrewsbury et al. 2001). The device geometry and flow rate, as well as the viscosity, concentration and relaxation time of the solution, influence the conformation of the DNA.

A 1.67:1 microfabricated abrupt planar contraction design was chosen as a test device (see figure). Experimental pressure drops for 450 μg/ml DNA (semi-dilute) solutions across abrupt contraction were measured. Velocity fields for flows of Newtonian and semi-dilute DNA solutions through abrupt contraction and abrupt expansion (inverted contraction device) were quantified using DPIV. No flow regions were observed near the contraction corners at slower flow rates and there is visual evidence of recirculation at faster flow rates.

Validation of the pressure testing system was conducted by measuring water flow through a narrow tube and comparing to theoretical solution. Pressure drop data was collected for flows of Newtonian (water) and non-Newtonian (DNA-laden) solutions through gradual and abrupt contraction devices. The effects of DNA solution concentration were explored and the onset of elastic, non-Newtonian behavior determined. DPIV was used to quantify flow fields for water flow in straight rectangular channels. A comparison study was conducted to determine the minimum interrogation region size for accurate data processing.



A microfabricated abrupt planar contraction design test device.at a ratio of 1.67 : 1.

Enabling Large-Scale Data Access

Principal Investigator **Ling Liu,** Georgia Institute of Technology

Collaborator **Terence Critchlow,** CASC

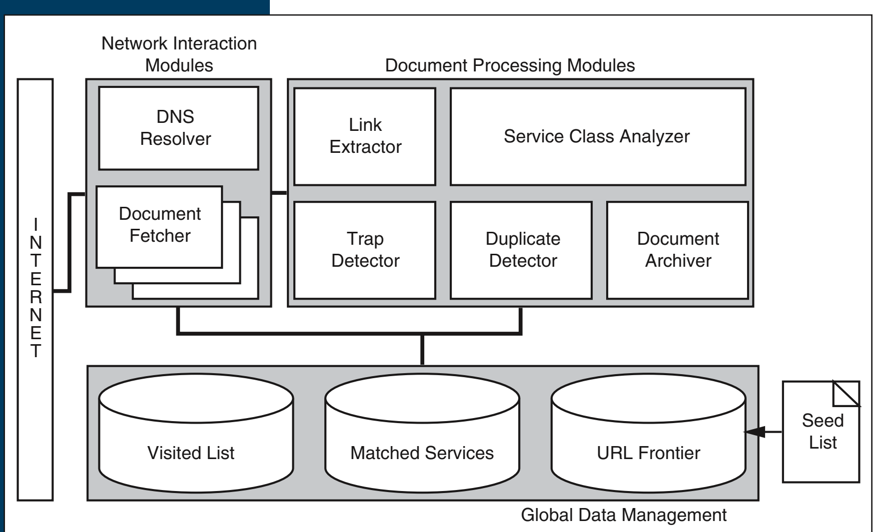
One of the ultimate goals of the Enabling Large Scale Data Access project is to produce a fully automated, end-to-end wrapper code generator through the design, development, and integration of service class descriptions with the XWRAP systems. The main idea is to provide mechanisms that enable XWRAP systems to take a generic description of a class of search interfaces (service classes), as well as the URL of a particular interface that is a member of the given class, to produce a functional wrapper that will take a class-specific query and produce XML views of the query results obtained through this particular interface. In the second year of this project, we focused on three main efforts.

(1) Extend the Web Spider developed under the previous subcontract to generate a set of Java 1.4 wrappers for interfaces that it identifies as matching a service-class description (SCD). Collectively, these wrappers will query all valid combinations of user input as specified by the SCD and the parameters enumerated by the interface. Each wrapper will be able to handle multiple types of errors (e.g., timeouts, connection refused, page not found) and return an appropriate message to the calling routine. The new version of the Web Spider is called DynaBot.

(2) In the DynaBot development, we have updated the Web-crawling component of the spider to be consistent with the interface identification component, updated the httpUnit library, and resolved JavaScript errors that prevent it from being used against some popular bio-portal sites, such as the National Center for Biotechnology Information's Basic Local Alignment Search Tool (BLAST) interface.

(3) We have tested the spider's ability to interact with the interface identification component by crawling the Web, starting at the dbCAT Web page and comparing the interfaces it identifies and the wrappers it generates with those identified and generated manually. This includes identifying and removing all obvious problems with the previous implementation of the spider, especially those identified in the second-year effort.

DYNABot System Architecture



The components that make up a crawler are divided among three major component groups: network interaction modules, global storage and associated data managers, and document processing modules. The simplest crawlers require mechanisms for retrieving documents and determining if a particular URL has been seen. More advanced crawlers will include features like mirror-site detection and trap-avoidance algorithms. DynaBot utilizes an advanced crawler architecture for source discovery and adds a document processor that can determine if a dynamic Web source is related to a particular domain of interest. The architecture of Dynabot is shown at left.

For a more detailed description of this research, see the full report at: **http://disl.cc.gatech.edu/LDRD/**

Large Graph Visualization

Principal Investigator **Kwan-Liu Ma,** University of California, Davis

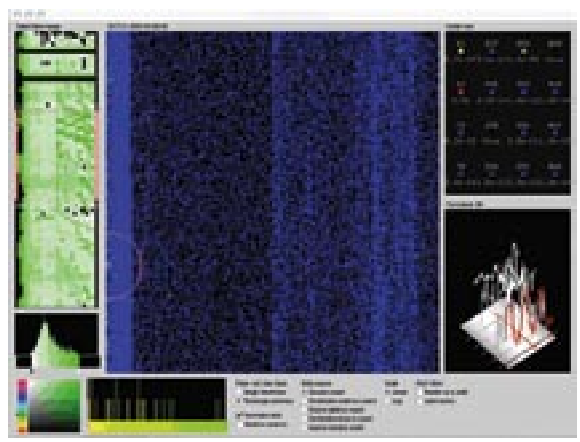
Collaborator **Marvin Christensen,** NAI

The Computer Incident Advisory Capability (CIAC) currently collects a great deal of data from sensors placed on a handful of computer networks. These sensors record activity on the networks so that the recorded information can be analyzed to provide useful information about both normal and anomalous activity on the networks.

This project focuses on one of the key significant technical challenges for interpreting this data. That challenge is a task in data mining—given the huge pool of data provided by the sensors, how can the interesting features be identified and classified? We address the problem by constructing a visualization system implementing novel methods of data analysis and information display. The process of exploring the space of a large data set using visualization is known as "visual data mining." It takes advantage of the pattern recognition facilities of the human visual system to detect patterns and anomalies in visual representations of abstract data.

We have developed a prototype system for visualizing TCP port data. The system permits analysts to discover the presence of any network security event that causes significant changes in the activity on ports. Since we currently only have access to very high-level data, the system is a very high-level tool and is useful mostly for uncovering high-level security events. Security events that consist of small details—an intrusion that includes only a few connections, for instance—are unlikely to be caught using this prototype system. Furthermore, since we have only obtained counts of activities rather than records of the activities themselves, the analysis can only go so far. The system can help identify suspicious traffic patterns, but it cannot see the traffic that caused the patterns. This is still useful, however, because analysts using this system can send the suspicious traffic signatures to analysts that have access to the full set of network traffic logs. The figure displays the current system interface.

Even in settings where only generalized information is available concerning network activity, we found many types of malicious activity can still be discovered using visualization. We have developed a tool that takes general, summarized network data and presents multiple, meaningful perspectives of the data and have demonstrated that this visualization leads to useful insights concerning network activity. Port scans of several types have been successfully detected, and many suspicious traffic patterns on individual ports have been uncovered. In addition, useful information about overall network traffic has been revealed, such as the rhythm of the traffic on commonly used ports as time progresses and the relationships between the various metrics used to describe port activity.



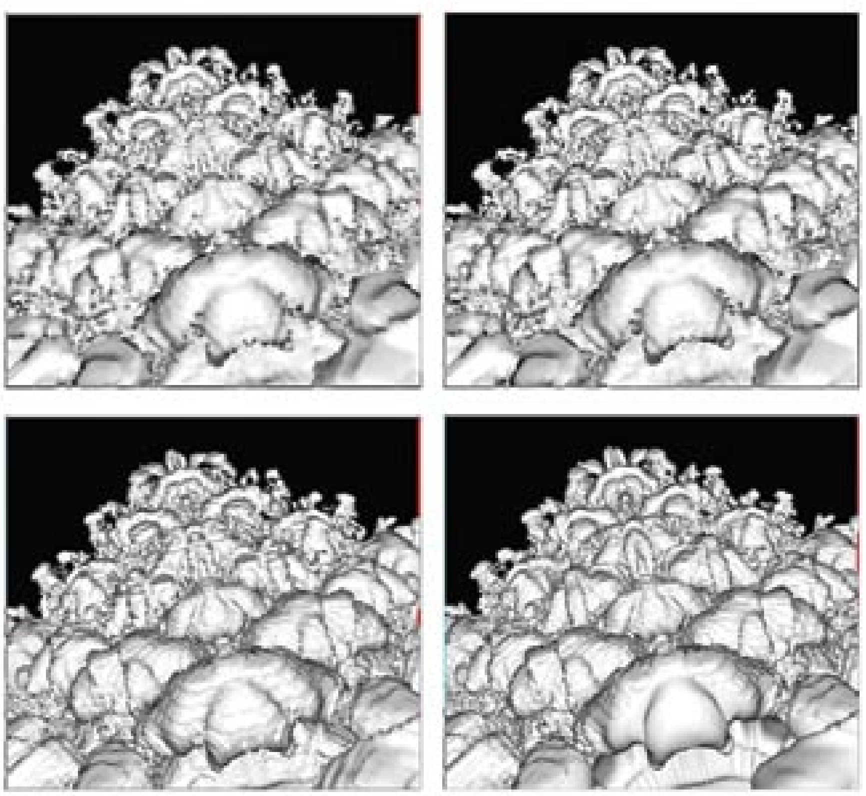
User interface of the TCP port data visualization system. The interface supports a drill-down process by allowing the analyst to explore high-level to lower-level information, from left to right, respectively.

Multi-Resolution Interactive Rendering of Large Scientific Data Sets Using an Image Cache

Principal Investigator **Kwan-Liu Ma,** University of California, Davis

Collaborator **Valerio Pascucci,** CASC

An example of the same scene rendered at different data resolutions. It is clear to see that as data resolution increases, images quality is greatly improved.



Large-scale simulations running on parallel supercomputers can produce very high-resolution information of the model phenomena, giving scientists the power to examine the phenomena at greater fidelity. Visualization enables scientists to better understand large amounts of data from such simulations. The problem with visualization is that conventional methods fail to handle the large amount of data. The data can take up too much system memory, as well as a considerable amount of time to render. Scientists are in need of tools that interactively search for and discover regions of interest within their simulations. Waiting for a long batch-mode visualization is not an option. In this project, we aimed at developing an interactive visualization solution enabling scientists to freely and effectively explore their simulation data.

This project began as an attempt to provide a system that is capable of rendering large multiresolution data sets while taking advantage of a cluster computer. Our work involved the extension of a multi-layered image cache system (MLIC) developed at LLNL in several ways. First, the rendering engine was modified to access data sets stored in the Visualizations Schemes for Ultimate Scalability (ViSUS) cache oblivious to multi-resolution representation. This has improved the performance of the rendering engine with minimal loss in accuracy. Second, we removed the reliance on the Visualization Toolkit (VTK), which has greatly simplified the system implementation and its performance. Next, the system now runs on LLNL's new cluster computers, making it possible to use high-resolution output devices like a large PowerWall display. Finally, the user interface was revamped to be more intuitive and easy to use. The new system gives an easy method of navigating through a large dataset while simultaneously creating high-quality visualizations.

The extended MLIC system is a capable multi-resolution rendering system. It not only retains but enhances system interactivity, which allows a user to navigate through a dataset while it is being rendered. Rendering can be stopped and settings can be modified before completion, allowing a scientist to make changes on the fly. The new rendering classes have made the system render independent, which allows for the use of almost any rendering algorithm or rendering software.

As seen in the figure, the system can render a very coarse-resolution image, as well as a clear high-resolution image and several levels in between for the large Richtmyer-Meshkov turbulence simulation. This allows a scientist to have an idea about what is going to be rendered and possibly make changes. This is made possible by decoupling the rendering and displaying of images, allowing the system to always remain responsive even as the rendering is taking place. Responsiveness leads to easy navigation.

Open Source Software Technology for Transforming Scientific Problems

Principal Investigator **John Mellor-Crummey,** Rice University

Collaborator **Dan Quinlan,** CASC

Two leading software systems for source-to-source transformation and optimization of scientific programs are the ROSE infrastructure being developed at LLNL and the Open64/SL infrastructure being developed at Rice University. These efforts are complementary. The Open64/SL infrastructure has focused on software support for parsing and transforming Fortran 90-based programming models. ROSE has focused on software support for parsing, analysis and transformations of C++ based programming models. This project will build an open-source software technology to bridge these two infrastructures to allow ROSE's transformation capabilities to be applied to Fortran 90-based programs.

The project began with a study of Open64's WHIRL infrastructure, a study of LLNL's ROSE infrastructure (including its SAGE IR) and the construction of test programs to familiarize ourselves with ROSE. Currently, we have begun constructing a framework for the translation between WHIRL and ROSE. We have begun adding a C++ namespace to the ROSE code so that it can interoperate with the Open64 infrastructure and WHIRL. The translator itself is being constructed by modifying code that performs a WHIRL-to-WHIRL transformation to compile Co-array Fortran programs.

A Perl script will orchestrate the WHIRL-to-ROSE translation. This script will invoke mfef90 (the Open64 front end for Fortran 90) on a Fortran 90 source file to generate a file that contains the WHIRL representation source. The WHIRL external representation will be read by the translator proper (under construction), which will walk the WHIRL abstract syntax tree and translate each WHIRL operator into its corresponding operator in the SAGE IR. Once in the SAGE IR, ROSE-based tools can be applied to transform the resulting code. Next, the process will be reversed. The SAGE internal representation will be traversed and WHIRL will be reconstructed. The reconstructed WHIRL will be written out in a file using its external representation. Finally, the driver script will invoke Open64's whirl2ftool to regenerate Fortran from WHIRL.

We have been using Los Alamos National Laboratory's Parallel Ocean Program (POP), the National Academy of Sciences (NAS) parallel benchmarks, and the Adjoint MIT Ocean General Circulation Model to exercise Open64's support for analysis and transformation of Fortran. Currently, we are working with Open64's infrastructure on the parsing and regeneration of the POP code. With the exception of a minor problem unparsing arrays of structures, Open64 regenerates legal Fortran 90 code for POP.

Over the last several months, we have resolved a half dozen issues in Open64's support for Fortran, including correcting the very-high level WHIRL intermediate representation produced by the Fortran 90 parser, correcting unparsing of formatted I/O, correcting unparsing of array initializations. Currently, work is under way to correct problems with the constant pool produced in the WHIRL intermediate form and to produce a general infrastructure for correctly unparsing hierarchies of structures and arrays of structures.

Once we resolve known correctness issues, we will begin assembling a suite of codes for regression testing of Open64.

A Tightly Coupled Particle–Fluid Model for DNA-Laden Flows in Complex Microscale Geometries

Principal Investigators **Gregory H. Miller,** University of California, Davis
P. Colella, D. T. Graves, D. F. Martin, P. O. Schwartz, LBNL

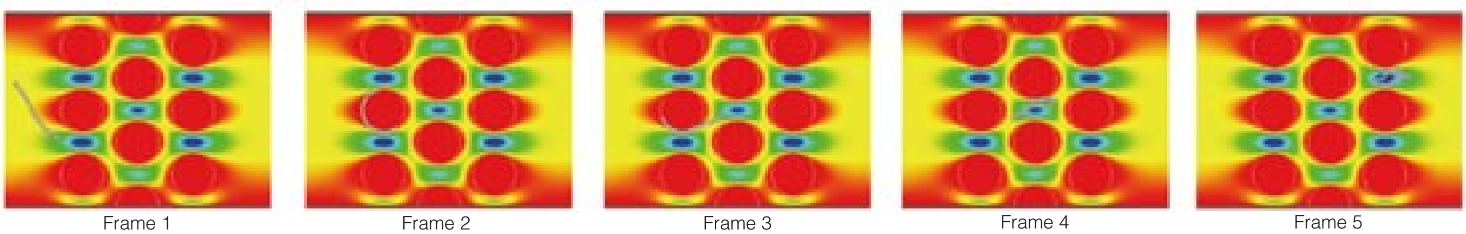
Collaborator **David P. Trebotich,** CASC

We present a stable and convergent method for the computation of flows of DNA-laden fluids in microchannels with complex geometry. The numerical strategy combines a ball-rod model representation for polymers tightly coupled with a projection method for incompressible viscous flow. We use Cartesian-grid embedded boundary methods to discretize the fluid equations in the presence of complex domain boundaries. A sample calculation is presented showing flow through a packed array microchannel in 2D.

Approximate DNA in an actual microfluidic device is used for extraction. The left boundary condition is plug flow with a velocity of 0.1 cm/s; the right boundary is outflow (homogeneous Neumann); the top and bottom boundaries and the interior circular boundaries are a solid wall. The polymer is a 26-node approximation of DNA introduced near the left boundary as an inclined linear array after the fluid-flow field reached steady state. The polymer's trajectory causes it to become wrapped around the first circular element, where it remains pinned until the stochastic perturbations work it loose.

The fluid dynamic steps of this method are subject to an advective Courant–Friedrichs–Lewy stability condition only. The particle steps, without constraints, are also stable with this CFL timestep. When particles move far from the constraint manifold, however, the Lagrange multiplier algorithm of Ciccotti, et al. may diverge. We have found that the maximum particle displacement per time step for which the Ciccotti, et al. algorithm is stable may be extended for most systems by centering the constraint force at the conclusion of the timestep instead of the beginning.

We use a backward Euler time-stepping strategy that is formally first order accurate. To make the overall method second-order, it will be necessary to replace at least this with a Runge–Kutta time-stepping strategy.



The time sequence of genomic DNA flowing in a 2D model of a packed-bed reactor PCR chamber. The DNA molecule enters from the left (frame 1), then wraps around a bead (frame 2), is loosened by hydrodynamic and Brownian forces (frame 3) and is swept out of the chamber by the flow field (frames 4 and 5). The color map indicates underlying flow field.

Clustered and Hierarchical Networks

Principal Investigator **Alex Pothen,** Old Dominion University

Collaborator **Edmond Chow,** CASC

We surveyed clustered and hierarchical representations of large networks available in the literature and made progress on two aspects of clustered graphs during the period of this contract.

First, we designed a new algorithm for discovering clusters in large networks that pays attention to the **local density** of the network. The density of a subgraph is the ratio of the number of edges joining nodes in the network to the maximum number of such edges possible (i.e., a complete subgraph or clique). In a large network, some subnetworks have high density, while others have low density, and a clustering algorithm that takes into account the local density would be able to do a better job of clustering.

In order to describe our algorithm, we need the concept of a ***k*-core**, a maximal subgraph in which every vertex is incident on at least k edges that belong to the subgraph (i.e., the degree of each vertex in the subgraph is at least k, when only edges in the subgraph are considered). The core number of a vertex is the maximum value of k as it belongs to a k-core, and the corresponding core is the maximum core to which the vertex belongs. The new clustering algorithm weights each vertex with the density of the connected component of the maximum core to which it belongs. The clustering proceeds by choosing a vertex with the highest weight to be the "seed vertex" for a cluster and then successively adds vertices whose weights are not too different from it to the cluster. After one cluster is complete, we begin with a new cluster in the same fashion.

This clustering algorithm has two advantages over existing ones. First, the weight of a vertex depends not on its degree but on the density of its neighborhood, thus clustering low-degree vertices together with high-degree vertices to which they are adjacent. Second, the algorithm can be efficiently implemented in time $O(K\,|E|)$, where K is the maximum core number of any vertex, and $|E|$ is the number of edges, in the network. We have implemented the weight computation phase of the algorithm and are currently working on the clustering phase.

Second, we are implementing a clustering algorithm for bipartite networks in which vertices in both parts need to be clustered. The algorithm relies on a singular vector of the vertex edge incidence matrix of the network. We followed the suggestion of Ravi Kannan, Santosh Vemula, and colleagues to speed up the singular-vector computation by computing with a submatrix of the incidence matrix. The columns in the submatrix are chosen with a probability proportional to the column weight. Then, the computed singular vector can be shown to be close to the singular vector of the larger matrix with high probability. Our preliminary results on the original incidence matrix shows that this approach works well, especially when combined with a combinatorial "local-smoothing" algorithm.

Spectral AMGe, ALE3D & FOSPACK

Principal Investigator **John W. Ruge,** Front Range Scientific Computations, Boulder, CO

Collaborator **Rob Falgout,** CASC

The majority of the work performed in FY 2004 has been on FOSPACK, a package for the discretization and solution of First-Order System Least-Squares (FOSLS) formulations of a range of PDEs. Algebraic Multigrid (AMG) is incorporated as the linear equation solver, and some of the effort has also gone into extending the range and improving the efficiency of that method. Progress has been made in several of the proposed areas of study and is summarized below.

AMG for Constraints

Slide-surface constraints were added to a two-dimensional, two-body elasticity problem. The constraints specified that two faces in contact remained in contact under application of a force, coupling the normal displacements along the two faces (and allowing for slippage in the tangential direction). Tests were conducted incorporating the constraints either strongly by essentially eliminating rows/columns from the matrix, or weakly by including them as weighted penalty terms in the functional to be minimized. Both resulted in an asymptotic V-cycle convergence factors that were nearly identical to those obtained without the constraints (where the two bodies were fully decoupled).

Parallel FOSPACK

Initial parallelization of 3D FOSPACK and coupling to Hypre/BoomerAMG has been completed for linear PDE systems with weak boundary conditions. The code has been ported to a 128-processor Beowulf cluster, and improvements in implementation resulted in better parallel performance. Treatment of nonlinear problems and coupled nonlinear multi-physics systems has also been improved. Tests performed on a coupled nonlinear fluid-elastic model and Maxwell's equations problem have resulted in solutions consistent with those obtained from the scalar FOSPACK code. The major effort for 3D FOSPACK is the introduction of higher-order discretizations, which are necessary for obtaining reasonable accuracy for 3D problems while keeping storage requirements within practical limits.

AMG Algorithm Enhancements

High-operator complexity can be a problem for AMG, especially in large 3D applications. In such cases, the coarse grid operators can be too large or too dense. Aggressive coarsening and multi-pass interpolation have been implemented in both the serial and parallel codes, resulting in markedly lower complexity, although some degradation of convergence is noted. A number of alternatives are being tested in the scalar code in an effort to both reduce complexity while retaining (or even improving) convergence.

Several variations of the basic AMG algorithm are being studied in an effort to improve its robustness. The use of compatible relaxation (CR) for choosing the coarser grids, combined with interpolation designed to minimize the trace of the coarse grid operator, looks quite promising. For some problems, the standard AMG interpolation is not accurate enough, particularly those with near-null space components. For this, an adaptive approach, in which interpolation is explicitly modified to better match the smoothest (slowest-to-converge) components, is being tested. The current approach is a hybrid of the recently developed adaptive Smoothed Aggregation (aSA) and a more straightforward adaptive AMG (aAMG).

Development of Efficient and Robust Algorithms for the Numerical Solution of Partial Differential Equations for High-Speed Reactive and Nonreactive Flow

Principal Investigator **Donald W. Schwendeman,** Rensselaer Polytechnic Institute

Collaborator **William Henshaw,** CASC

The work carried out under this subcontract involved the development and use of an adaptive numerical method for the accurate calculation of high-speed reactive and nonreactive flows on overlapping grids. In the reactive case, the flow is modeled by the reactive Euler equations with various choices for the equation of state and reaction rate model.

A numerical method has been developed to solve the nonlinear hyperbolic partial differential equations in the model. The method uses an unsplit, shock-capturing scheme, a Godunov-type scheme to compute fluxes, and a Runge–Kutta error control scheme to compute the source term modeling the chemical reactions. Two approximate Riemann solvers are now available and can be used for a general (non-ideal) equation of state. An exact Riemann solver is also available for an ideal, gamma-law equation of state.

Adaptive mesh refinement (AMR) has been implemented in order to locally increase grid resolution. The numerical method may also be applied to nonreactive flow problems, in which case the reactive source term is simply set to zero. The numerical method uses composite overlapping grids to handle complex flow geometries in either two or three dimensions. The code is part of the Overture–OverBlown framework of object-oriented codes, and the development has occurred in close collaboration with Bill Henshaw and Lori Diachin, and other members of the Overture group within the Center for Applied Scientific Computing.

During the period of this subcontract, a number of tasks were accomplished, including

- Further development of the numerical method for 2D moving meshes
- Implementation of additional Riemann solvers, including a Harten–Lax–van Leer (HLL) approximate Riemann solver for a general equation of state and an exact Riemann solver for an ideal equation of state
- An extension of the numerical method to handle axisymmetric flow
- An extension of the numerical method to handle multi-material, non-reactive flow

Workshops and Conferences

The ISCR-sponsored or co-sponsored 10 scientific workshops in FY 2004. Two of these were hosted locally and exclusively by the Laboratory. The rest were hosted in cooperation with other organizations, such as the Society for Industrial and Applied Mathematics (SIAM), the Institute for Pure and Applied Mathematics (IPAM), the Department of Homeland Security, or Argonne National Laboratory and held off site. Some ISCR workshops are one-of-a-kind exploratory workshops that assemble experts to scope out possible new programs. Others have become part of the fabric of their disciplines and are held at regular intervals. In each case, there is a vital LLNL interest and typically, several Laboratory researchers participate.

The 2003 International Conference on Preconditioning Techniques for Large Sparse Matrix Problems in Scientific and Industrial Applications

Dates October 27–29, 2003

Location Napa, California

The 2003 International Conference on Preconditioning Techniques for Large Sparse Matrix Problems in Scientific and Industrial Applications (Preconditioning 2003) was the third conference of its kind to focus on preconditioning techniques for solving sparse matrix problems. The first conference took place in Minneapolis, MN, in June 1999 and the second in Tahoe City, CA, at the end of April 2001. The first conference drew close to 100 participants and the second about 70 participants, while the third drew about 80 participants.

One of the characteristic themes of the meeting is its emphasis on real-life ("industrial") problems. For this reason, all three conferences enjoyed a healthy balance between academia and industry/government labs in its mix of participants. In addition, there is a rather important contingent of participants from overseas (mainly Europe). The preconditioning meetings have been quite successful and are now being viewed by the community as the premier specialized conference on preconditioners.

The Napa conference featured 7 invited speakers, 29 contributed papers and 8 posters. One of the goals of the meeting is to foster dialogue between practitioners and academics. As a rule, the organizing committee gives a charge to the program committee to nominate invited speakers with a goal of keeping a good balance between the number of talks on "methods" and those on "applications." The organizing committee then finalizes the selection to reach this goal. This particular meeting reached a good balance. There were four invited talks on applications and three invited presentations on algorithmic aspects.

One feature that distinguished this meeting from the previous two was that we allowed more parallel sessions. Previously, few (Tahoe City) or zero (Minneapolis) parallel sessions were scheduled. This change was necessary to increase the number of presentations as there were many high-quality abstracts.

There was a good mix of attendees from academia, research laboratories, and industries. In particular, the DOE was well represented; with more than 20 participants from all of the major DOE laboratories (ANL, LBNL, LLNL, LANL, and SNL). Industry participation was also relatively strong. Overseas participants came from Belgium, China, Denmark, France, Germany, Great Britain, Israel, Italy, Japan, The Netherlands, Russia, and Tanzania.

Multi-Algorithm Methods for Multiscale Simulations Workshop

Dates January 14–16, 2004

Location Livermore, California

The Multi-algorithm Methods for Multiscale Simulations Workshop was held January 14–16, 2004, at the Hilton Garden Inn in Livermore, California. The event was hosted by CASC, ISCR and LLNL. It was sponsored by the Institute for Terascale Simulation (ITS) on behalf of the ASC Program. In all, 45 attendees, including 26 from U. S. Department of Energy laboratories and the balance from academia and industry, participated in the two-and-one-half day workshop.

Multiscale simulation is a central emerging numerical modeling paradigm for many science and engineering problem areas at LLNL and throughout the scientific community. Application areas include: materials design (nano-wires in computer chips, photonics, micro-electromechanical systems); biological systems (protein docking); medicine (drug delivery systems); and nuclear and aerospace technologies where materials failure and response in severe environments is a primary concern (dislocation patterns in fatigue and creep, surface roughening and crack nucleation in fatigue, etc.). Important scales in such problems range from macroscopic, where continuum models based on differential equations are usually employed, to atomistic, with quantum mechanical models applied at the finest resolution, plus all intermediate scales.

The Multi-algorithm Methods for Multiscale Simulations Workshop comprised five half-day sessions on various multiscale topics. Sessions were sorted by the organizers into methods and applications in solids, liquids, and gases and cross-cutting techniques. Each session concluded with a panel discussion in which the speakers and audience interacted richly and informally.

An important aim of the workshop organizers was to gather practitioners of multi-algorithm methods from various fields to find common ground in their work and engage in intellectual cross-fertilization. For example, materials scientists and gas dynamicists are not accustomed to working closely together, let alone with the mathematicians and computer scientists also present, since traditional conferences and meetings are usually focused on specific disciplines or problem areas. Many of the workshop participants explicitly stated that they found the workshop format and technical exchanges refreshing and informative. Several participants expressed strong interest in making the workshop an annual event.

Common challenges identified at the workshop included locating relevant scale boundaries and in grafting together representations of the solution from different methods in such boundary regions in which both are valid, in order to model a global system for which no single method is everywhere valid or efficient. Participants also discussed important challenges in understanding physical and mathematical error analysis in hybrid computational models and implementation challenges associated with complex models for large-scale parallel platforms.

An especially important outcome of the workshop is that DOE and academic researchers gained familiarity with many commonalities, as well as differences, among the problems on which they work. For example, it was revealed that understanding dissipative processes in solid mechanics (e.g., dislocations) can benefit from work done in modeling liquids. Also, the principle of entropy production used to understand the breakdown of continuum models in liquid and gas simulations appears to possess similarities with the principle of power dissipation used in solid mechanics. Other important areas identified as requiring further exploration included methods for bridging widely disparate time scales in hybrid simulations and methods for fluid-structure/gas-surface interactions, as well as incorporating more detailed chemistry in biological and nanoscale problems.

https://www.llnl.gov/casc/workshops/multiscale_simulations

Copper Mountain Conference on Iterative Methods

Dates March 28 – April 2, 2004

Location Copper Mountain, Colorado

The Copper Mountain Conference Series, held annually in early April at Copper Mountain, CO, alternates subjects between Multigrid Methods in odd-numbered years and Iterative Methods in even-numbered years. It represents an important forum for the exchange of ideas in these two closely related fields.

The Copper Mountain Conference on Iterative Methods was held March 28 – April 2, 2004. A total of 185 mathematicians from all over the world attended the meeting, which began with a reception on Sunday, March 28. During the following 5 days of the meeting, 131 talks on current research topics were given. Talks were organized into the following sessions.

1. Multigrid Solvers and Algebraic Multigrid
2. Saddle-Point Solvers
3. PDE Methods
4. Preconditioning Methods
5. Eigenvalue Methods
6. Multi-Physics Solution Methods
7. Krylov Subspace Methods
8. First Order System Least Squares Methods (FOSLS)
9. Nonlinear Solvers
10. Continuation Methods
11. Stochastic Systems
12. Parallel Algorithms
13. Applications
14. Software

In addition to the regular sessions, three evening workshops were offered. Monday night's workshop was a mini-symposium organized by Henry Tufo, who represented NCAR and the University of Colorado. Tuesday night, Michael Heroux from SNL organized a workshop on Sandia's Trilinos project. Wednesday night, a workshop organized by Eldad Haber from Emory University highlighted PDE-Constrained Optimization.

The sessions were all very well attended. The Copper Mountain Conference Series is known for having a very relaxed atmosphere and for fostering open, active discussions. This collaborative environment has characterized the meeting since the series began in 1983, and is one of the reasons many attendees come back repeatedly.

A student paper competition was held to stimulate student participation in the Conference. A panel of judges made up of members of the Program Committee selected the winners: Yair Koren (Technion, Israel); Ruth Holland (Oxford University, England); and Andrei Draganescu (University of Chicago).

http://amath.colorado.edu/faculty/copper/2004

Statistics and Practical Applications of Data Mining: Highlights from SDM04

Dates April 22–24, 2004

Location Orlando, Florida

The Fourth SIAM International Conference on Data Mining, held in Orlando, FL, April 22–24, 2004, continued the tradition of providing an open forum for the presentation and discussion of innovative algorithms, as well as novel applications of data mining. A record number of paper submissions this year marked not only a growing interest in the field, but also a greater acceptance of the conference among data mining researchers and practitioners.

Student authors accounted for a large percentage of the accepted papers, and their papers were reviewed under the same stringent guidelines as regular papers. The best student paper award was given to Martin Law from Michigan State University for his work on manifold learning. The award for the best algorithms paper went to a team from the University of Texas at Austin for their work on clustering, while the best applications paper was on enhancing communities of interest by a team from AT&T Laboratories.

A running theme of the conference was the practical application of data mining, including opportunities in various problem domains and practical lessons learned by those solving real data analysis problems in these domains. This was reflected in the topics covered in the three tutorials: analysis of patients' medical data, data mining for computer security, and mistakes commonly made in data mining and ways to avoid them.

In an industry–government session, speakers discussed problems encountered in the telecommunications industry, the role of information visualization, and data mining in such diverse domains as aviation safety and security, performance of computer networks, and earth sciences. Applications of data mining were also the subject of three of the keynote talks: Sara Graves of the University of Alabama at Huntsville considered issues of data usability; David Page of the University of Wisconsin Medical School elaborated on data mining questions raised by biology data; Ted Senator from the Defense Advanced Research Projects Agency (DARPA) discussed "connecting the dots." The increasing importance of homeland security was also reflected in many of the conference workshop topics, which ranged from link analysis, counterterrorism and privacy, to data mining in resource-constrained environments. More traditional topics, such as bio-informatics, mining of scientific and engineering data sets, and high-performance and distributed mining, also continued to attract participants.

Conference attendees clearly welcomed the focus on applications, which led to animated discussions in the industry-government presentations. One workshop speaker took the tutorial by John Elder on common mistakes in data mining to heart; she did some real-time editing of her presentation to point out the mistakes in her application domain, such as a lack of caution in sampling the data and discounting pesky cases though they might reveal a larger problem in the data.

A new aspect of this year's conference was the increasingly important role of statistics in data mining. Keynote speaker Chris Bishop of Microsoft Research: Cambridge discussed recent advances in Bayesian inference techniques and several technical sessions focused on statistical techniques in data mining. This connection between statistics and data mining will be exploited further in the next conference in the series (scheduled for Newport Beach, April 21–23, 2005), which will be co-sponsored by the American Statistical Association and SIAM (http://www.siam.org/meetings/sdm05/). We encourage statisticians and data miners to submit papers and attend the conference, and help us narrow the gap between the two fields to bring together the best of both worlds.

The proceedings of SDM04, including the keynotes and the presentations at the industry/government session, are available on-line at http://www.siam.org/meetings/sdm04.

Department of Homeland Security Advanced Scientific Computing Workshops

Dates	May 12, 2004	September 22–23, 2004
Locations	Washington, DC	Alexandria, Virginia

As part of the Department of Homeland Security (DHS) Advanced Scientific Computing program, ISCR co-hosted three requirements-gathering workshops during FY 2003. The first of these, the Advanced Scientific Computing Requirements Workshop was held October 8-9, 2003; a summary of this workshop can be found in the *ISCR 2003 Annual Report*. The two additional DHS workshops held this year were the Incident Management Simulation Workshop and the Data Sciences Workshop. The Krell Institute participated in the development of the content of these workshops and handled all workshop logistics and developed the workshop Web sites.

The Incident Management Simulation Workshop was held on May 12, 2004 at the Westin Grand Hotel in Washington, DC. This workshop brought together senior representatives of the emergency response and incident management communities with modeling and simulation technologists from DOE laboratories. The workshop provided an opportunity for incident responders to describe the nature and substance of the primary personnel roles in an incident response, identify current and anticipated roles of modeling and simulation in support of incident response, and begin a dialog between the incident response and simulation technology communities that will guide and inform planned modeling and simulation development for incident response.

The workshop was a joint effort of the Advanced Scientific Computing Program and the DHS Emergency Preparedness and Response Portfolio, both elements of the DHS Science and Technology Directorate. Based on the interactions at the workshop, a panel of computational science technologists prepared a summary report on incident management practice and the potential roles that computational simulation might play in supporting incident management. In particular, the panel prepared a summary of simulation capabilities that are relevant to incident management training and recommendations for the use of simulation in both incident management and in incident management training. In addition, the final report discusses areas where further research and development will be required to support future needs in this area.

The DHS Data Sciences Workshop was held September 22–23, 2004 at the Hilton Old Town in Alexandria, VA. The purpose of this workshop was to thoroughly review the data sciences mission needs of DHS and to identify specific mathematics and computer science research and development (R&D) topic areas required to address those needs. During the workshop, approximately 50 invited participants representing DHS, DOE and its national laboratories, academia and industry, identified specific R&D topic areas in the data sciences, their ties to the mission needs of DHS, and the potential impact of the proposed R&D. This effort will specify five years of relevant research topics in the data sciences area to support the Threat and Vulnerability Testing and Assessment Portfolio of DHS. These activities are important since they will immediately feed into the current DHS planning process for FY 2006.

Domain-Specific Languages for Numerical Optimization

Dates August 18–20, 2004

Location Argonne, Illinois

On August 18–20, 2004, ANL hosted a workshop on Domain-specific Languages for Numerical Optimization, co-sponsored by LLNL. There were 36 participants, including students, faculty, and staff from 12 universities, plus scientific staff from ANL, LLNL, and Sandia. The purpose of the meeting was to bring together experts in programming languages and compilers together with experts in numerical optimization and partial differential equations (PDEs) to stimulate discussion on the design and implementation of next-generation domain-specific languages for numerical optimization, with an emphasis on stochastic optimization and PDE-constrained optimization.

The participants discussed the design and implementation of current languages for numerical optimization, called modeling languages. These languages are mostly declarative, but take on an imperative flavor when a function or its derivatives must be evaluated. Native data types include scalars, sets, and ordered sets. Typically, the models and data are "compiled" into an internal representation or bytecode that is then interpreted. Several examples were given in which a problem that might have taken thousands of lines of code to express in Fortran or C required only 30-100 lines of code in AMPL or GAMS.

On the other hand, two examples were cited where converting a problem from a modeling language to C or Fortran (plus an FFT library in one case) resulted in a hundred- or thousand-fold speedup. Although the granularity of objects in modeling languages is typically much finer than that in other domain specific languages, it seems likely that static or dynamic compilation could provide the performance needed for some large problems without sacrificing the expressiveness of the modeling languages.

One of the obstacles to successfully extending modeling languages to support stochastic optimization is the wide variety of ways that randomness can enter an optimization problem and hence the many kinds of stochastic optimization problems. Even when the scope is restricted to a particular kind of stochastic optimization, multistage linear recourse problem, the specification of a problem is nontrivial. Part of the challenge arises from the fact that multistage problems can lead to enormous scenario trees. However, for problems in which stochasticity enters in a structured manner, extending a modeling language with time (stage) information and mechanisms to specify the probability distributions for random variables may suffice. The primary obstacle to extending modeling languages to support PDE-constrained optimization is that effective mechanisms for specifying partial differential equations themselves have not been developed. However, emerging systems, such as Sundance, FIAT, and PETSc 3, offer some hope that effective mechanisms for specifying and solving PDEs can be developed.

Several systems for analysis and transformation of general-purpose languages, domain-specific languages, and meta-languages were presented. The DMS Software Reengineering Toolkit supports automated source code analysis and modification. It utilizes Unicode lexers, GLR parsers for arbitrary context free grammars, analysis via multipass attribute grammars, and conditional source-to-source transformations. The extensible C (xtc) system uses a packrat parser to support arbitrary syntactic extensions to C, AST transformation rules to reduce and optimize, and typing rules to support safety constraints. Several researchers presented their work in the area of telescoping languages and related techniques. These methods exploit domain-specific analysis and optimization to improve the performance of general purpose languages extended with domain-specific libraries. Examples of systems supporting this paradigm are ROSE, Broadway, and libGen. In many cases, these systems are able to achieve performance superior to voluminous hand-developed Fortran or C implementations using concise implementations in C++ or MATLAB. It was also demonstrated that generic programming techniques can also provide high performance and high levels of expressiveness.

http://www-unix.mcs.anl.gov/workshops/DSLOpt/

Short-Pulse Laser Workshop

Dates August 25 – 27, 2004

Location Pleasanton, California

For three days at the end of August 2004, 55 plasma scientists met at the Four Points by Sheraton in Pleasanton to discuss some of the critical issues associated with the computational aspects of the interaction of short-pulse, high-intensity lasers with matter. The workshop was organized around the following areas of key interest to the Laboratory.

- Laser propagation / interaction through various density plasmas: micro scale.
- Anomalous electron transport effects: from micro to meso scale.
- Electron transport through plasmas: from meso to macro scale.
- Ion beam generation, transport, and focusing.
- "Atomic-scale" electron and proton stopping powers.
- Kα diagnostics.

Each area had a coordinator who drew up a list of questions, moderated discussions, and wrote a working group summary.

Many important problems in fast ignition are related to laser-plasma interactions, including laser propagation in the underdense corona plasma, laser hole-boring in the overdense plasma, laser absorption and energetic electron production at the critical surface, and electron transport in the mildly-dense plasma region. Participants worked out a set of benchmark computational simulation problems to compare their modeling capability in these areas.

Meso to macro scale electron transport discussions focused primarily on the correct method of initiating the electron beam. Several phenomenological techniques were discussed, such as injection at a plane in free space, promotion of ambient electrons, and the use of a ponderomotive force. It was generally agreed that the boundary conditions in the laser–plasma interaction (LPI) region were critical to the problem setup. A two-region approach, in which the LPI is simulated in the blow-off plasma and hybrid methods are used in the solid density material, might be a reasonable intermediate step. Since most codes do not have a laser-deposition package, it is necessary to choose and standardize the beam parameters, and a set was proposed.

The ion beam generation, transport, and focusing group discussed the following questions, from general to application specific.

- What are the proton generation mechanisms?
- What are their efficiencies?
- How sensitive to resolution are the answers?
- What codes can be used?
- How does electron flow affect proton generation?
- How can we control the generation and focusing of the protons?
- What is the optimum proton energy for radiography?
- What are the qualities that set ions using these mechanisms apart from "standard" ion beams?
- What governs ion flux?
- What is the optimal distance of the "proton lens" from the target?

The "atomic-scale" stopping powers session discussed first the stopping power of relativistic electron beams (REB) with energies of 1 to 10 MeV, stopping in pre-compressed deuterium–tritium (DT) targets, and the stopping power of non-relativistic (NR) protons with energies of 1 to 100 MeV. They set benchmark problems for multiple scattering of REB on target ions and of multiple scattering of NR protons in thin foils of high-Z materials disposed in front of laser proton sources (LPS). Finally, they considered REB and NR proton stopping in strongly magnetized fast ignition targets.

The purpose of the Kα diagnostic sessions was to discuss some of the progress being made in modeling Kα emission in short-pulse petawatt laser experiments and to discuss with experimentalists some of their latest results.

The workshop was made possible by the joint financial support of the Institute for Laser Science and Applications and the ISCR at LLNL.

Computational Methods in Transport

Dates September 11–16, 2004

Location Tahoe City, California

The Computational Methods in Transport Workshop was devoted to providing a forum where computational transport researchers could discuss their methods, successes and failures across disciplinary boundaries. Typically, the numerical methods used in a given field are communicated to other researchers in that field. Rarely, however, are those methods communicated between one application domain of radiation transport and another.

For example, nuclear engineers and astrophysicists rarely attend the same meetings or read the same literature. The goal of the Computational Methods in Transport Workshop was to address this discrepancy and open channels of communication and cooperation so that (1) existing methods used in one field could be applied to other fields and (2) greater scientific resources could be mobilized to help solve outstanding problems.

Beginning on the afternoon of September 11, 2004 and ending with lunch on September 16, 2004, the workshop was held at the Granlibakken Conference Center in Tahoe City, CA. The first day of the meeting consisted of a series of one-hour talks reviewing one of the major fields represented at the workshop. The areas covered included astrophysics, atmospheric physics, mathematics, plant canopies, nuclear engineering, oceanography and high-energy density physics. One afternoon was reserved for a poster session where 30 posters were presented in a very lively and well-attended event. The following days were filled with focused 45-minute talks by each of the representative fields that delved into more technical detail. Substantial time was reserved for individual networking and communications.

Speakers were chosen based on their international recognition and covered various topics:

- **Ed Larsen** (University of Michigan) — Numerical methods used in neutron transport.
- **Tony Mezzacappa** (ORNL) — Applications and numerical methods used in supernova core collapse.
- **David Levermore** (University of Maryland) — Moment and closure approximations used in approximating transport equations.
- **Marty Marinak** (LLNL) — Transport needs in high-energy density physics.
- **George Kattawar** (Texas A&M) – Polarization and radiative transfer in oceanography.

The conference structure and venue work extremely well. Participants from different fields, who would never have had the opportunity to speak with other participants, were engaged in stimulating and very fruitful discussions. The atmosphere was collegial with all participants willing to learn and teach. As the conference week progressed, atmospheric scientists were learning about methods used in nuclear engineering.

A radiation physicist related, "Using the Fokker–Planck equation for studying our scattering problems never occurred to me. We are going to look into this." A mathematician who gave a talk on medical imaging and radiation oncology forged a bond with a nuclear engineer and is taking a sabbatical leave to apply his knowledge to radiation oncology. Another outgrowth of the workshop is that Ryan Clement (LLNL) and organizer Frank Graziani are setting up an e-print server that will serve as a repository for computational transport papers.

Overall, feedback from the workshop has been very positive and most expressed hope that it would be done again. The plan is to do a smaller, multi-disciplinary special-topics meeting next year followed by a larger meeting in 2006 similar to the 2004 workshop. Like the 2004 workshop, the future ones will be organized in conjunction with the Institute for Pure and Applied Mathematics at UCLA.

http://www.ipam.ucla.edu/programs/tr2004/

AMG / FOSLS Summit

Dates September 27 – October 3, 2004

Location Lake City, Colorado

In 1997, an informal meeting was held in Frisco, CO, between researchers in CU Boulder's Applied Mathematics Department and LLNL's CASC Division to discuss their collaboration on Algebraic Multigrid (AMG) methods. They met again in Boulder in 1998. In 2000, the meetings became annual, held in Lake City, CO. In 2002, they were expanded to include discussions on the First Order Systems Least Squares (FOSLS) methodology.

The summit is structured as a "working meeting" with an emphasis on exposing open research issues and generating ideas for solving them. Formal talks are strongly discouraged in favor of whiteboard discussions and individual interactions, a format that distinguishes it from typical meetings and conferences. Participation is by invitation only, consisting primarily of researchers from CU Boulder and CASC, but also including a small number of leading experts from other institutions around the world.

The Summit was held September 27–October 3, 2004. The first half of the meeting focused on AMG, and the second half focused on FOSLS, with an overlap day in between. There were 31 attendees this year. The main CU/CASC group consisted of 11 from CU Boulder, 9 from CASC, plus 4 recent CU graduates. The other attendees were: Irad Yavneh (Technion, Israel), Ludmil Zikatanov (Penn State), Ira Livshits (Ball State), Achi Brandt Weizmann Institute, Israel), Bobby Philip (LANL), Marzio Sala and Michael Gee (SNL).

The first half hour of the meeting was spent setting the agenda. The topics suggested this year by the attendees were as follows (as written on the whiteboard).

- Weighted Functionals
- Smooth Aggregation & e-Free AMGe
- Nonlinearity
- Almost Zero Modes
- Measures
- Coarse Variable Types
- Sparsity of P/Dilution
- Relaxation
- Sharp Theory
- Trace Minimization
- Iradism
- Upscaling
- Wavelet AMG

During the remainder of the meeting, each topic item (and its associated issues) was discussed in detail and solution approaches were proposed and debated. Two sample outcomes of the meeting were an improved theoretical foundation for a new trace minimization approach for defining interpolation in AMG, and an idea for relating a new compatible relaxation method (one that defines the coarse variables as averages) to a recent AMG theory and framework.

Summer Research Program

The ISCR put on its largest summer visitor program ever in FY 2004 with approximately 100 students in residence, as well as 11 faculty. Students are assigned individual LLNL mentors and given specific projects, ranging from programming tasks to original research, to which they will contribute based on their background and skills. Reports for most of the projects in the following table have been omitted from the printed version of the ISCR Annual Report in the interest of space, but can be found on the accompanying PDF. You can also obtain a CD-ROM containing these reports by calling (925) 423-3691.

Summer Research Program

Name, Affiliation — Mentor, Host Organization

Project Title

Sabbatical Visitors

Carl Ollivier-Gooch, University of British Columbia — Lori Diachin, CASC

Unstructured Mesh Technology Development

Anne Greenbaum, University of Washington — Peter Brown, CASC

Boundary Integral Equation Methods and Software for Poisson's Equation in Three Dimensions

Bjorn Sjogreen, Royal Institute of Technology, Stockholm, Sweden — Anders Petersson, CASC

Embedded Boundary Methods for Partial Differential Equations

Summer Faculty

Constantin Bacuta, University of Delaware — Rob Falgout, CASC

A New Approach for Solving the Stokes Problem Based on the Distributed Relaxation Method

James Brannick, University of Colorado at Boulder — Rob Falgout, CASC
Ludmil Zikatanov, Pennsylvania State University — Panayot Vassilevski, CASC

Compatible Relaxation and Optimal AMG Interpolation

Zhiqiang Cai, Purdue University — Panayot Vassilevski, CASC

Mixed and Multigrid Methods for Stokes and Navier–Stokes Equations

Tim Chartier, Davidson College — Van Emden Henson, CASC

Adaptive Multigrid via Subcycling on Complementary Grids

Irene Livshits, University of Central Arkansas — Rob Falgout, CASC

AMG Algorithm for Finding an Eigenbasis for the Schroedinger Operator

Anne Ngu, Southwest Texas University — Terence Critchlow, CASC

Large-Scale Integration of Web Sources

Summer Students

David Alber, University of Illinois — Jim Jones, CASC

Adapting Algebraic Multigrid for the Solution of the Curl–Curl Formulation of Maxwell's Equations

Kristopher Andersen, Unversity of California, Davis — John Pask, PAT

Locally Optimal Methods to Solve Eigenvalue Problems in Electronic Structure Calculations

John Anderson, University of California, Davis — Benjy Grover, DCOM

QTester

Dustin Anderson, California Polytechnic State University, San Luis Obispo — Don MacQueen, EPD

Real-Time Radiation Area Monitoring: Emergency Response and Regulation

Benjamin Apodaca, Northern Arizona University — Carolyn Wimple, NIFE

PIMS Regression Test Development and Product Enhancement

Name, Affiliation	Mentor, Host Organization
Abraham Bagherjeiran, University of Houston Design and Implementation of an Anomaly Detector	Erick Cántu–Paz, CASC
Paul Baginski, University of California, Berkeley Scalable Graph Algorithms	Van Emden Henson, CASC
Jeffrey W. Banks, Rensselaer Polytechnic Institute High-Speed Multicomponent Flows	William Henshaw, CASC
Lerone Banks, University of California, Davis Firewall Egress Filtering	Jerry Rayome, CSP
Janine Bennett, University of California, Davis Volumetric Mesh Parameterization towards Slow-Growing Subdivision	Valerio Pascucci, CASC
James Brannick, University of Colorado at Boulder Adaptive Algebraic Multigrid Preconditioners in Quantum Chromodynamics	Rob Falgout, CASC
Steve Callahan, University of Utah A Memory-Insensitive Format for Out-of-Core Access to Unstructured Volumetric Meshes	Valerio Pascucci, CASC
Jedidiah Chow, Granada High School Discovery Center Display	Jean Shuler, CSG
Kevin Chu, Massachusetts Institute of Technology Incorporating Electrokinetic Effects into the EB Navier–Stokes Embedded Boundary Incompressible Fluid Solver	David Trebotich, CASC
Kree Cole-McLaughlin, University of California, Los Angeles Applying Morse Theory to Computational Datasets	Valerio Pascucci, CASC **Terence Critchlow, CASC**
Dylan Copeland, Texas A & M University Geometric Multigrid for Variable Coefficient Maxwell's Equations	Barry Lee, CASC
Steven Davis, California State University, Hayward Web Designer, Program for Climate Model Diagnosis and Intercomparison	Charles Doutriaux, EEBI
Stanko Dimitrov, University of Arizona Edge Betweenness Properties in Complex Networks	Edmond Chow, CASC
Veselin Dobrev, Texas A & M University Element Agglomeration AMGe Solvers for Unstructured Finite-Element Problems	Panayot Vassilevski, CASC
Nina Dokeva, University of Southern California Parallel Implementation of a FETI–DP Method for Elliptic Problems with Mortar Finite Element Discretization	Panayot Vassilevski, CASC
Emily Eder, University of California, Los Angeles Graph Viewer Improvements	Robert Fernandes, IOAC

Name, Affiliation / Project Title	Mentor, Host Organization
Christopher Egner, Rochester Institute of Technology Platform Independence for Cale: wxWidgets and GNU Autotools	Paul Amala, AX Division
Chris Elofson, University of Arizona Wavelet-Based Opacity Data Compression	Vijay Sonnad, DCOM
Dayo Esho, University of California, Berkeley ODE Visualization and Archive Tool in Python	Steven Lee, CASC
Pak-Wing Fok, Massachusetts Institute of Technology 2D Numerical Modelling of Soap Film in Overture	Petri Fast, CASC
Camille Fournier, University of Wisconsin-Madison Implementing Cray Pointers in the GNU gFortran Compiler	Michael Kumbera, DCOM
Jeffrey L. Freschl, University of Wisconsin-Madison Performance Analysis of Monitoring and Information Systems Using NetLogger and an Investigation into P2P Overlay Topologies	John Johnson, DCOM
Evan Geller, Summit Preparatory High School Bringing Order to CHAOS	Albert Chu, HPSD
Tobias Gradl, Technische Universität München, Germany Assessing Performance of Hybrid MPI/OpenMP Programs on SMP Clusters	Edmond Chow, CASC
Rachel Greenstadt, Harvard University Censorship Resistance	David Youd, NAIC
Attila Gyulassy, University of California, Davis Hierarchical Morse–Smale Complexes	Valerio Pascucci, CASC
Daniel Han, University of Southern California MPI Profiling	Terry Jones, ICC Services and Development Division
Kevin Hoffman, Brigham Young University Parallel Analysis of Asymmetries in Symmetric Simulations in ALE3D	Robert Cooper, DCOM
W. Taylor Holliday, University of California, Davis Combinatorial Feature Extraction for a Streaming Framework	Valerio Pascucci, CASC
Gary Hon, University of California, San Diego Virtual PCR (vPCR)	David Clague, EE-EETD
Jason Howell, Clemson University A Two-Grid Method for Radiation Diffusion	Carol Woodward, CASC
David Hoyt, Brigham Young University DHS Countermeasures Test Bed–Command Center	James Schek, NAIC

Name, Affiliation	Mentor, Host Organization
Project Title	
Kevin Huck, University of Oregon	Brian Miller, CASC
PerfTrack	
Lukas Jager, Universität Bonn, Germany	Radu Serban, CASC
A Multilevel Preconditioner for a PDE-Constraint Optimization Problem	
Abram Jujunashvili, University of Colorado at Denver	Charles Tong, CASC
Integration of Multivectors to Hypre	
Alin Jula, Texas A&M University	Dan Quinlan, CASC
Automatic Detection of Hot Spots in C/C++ Programs	
Kirk Kelsey, University of Rochester	Tom Epperly, CASC
Automatic SIDL Generation with ROSE	
Jason Kimball, University of California, Irvine	Mark Duchaineau, CASC
Large-Scale Atom Data Visualization	
Nicholas Kridler, University of Colorado at Boulder	Barry Lee, CASC
Algebraic Multigrid for Maxwell's Equations with Variable Coefficients	
Brian Lum, University of California, San Diego	Amber Marsh, EEBI
IMAGE Pipeline	
Eric Machorro, University of Washington	Britton Chang, CASC
Numerical Solution Methods and Error Analysis for the Neutron Transport Equation in Absorbing, Monoenergetic, Non-Scattering Media	
Anna Majkowska, University of California, Riverside	Valerio Pascucci, CASC
Out-of-Core Visualization of Climate Modeling Data	
Robert Dean Malmgren, Northwestern University	Scott Brandon, AX Division
Code-Independent Analysis Tools for Physics Simulation Codes	
Peter Manning, University of Arizona	Glenn Goderre, NIFE
Development of a Graphical User Interface for the Virtual Beam Line (VBL) Simulation Software	
Ajith Mascarenhas, University of North Carolina at Chapel Hill	Valerio Pascucci, CASC
Time-Varying Reeb Graphs for Space-Time Data	
Tammara Massey, University of California, Los Angeles	Terry Brugger, NAIC
Ultra-Wideband Token Ring Simulation and Security in Sensor Networks at LLNL	
Kathryn Mohror, Portland State University	John May, CASC
PerfTrack	
Christopher Muelder, University of California, Davis	Marv Christensen, NAIC
Summarizing Network Traffic with Information Visualization	

Summer Research Program

Name, Affiliation / Project Title	Mentor, Host Organization
Spencer Nielsen, Brigham Young University Software Failure Risk Analysis	Darrel Whitney, CADSE
Daniel Oeltz, Universität Bonn, Germany A Space-Time Sparse Grid Approximation Space	Panayot Vassilevski, CASC
Christopher Olson, University of California, Santa Cruz Reviewing Vista	Robert Cooper, DCOM
Timothy Paik, University of California, Berkeley Subspace Detectors	Shawn Larsen, ISCR Dave Harris, ISCR
Sung W. Park, University of California, Davis Streaming Pointsets	Peter Lindstrom, CASC
Bryan Parno, Harvard University Internet Ballistics: Retrieving Forensic Data From Network Scans	Tony Bartoletti, NAIC
Dan Phung, Columbia University SLURM for BlueGene/L	Moe Jette, CASC
Sriram Polepeddi, Carnegie Mellon University Software Vulnerability Taxonomy Consolidation	Noel Tijerino, IOAC
Raymond Pon, University of California, Los Angeles Performance-Oriented, Privacy-Preserving Data Integration	Terence Critchlow, CASC
Serban Porumbescu, University of California, Davis Out-of-Core Parameterization and Remeshing of Surfaces	Mark Duchaineau, CASC
Ashley President, Carnegie Mellon University Intrinsic Function Testing and an OpenMP Runtime Library Summer	Michael Kumbera, A Division
Davinder Rama, California State University, Sacramento Amino-Acid Sequence into Tertiary Structure Web Site	Adam Zemla, EEBI
George Roberts, Georgia Institute of Technology Variance-Based Feature Tracking	Chandrika Kamath, CASC
Rolf Ryham, Penn State University Multigrid Prolongation Based on Sharp Convergence Theory	Rob Falgout, CASC
Andreas Saebjornsen, Universitete i Oslo, Norway The AST Query Mechanism and the C/C++ Graphing Mechanism	Dan Quinlan, CASC
Elmer Salazar, California State University, Stanislaus Implementation of OpenMP Support in the gFortran Compiler	Michael Kumbera, DCOM
Jennifer Sirp, California State University, Sacramento Extending ReiserFS for Automatic File Queuing	Terry Brugger, NAIC

Name, Affiliation	Mentor, Host Organization
Project Title	
Yoshifumi Suzuki, University of Michigan	Jeff Hittinger, CASC
Efficient Schemes for Hyperbolic Systems with Stiff Relaxation Source Terms	
Valerie Szudziejka, University of California, Davis	Valerio Pascucci, CASC
Streaming Computation of Structural Graphs	
Ryan Szypowski, University of California, San Diego	Ulrike Yang, CASC
Parallel AMG for Systems of PDEs	
Brian Taylor, University of Illinois at Urbana-Champaign	Bill Henshaw, CASC
Computation of Cellular Detonation	
Robert Taylor, Northern Arizona University	Jody Malik, CSP
Vulnerability Tracking Database 2.0	
Nils Thuerey, Universität Erlangen-Nürnberg, Germany	Dan Quinlan, CASC
Global Analysis of the ROSE Infrastructure	Jody Malik, CSP
Peter Tipton, University of Southern California	Brandon Scott, AX Division
Code Validation Made Easy	
Chunbo Wang, Purdue University	Charles Tong, CASC
A MATLAB Implementation of Mixed Finite-Element Method for Incompressible Newtonian Flows: Pseudostress–Velocity Formulation	
Rebecca Wasyk, Worcester Polytechnic Institute	Carol Woodward, CASC
Newton–Krylov Methods for Expensive Nonlinear Function Evaluations	
Sage Weil, University of California, Santa Cruz	Tyce McLarty, CADSE
Metadata Management for Petabyte-Scale File Systems	
Dan Wendlandt, Stanford University	Martin Casado, NAIC
Information Leakage Due to Geographic Properties of Internet Routing	
Brian White, Cornell University	Dan Quinlan, CASC
Ameliorating the Performance Degradation of User-defined Abstractions and Indirect Memory Accesses	
Ryan M. White, University of California, Berkeley	Shawn Newsam, CASC
Matching Shapes Using Local Descriptors	
Jeremiah Willcock, Indiana University	Dan Quinlan, CASC
Additions to the ROSE Compiler Infrastructure	
Suzanne Wingenter, San Diego State University	Petri Fast, CASC
Shallow-Water Equations on Curvilinear Grids	
Christopher Wojtan, University of Illinois, Urbana-Champaign (2000–2004) Georgia Institute of Technology (2004)	Jeremy Meredith, DNT
A Hybrid Sort-First/Sort-Last Approach for Rendering Translucent Geometry in the VisIt Visualization Tool	

Summer Research Program

Name, Affiliation — Mentor, Host Organization

Project Title

Sung-Eui Yoon, University of North Carolina at Chapel Hill — Peter Lindstrom, CASC

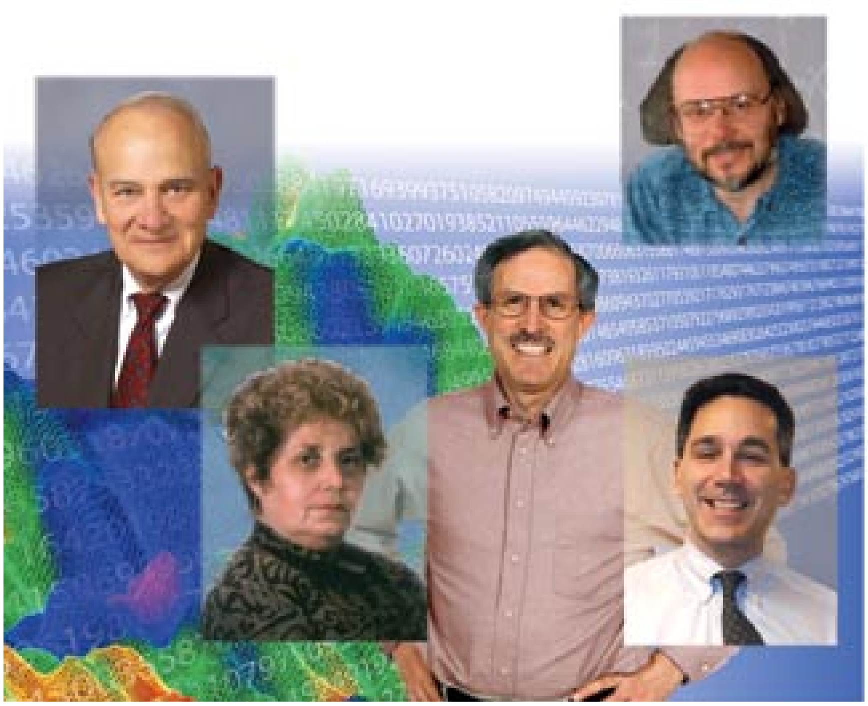
Cache Coherent Mesh Layout

Beth Yost, Virginia Polytechnic Institute and State University — Terence Critchlow, CASC

A Visual Interface for the Promoter Identification Workflow

ISCR Seminar Series

The ISCR hosted 70 seminars from visitors in FY 2004 covering a wide spectrum of research areas, and recruited an additional 35 speakers from LLNL ranks to speak to visiting students. The ASC Institute for Terascale Simulation Lecture Series was established in 2000 to enrich the intellectual atmosphere of LLNL's large simulation community through the visits of leaders representing the diverse areas of computation. In FY 2004, we hosted William Wulf, John Grosh, Bjarne Stroustrup, Mary Wheeler, and David Bailey. The ISCR Summer Student Lecture Series was also established in 2000 and forked into three different series in Summer 2004—Internships in Computational Modeling at the Terascale (ICMT), Internships in Computer Science at the Terascale (ICST), and College Cyber Defenders Computer Security. Seminar reports listed in the following table have been omitted from the printed version of the ISCR Annual Report in the interest of space, but can be found on the accompanying PDF. You can also obtain a CD-ROM containing these reports by calling (925) 423-3691.



ISCR Seminar Series

Date	Speaker, Affiliation / Title of Seminar

ITS Lecture Series

Date	Speaker, Affiliation / Title of Seminar
12/4/03	**William Wulf**, National Academy of Engineering Challenges for Computing and Information Technology in the 21st Century
2/17/04	**John Grosh**, Department of Defense A DoD Perspective on High-End Computing
3/1/04	**Bjarne Stroustrup**, Texas A&M University Programming, Language and Libraries
6/24/04	**Mary Wheeler**, University of Texas at Austin Mathematical and Computational Modeling of Multiphysics Couplings
8/4/04	**David Bailey**, Lawrence Berkeley National Laboratory Twelve Ways to Fool the Masses: Back to the Future

Visitor Seminars

Date	Speaker, Affiliation / Title of Seminar
10/1/03	**Kenneth Chiu**, Indiana University Bloomington A C++ Reflection Library
10/3/03	**Branden Fitelson**, University of California, Berkeley A User-Friendly Probability Machine with Applications
10/9/03	**Olav Beckmann**, Imperial College London,United Kingdom Dynamic Code Generation in C++ as a Foundation for Domain-Specific Optimization
10/23/03	**Charles Elkan**, University of California, San Diego Progress in Clustering
10/23/03	**James Glimm**, Stony Brook University Statistical Models for Simulation Errors and Their Role in Prediction and Uncertainty Quantification
10/24/03	**Xiaolin Li**, Stony Brook University Simplification, Adaptivity and Conservation in Front Tracking Method
10/27/03	**Achi Brandt**, Weizmann Institute of Science, Israel Image Processing: From Segmentation to Recognition
10/30/03	**Douglas Swesty**, Stony Brook University Efficient Solution of the Discretized S_N Time-Dependent Boltzmann Transport Equation on Parallel Platforms
10/31/03	**Frank Mueller**, North Carolina State University Detailed Cache Coherence Characterization for OpenMP Benchmarks
11/7/03	**Daniel Segre**, Harvard Medical School Computational Models of Cellular Metabolic Fluxes

Date	Speaker, Affiliation Title of Seminar
11/13/03	**Richard Scalettar,** University of California, Davis Linear Algebra and Quantum Simulations
11/14/03	**Marsha Berger,** New York University Cartesian Grids with Embedded Geometry
11/24/03	**Johannes Gehrke,** Cornell University Techniques for Processing Large Data Streams
12/3/03	**John Feo,** Cray Inc. Cray Cascade Project
12/3/03	**Ravi Samtaney,** Princeton Plasma Physics Laboratory The Magneto–Hydrodynamic Richtmyer–Meshkov Instability
12/9/03	**Margot Gerritsen,** Stanford University Why Are Streamline Methods Attractive for Simulation of Gas-Injection Processes?
12/9/03	**Jonas Nilsson,** Stanford University A Hybrid High-Order Method for the Incompressible Navier–Stokes Equations
12/12/03	**Barney Maccabe,** University of New Mexico Early Experience in Splintering Communication Protocols
12/12/03	**Jennifer Widom,** Stanford University The Stanford Data Stream Management System
12/17/03	**Daniel Crichton,** Jet Propulsion Laboratory A Data Grid Framework for Managing Planetary Science Data
12/17/03	**William Dally and Patrick Hanrahan,** Stanford University Merrimac: Supercomputing with Streams
12/18/03	**Terran Lane,** University of New Mexico From Security to Cells: Ongoing Machine Learning Research at the University of New Mexico
1/9/04	**Christopher Jermaine,** University of Florida Approximate Query Processing with Sampling and Pre-Aggregation
1/12/04	**Thomas Seidman,** University of Maryland, Baltimore County Hybrid Systems: Discontinuous Dynamics in a Continuous World
1/16/04	**Serge Belongie,** University of California, San Diego Three Brown Mice: See How They Run — Monitoring Rodent Behavior in the Smart Vivarium
1/23/04	**Cheryl McCosh,** Rice University Type-Based Specialization in a Telescoping Compiler for MATLAB
1/26/04	**Elizabeth Post,** Lincoln University On the Farm: Parallel Small Talk for Simulating Dairy Operations

ISCR Seminar Series

Date	Speaker, Affiliation Title of Seminar
1/27/04	**Steven Parker,** University of Utah What's New with SCIRun2?
1/29/04	**Miguel Argaez,** University of Texas at El Paso An Optimization Technique for Large-Scale Nonlinear Programming
1/29/04	**Leticia Velazquez,** University of Texas at El Paso A Global Optimization Technique for Solving Zero or Very Small Residual Nonlinear Least-Squares Problems
2/10/04	**James Hobart,** Classic Systems Solutions Designing for Usability
2/10/04	**Alan Laub,** University of California, Davis Statistical Condition Estimation
2/11/04	**Demet Aksoy,** University of California, Davis PLASMA (PLAnetary Scale Monitoring Architecture)
2/18/04	**Donald Schwendeman,** Rensselaer Polytechnic Institute Numerical Method for High-Speed Reactive Flow on Overlapping Grids
2/20/04	**Matteo Pellegrini,** University of California, Los Angeles PROLINKS: A Database of Co-Evolving Proteins
2/25/04	**Laxmikant Kale,** University of Illinois, Urbana-Champaign Adaptive Resource Management via Processor Virtualization: Charm++ and AMP
3/3/04	**Lada Adamic,** Hewlett-Packard Laboratories How to Search a Social Network
3/4/04	**Christoph Pflaum,** Universität Erlangen-Nürnberg, Germany 3D Computation of Laser Cavity Eigenmodes by Finite Elements
3/5/04	**Joel Saltz,** Ohio State University Middleware Support for Data Ensemble Analysis
3/10/04	**Stephen Neuendorffer,** University of California, Berkeley Actor-Oriented Metaprogramming
3/15/04	**Luiz DeRose,** IBM Research DPOMP: An Infrastructure for Performance Monitoring of OpenMP Applications
3/18/04	**Boleslaw Szymanski,** Rensselaer Polytechnic Institute Analyzing Evolution of Virulence through Spatially-Explicit Epidemic Models
3/26/04	**Gunther Weber,** University of California, Davis Topology-Based Exploration of Scalar Fields

Date	Speaker, Affiliation Title of Seminar
4/6/04	**Kirk Hays and Max Alt,** Intel Corporation Performance Analysis and Tuning of MDCASK and PF3D Codes on Itanium Processors
5/6/04	**David Liu,** University of California, Berkeley GridDB: A Data-Centric Overlay for Scientific Grids
5/14/04	**Scott Baden & Jacob Sorenson,** University of California, San Diego Data-Driven Execution of Communication Tolerant Algorithms
5/17/04	**Homer Walker,** Worcester Polytechnic Institute Globalization Techniques for Newton–Krylov Methods
5/24/04	**Jeffrey Heys,** University of Colorado at Boulder Numerical Issues When Modeling Fluid-Elastic Interaction in 3D with First-Order System Least Squares
5/24/04	**Chad Westphal,** University of Colorado at Boulder First-Order System Least-Squares for Problems with Boundary Singularities
6/1/04	**Martin Bazant,** Massachusetts Institute of Technology Induced-Charge Electro-osmosis
6/3/04	**Marc Barthelemy,** Commissariat á l'Energie Atomique, France Structure and Modeling of Weighted Complex Networks
6/8/04	**David Jensen,** University of Massachusetts Amherst Knowledge Discovery in Networks
6/10/04	**Steven Knight,** SCons Project SCons: A Next-Generation Build Tool
6/16/04	**Wenke Lee,** Georgia Institute of Technology Worm Detection and Response: Local Strategies and Analytical Models
7/21/04	**Frank Mueller,** North Carolina State University Detailed Cache Coherence Characterization for OpenMP Benchmarks
7/23/04	**Sameer Agarwal,** University of California, San Diego On Refractive Optical Flow
7/23/04	**Gabriele Jost,** NASA Ames Research Center What Multilevel Parallel Programs Do When You Are Not Watching
7/29/04	**David Forsyth,** University of California, Berkeley Words and Pictures
7/29/04	**Ling Liu,** Georgia Institute of Technology ReFlex: Flexible and Reliable Systems Technologies for Responding to Massively Disruptive Events

ISCR Seminar Series

Date	Speaker, Affiliation / Title of Seminar
7/29/04	**Nikolaos Nikiforakis,** Cambridge University, United Kingdom Emergence of Detonation in the Flowfield Induced by Richtmyer-Meshkov Instability
8/13/04	**Chris Wiggins,** Columbia University Data-Driven Approaches for Biological Networks: Inference, Organization and Analysis
8/27/04	**Alex Schweitzer,** Universität Bonn, Germany Efficient Implementation and Parallelization of Meshfree and Particle Methods: The Parallel Multilevel Partition of Unity Method
9/8/04	**E. Ann Stanley,** Los Alamos National Laboratory Using Mathematical Models to Understand the AIDS Epidemic and Guide Policy
9/9/04	**Jia Li,** University of Alabama, Huntsville Mathematical Modeling of Malaria, Early Warning System, and Transgenic Mosquitoes
9/30/04	**Reagan Moore,** San Diego Supercomputer Center Digital Libraries and Data-Intensive Computing

LLNL Summer Seminars, ICST Computer Science

Date	Speaker, Affiliation / Title of Seminar
6/15/04	Pat Miller, CASC FlashMob Instant Supercomputing
6/22/04	David Jefferson, CASC The Time Warp Method of Parallel Discrete Event Simulation.
6/29/04	Kim Yates, CASC BlueGene/L: The Next Generation of Scalable Supercomputer
7/13/04	Chandrika Kamath, CASC Scientific Data Mining: The Sapphire Project
7/20/04	Tom Epperly, CASC Babel Language Interoperability Tool
7/27/04	Gregory M. Pope, CADSE Why Software Quality Assurance Practices Become Evil!
8/3/04	Kim Minuzzo, NIF NIF Control System
8/10/04	Valerio Pascucci, CASC Multiresolution Computation and Presentation of Topological Structures
8/17/04	Erick Cantú-Paz, CASC Solving Problems with Evolutionary Algorithms

Date	Speaker, Affiliation Title of Seminar

LLNL Summer Seminars, ICMT Computational Science

6/9/04	David Keyes, Columbia University and ISCR A Science-Based Case for Large-Scale Simulation
6/16/04	Gary Kumfert, CASC Power Presentations
6/30/04	Peter G. Eltgroth, CASC Overview of Research Activities in the Center for Applied Scientific Computing
7/6/04	Paul Dubois, CASC Effective Architectures for Scientific Programs
7/7/04	Steven F. Ashby, CAR Computational Science at Lawrence Livermore National Laboratory
7/14/04	David Trebotich, CASC Computational Microfluidics
7/21/04	Frank Graziani, B Division Radiation Transport in 3D Random Media: Direct Numerical Simulation
7/28/04	Don Dossa, CASC Architectural Overview of BlueGene/L Supercomputer
8/11/04	Rose McCallen and Kambi Salari, CASC DOE's Effort to Reduce Truck Aerodynamic Drag — Joint Experiments and Computations Lead to Smart Design
8/18/04	Lori Freitag-Diachin, CASC Developing Interoperable Meshing and Discretization Components
8/25/05	Rob Falgout, CASC Scalable Linear Solvers: Multigrid Methods

LLNL Summer Seminars, CCD Computer Security

6/10/04	Terry Brugger, NAIC Data Mining for Network Intrusion Detection
7/1/04	Tony Bartoletti, NAIC Identifying Internet Adversaries Despite Falsified Source Addressing
7/8/04	Bill Orvis, EE-EETD The Tale of Two Rootkits — SucKit and Hacker Defenders

Acronyms and Abbreviations

1D 2D 3D AAWP AC AIDS ALC ASC ALE3D AMG
aAMG AMGe AMPI AMPL AMR ANL AP API
ARPACK ASC AST AS2TS BELCh BGK BGL BLAS
BLAST BMP BOOST BSP CADSE CAR CASC
dbCAT CCA CCD CES CFD CFL CHAOS CHOMBO
CIAC CJ COCs CORBA CR CRON C-SAFE CSGF CSS
CSV CU CVDB DAKOTA DAG DARPA DC DCOM
DDT DG DHS DMS DMTs DoD DNA DNS DNS
DOE DPCL DPIV DPOMP DSCEAs EC EE EETD EEBI
EFGM EOC EPIC ERD ERI ESS EGG FAS FASTA FBA
FEM FETI-DP FFLO FFT FIAT FORTRAN FOSPACK
FOSLS FY GAMS GCC GEM GHHE GLR GMRES
GNU GridDB GUI HCI HIV HkDef HLL HP HPC
HTML HTTP HVAC HYDRA HYPRE IBM ICAT ICCD
ICCS ICEO ICMT ICST IDS ISO IDX ILP IMAGE
IMPACT IOR IP IPAM I/O ISCR ISMG IT ITER ITS
JDO JPEG JSP KINSOL KOJAK KULL LANL LDA
LDRD LES LGA LINPACK LLNL LOBPCG LOCKSS
LOD LPI LPS MATLAB MCR MDCASK MEMS METIS
MHD MIRANDA MLIC MOMA MPI MRF fMRI NAI
NAIC NAMD NASA NCAR NFS NIF NIFE NNSA NR
ODE OODT OpenMP ORNL OS OSVDB PAPI PAT
PaToH PC PCG PCMDI PCR vPCR PDB PDEs PDS
PERC PerfTrack PF3D PHP PHP/MySQL PI PIMS PIW
PLASMA PMaC PMPI POMP POP POSIX PROLINKS
PUM PVODE QBF QC QCD RANS RANSAC RBF REB

Acronyms and Abbreviations

A

1D – one-dimensional

2D – two-dimensional

3D – three-dimensional

AAWP – Analytical Active Worm Propagation

AC – alternating current

AIDS – acquired immune deficiency syndrome

ALC – ASC (see) Linux Cluster, provides computing cycles for ASC Alliance users and unclassified ASC code development. ALC and MCR (see) are sibling Linux clusters with ~1000 Intel P4 nodes with dual 2.4GHz processors.

ALE3D – Arbitrary Lagrangian-Eulerian 3D code

AMG – Algebraic Multigrid method developed to solve matrix equations resulting from the discretization of an elliptical PDE on an unstructured mesh

aAMG – adaptive AMG

AMGe – Algebraic Multigrid based on element interpolation

AMPI – Adaptive Message Passing Interface

AMPL – A Mathematical Programming Language

AMR – adaptive mesh refinement

ANL – Argonne National Laboratory

AP – asymptotic preservation

API – Application Program Interface

ARPACK – a collection of Fortran77 subroutines designed to solve large-scale eigenvalue problems

ASC – Advanced Simulation and Computing

AST – abstract syntax tree

AS2TS – Amino-acid sequence into tertiary structure

B

BELCh – Boundary Exponential Log Characteristic

BGK – Bhatnagar-Gross-Krook

BGL – BlueGene/L, an IBM computer, currently the world's fastest

BLAS – Basic Linear Algebra Subprograms

BLAST – Basic Local Alignment Search Tool

BMP – file extension for bitmap graphic files

BOOST – C++ libraries aimed at providing quality software components to developers using Standard Template Library styles.

BSP – bulk synchronous parallelism

C

CADSE – Center for Applications Development & Software Engineering

CAR – Computing and Applied Research

CASC – Center for Applied Scientific Computing

dbCAT – Catalog of Databases

CCA – Common Component Architecture

CCD – College Cyber Defenders

CES – Cost Effective Sampling

CFD – Computational Fluid Dynamics

CFL – Courant-Friedrichs-Lewy

CHAOS – Clustered High-Availability Operating System

CHOMBO – a set of software tools for implementing finite difference methods for the solution of PDEs (see) on block-structured adaptively refined rectangular grids

CIAC – Computer Incident Advisory Capability

CJ – Chapman-Jouguet

COCs – chain of custodies

CORBA – Common Object Request Broker Architecture

CR – compatible relaxation

CRON – not an acronym

C-SAFE – Center for Simulation of Accidental Fires and Explosions

CSGF – Computational Science Graduate Fellowship

CSS – Cascading Style Sheets

CSV – Comma-Separated Values

CU – University of Colorado

CVDB – consolidated vulnerability database

D

DAKOTA – Design Analysis Kit for Optimization and Terascale Applications

DAG – Directed Acyclic Graph

DARPA – Defense Advanced Research Projects Agency

DC – direct current

DCOM – DNT and PAT Computing Applications Division

DDT – deflagration to detonation transition

DG – Discontinuous Galerkin

DHS – Department of Homeland Security

DMS – Design Maintenance Systems

DMTs – data management tools

DoD – Department of Defense

DNA – deoxyribonucleic acid

DNS – direct numerical simulation

DNS – domain name system

DOE – Department of Energy

DPCL – Dynamic Probe Class Library

DPIV – Digital Particle Image Velocimetry

DPOMP – Dynamic Performance Monitor for OpenMP

DSC – Destination Source Correlation

E

EAs – evolutionary algorithms

EC – exponential characteristic

EE – Electronics Engineering

EETD – EE Technologies Division

EEBI – Energy & Environment, Biology & Biotechnology, and Institutional Computing

EFGM – element-free Galerkin methods

EOC – Emergency Operations Center

EPIC – Explicitly Parallel Instruction Computing

ERD – Environmental Restoration Division

ERI – Exploratory Research in the Institutes

ESS – European Simulation Symposium

EGG – elliptic grid generation

F

FAS – Full-approximation Scheme

FASTA – Search database that compares a protein sequence to another protein sequence or protein database, or a DNA sequence to another DNA sequence or DNA library.

FBA – flux balance analysis

FEM – finite element method

FETI-DP – Finite Element Tearing and Interconnecting Dual Primal

FFLO – Fulde, Farrell, Larkin and Ovchinnikov, discoverers of ferromagnetic-superconducting state

FFT – Fast Fourier Transform

FIAT – Framework for Interprocedural Analysis and Transformation

FORTRAN – *for*mula *tran*slator, the first compiled high-level programming language.

FOSPACK – a package developed for automatic discretization and solution of FOSLS

FOSLS – First-Order System Least-Squares

FY – fiscal year

G

GAMS – General Algebraic Modeling System

GCC – GNU Compiler Collection

GEM – Geometric Efficient Matching

GHHE – generalized hyperbolic heat equations

GLR – Generalized left right parsers

GMRES – Generalized Minimal Residual (GMRES)

GNU – GNU's not UNIX

GridDB – A Database Overlay for the Scientific Grid

GUI – Graphical User Interface

H

HCI – human-computer interaction

HIV – human immunodeficiency virus

HkDef – Hacker Defender

HLL – Harten–Lax–van Leer, developers of approximate Riemann solver for the Euler equations of inviscid gas dynamics

HP – Hewlett-Packard

HPC – High-Productivity Computing

HTML – HyperText Markup Language

HTTP – Hypertext Transfer Protocol

HVAC – heating, ventilation, and air conditioning

HYDRA – Hydrological routing algorithm that simulates the flow of water

HYPRE – high-performance conditioners

I

IBM – International Business Machines

ICAT – Internet Catalog of Assailable Technologies

ICCD – Integrated Computing & Communications Department

ICCS – Integrated Computer Control System

ICEO – induced-charge electro-osmosis

Acronyms and Abbreviations

ICMT – Internships in Computational Modeling at the Terascale

ICST – Internships in Computer Science at the Terascale

IDS – intrusion detection system

ISO – International Organization for Standardization

IDX – File format used in ViSUS (see) based on a multiresolution space-filling curve index that allows for fast multiscale data access.

ILP – instruction level parallelism

IMAGE – Integrated Molecular Analysis of Genomes and their Expression

IMPACT – Integrated Map and Particle Accelerator Tracking code

IOR – intermediate object representation

IP – Internet Protocol

IPAM – Institute for Pure and Applied Mathematics

I/O – input/output

ISCR – Institute for Scientific Computing Research

ISMG – Information System Management Group

IT – information technology

ITER – International Thermonucular Experimental Reactor

ITS – Institute for Terascale Simulation

J

JDO – Java Data Objects

JPEG – Joint Photographic Experts Group, a graphics format ideal for photographs, artwork, and paintings; not suited to line drawings, text, or simple cartoons

JSP – Java Server Pages

K

KINSOL – solves nonlinear algebraic systems, see SUNDIALS

KOJAK – POMP-compliant (see) library for profiling and tracing OpenMP (see) applications

KULL – unclassified designation for AX-Division code used to model inertial confinement fusion (ICF, see)

L

LANL – Los Alamos National Laboratory

LDA – mathematical library integer description: On entry, LDA specifies the first dimension of A as declared in the calling (sub) program. Also – local density approximation (first-principles physics approximation within the Linear Expansion in Geometric Objects (LEGO) approach

LDRD – Laboratory Directed Research and Development

LES – Large-Eddy Simulation

LGA – Local Global Alignment

LINPACK– benchmark code for testing supercomputer TF (see) capabilitity by solving systems of linear equations

LLNL – Lawrence Livermore National Laboratory

LOBPCG – locally optimal block preconditioned conjugate gradient

LOCKSS – Lots of Copies Keep Stuff Safe

LOD – level-of-detail

LPI – laser–plasma interaction

LPS – laser proton sources

M

MATLAB – MathWorks proprietary scientific computing and graphics capable programming language, now at V7.0

MCR – Multiprogrammatic Capability Cluster, ALC (see) and MCR are sibling Linux clusters with ~1000 Intel P4 nodes with dual 2.4GHz processors

MDCASK – Molecular dynamics code for radiation damage, to be used as one of the benchmark codes for testing ASC Purple C (see)

MEMS – Micro-Electro-Mechanical Systems

METIS – linear algebra package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices, written in FORTRAN

MHD – magneto-hydrodynamics

MIRANDA – research hydrodynamics code ideal for simulating Rayleigh–Taylor and Richtmyer–Meshkov instability growth. Runs on Fortran 95 with MPI (see). Important for these four factors: incompressible and compressible forms; explicit time solution (Poisson solve for incompressible); Eulerian (fixed), Cartesian mesh; and high-order-accurate derivatives

MLIC – multi-layered image cache system

MOMA – minimization of metabolic adjustment

MPI – Message Passing Interface.

MRF – Markov Random Field

fMRI – functional magnetic resonance imagingMRO – Mars Reconnaissance Orbiter

N

NAI – Nonproliferation, Arms Control, & International Security Directorate at LLNL

NAIC – NAI (see) and Computing Applications Division within CAR (see)

NAMD – object-oriented molecular dynamics code designed for simulation of large biomolecular systems

NASA – National Aeronautics and Space Administration

NCAR – National Center for Atmospheric Research located in Boulder, CO; and the software products NCAR Command Language (NCL) and NCAR Graphics that facilitate forecasting and visualization

NFS – Network File System

NIF – National Ignition Facility

NIFE – NIF (see) and Engineering Computing Applications Division within CAR (see)

NNSA – National Nuclear Security Administration

NR – non-relativistic

O

ODE – ordinary differential equation

OODT – object-oriented data technology

OpenMP – open message-passing, de facto standard for shared-memory parallel programming of scientific applications

ORNL – Oak Ridge National Laboratory

OS – operating system

OSVDB – Open Source Vulnerability Database, a project to catalog and describe global security vulnerabilities, opened a vendor dictionary as a centralized resource on August 31, 2004. OSVDB is sponsored by Digital Defense, Inc.(1999), a private global network security provider; and by Churchill & Harriman (1986), security business partner to mid-market and Fortune 500 companies, headquartered in Princeton, NJ.

P

PAPI – Performance Analysis Programmable Interface sfotware tool: open, cross-platform interface to the performance analysis hardware found in most modern microprocessors

PAT – Physics and Advanced Technologies Directorate at LLNL

PaToH – Partitioning Tools for Hypergraph

PC – personal computer

PCG – preconditioned conjugate gradient

PCMDI – Program for Climate Model Diagnosis and Intercomparison

PCR – polymerase chain reaction assays can amplify a target segment of DNA in suspect biological organisms

vPCR – Virtual PCR (see above)

PDB – Protein Data Bank

PDEs – partial differential equations

PDS – Planetary Data System

PERC – Performance Evaluation Research Center, a SciDAC (see) integrated software infrastructure center with four strategies for maximizing memory hierarchy: discipline-specific benchmarks; performance analysis tools; performance modeling; and performance optimization tools

PerfTrack – a database-based tool for storing, navigating, and analyzing very large amounts of performance data

PF3D – LLNL 3D laser–plasma interaction code

PHP – recursive acronym for "PHP: Hypertext Preprocessor": general-purpose scripting language well suited for Web development and easy to embedded into HTML, commonly used with the Apache HTTP server and included in Red Hat Linux versions

PHP/MySQL – These functions allow the user to access MySQL database servers. More information about MySQL can be found at http://www.mysql.com/

PI – Principal Investigator

PIMS – the LLNL Engineering Directorate Personnel Information Management System

PIW – Promoter Identification Workflow

PLASMA – PLAnetary Scale Monitoring Architecture

PMaC – Performance Modeling and Characterization

PMPI – performance-monitoring programmable interface

POMP –standard performance monitoring interface for OpenMP (see), an API (see) to be called by probes inserted into the application by a compiler, a pre-processor, or via a binary or dynamic instrumentation mechanism

POP – Parallel Ocean Program

POSIX – Portable Operating System Interface incorporates the IEEE and Open Group set of fundamental services needed for the efficient construction of application programs.

PROLINKS – database for co-evolving proteins, used in biological studies and comparisons

PUM – partition of unity methods

PVODE – parallel ODE (see) integrator, a special case of the scaled nonlinear solver, see SUNDIALS

Acronyms and Abbreviations

Q

QBF – query by form

QC – quality control

QCD – Quantum Chromodynamics

R

RANS – Reynolds-Averaged Navier–Stokes equations

RANSAC – Random Sample Consensus

RBF – radial-basis functions

REB – relativistic electron beams

ReiserFS – journalling filesystem included in Linux 2.4, designed and developed by Hans Reiser and his team at Namesys at, creates a single shared environment, or namespace, where applications can interact more directly, efficiently and powerfully. Users can access the filesystem directly rather than building special-purpose layers that run on top of the filesystem, such as databases.

RKPM – reproducing kernel-particle methods

RM – Richtmyer-Meshkov

RMI – Remote Method Invocation.: Java RMI allows the user to invoke a method on an object that exists in another address space — on the same machine or a different one.

RNA – ribonucleic acid: Genetic code is stored in the DNA sequence, which is transcribed into RNA and translated into a polypeptide—proteins, enzymes, or peptide hormones.

ROAM – Real-time Optimally Adapting Meshes: AMR (see) tool in which two priority queues drive split and merge operations that maintain continuous triangulations built from pre-processed bintree triangles

ROSE – name, not an acronym for the Overture Suite preprocessor that recognizes user-defined objects and substitutes optimized code before compilation

RTRAM – Real-Time Radiation Area Monitoring Network

RTS – runtime system

S

SAGE IR – at the request of LANL (see), the SAGE ASC (see) benchmark code is no longer publicly accessible

SAMRAI – Structured Adaptive Mesh Refinement Application Infrastructure

SCaLeS – Science-based Case for Large-scale Simulation

SCD – service-class description

SCE – statistical condition estimation

SCI – Scientific Computing and Imaging Institute at the University of Utah utilizes component-based environments for biomedical computing, computational combustion and other applications

SciDAC – Scientific Discovery through Advanced Computing

SCons – Open Source Software Carpentry tool that is, a next-generation build tool, an improved, cross-platform substitute for the classic `Make` utility with integrated functionality similar to `autoconf/automake` and compiler caches such as `ccache`.

SCWRL – Side Chain placement With a Rotamer Library

SDM – scientific data management

SDSC – San Diego Supercomputer Center

SEGRF – Student-Employee Graduate Fellowship

SGS – Slow Growing Subdivision

SIAM – Society for Industrial and Applied Mathematics

SIDL – Scientific Interface Definition Language

ccSIM – cache-coherent memory simulator

SLURM – Simple Linux Utility for Resource Mnaagement: Open Source, fault-tolerant, and highly scalable cluster management and job scheduling system for large and small Linux clusters

SMPs – symmetric multiprocessors

SPASE – Space Physics Archive Search and Extract

SOAP – an emerging communication standard that encodes remote method invocations using XML payloads over network transport mechanisms such as HTTP

SPH – smoothed-particle hydrodynamics

SPMD – Single Program Multiple Data, parallel programs that use multiple processes running the same code working on different data to solve a problem

SQA – software quality assurance

STL – Standard Template Library

STREAM – Stanford Stream Data Manager

SK – SucKIT – a fully working rootkit that is loaded through `/dev/kmem`

SUIF – Stanford University Intermediate Format

SUNDIALS – (**SU**ite of **N**onlinear and **DI**fferential/**AL**gebraic equation **S**olvers) consists of the following four solvers.

CVODE solves initial value problems for ordinary differential equation (ODE) systems.

CVODES solves ODE systems and includes sensitivity analysis capabilities (forward and adjoint).

IDA solves initial value problems for differential-algebraic equation (DAE) systems.

KINSOL solves nonlinear algebraic systems.

SUPRI – Stanford University Petroleum Research Institute, research group interested in the design of efficient and accurate simulation tools for compositional problems, such as those occurring in gas injection processes

SVM – Support Vector Machine

SWA – segmentation by weighted aggregation

T

TAMM – Terrestrial and Atmospheric Monitoring and Modeling

TAU – Tuning and Analysis Utilities, a program and performance analysis tool framework for high-performance parallel and distributed computing

TCP – transmission control protocol

Teraflop/s or TF – trillion floating-point operations per second

TIFF – Tagged Image File Format, a file format used for scanning, storage, and interchange of gray-scale graphic images

TPS – thin-plate spline

TRANSFAC – Transcription Factor, BIOBASE proprietary database on eukaryotic transcription factors, their genomic binding sites and DNA-binding profiles.

TSTT – Terascale Simulation Tools and Technologies

U

UC – University of California

UCI – University of California, Irvine

UCRP – University Collaborative Research Program

UI – user interface

UIUC – University of Illinois, Urbana-Champaign

URL – Uniform Resource Locator

URP – University Relations Program

UWB – ultra-wideband

V

VBL – Virtual Beam Line

VETFEM – Variable-Element-Topology Finite Element Method, a general-purpose finite in which each element is free to take essentially any polygonal (polyhedral in 3D) shapeVisIT – contraction of Visualize It, a free DOE/ASC (see both) interactive parallel visualization and graphical analysis tool for viewing scientific data on Unix and PC platforms

ViSUS – Visualization Streams for Ultimate Scalability

VTDB – Vulnerability Tracking Database

VTK – Visualization Toolkit

W

WHIRL – Word-based Information Representation Language, a representation system that combines some of the properties of relational databases, and some of the properties of statistical ranked-retrieval systems.

WPI – Worcester Polytechnic Institute

X

XML – eXtensible Markup Language

XWRAP – an XML-enabled software system for semi-automatic generation of wrapper programs for Web sources