SANDIA REPORT

SAND2003-1237
Unlimited Release
Printed April 2003

Computer Science Research Institute 2002 Annual Report of Activities

David E. Womble, Barbara J. DeLap , and Deanna R. Ceballos

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/ordering.htm

SAND2003-1237
Unlimited Release
Printed April 2003

Computer Science Research Institute 2002 Annual Report of Activities

David E. Womble, Barbara J. DeLap, and Deanna R. Ceballos
Computer Science Research Institute
Sandia National Laboratories
P. O. Box 5800
Albuquerque, NM 87185-1110

Abstract

This report summarizes the activities of the Computer Science Research Institute (CSRI) at Sandia National Laboratories during the period January 1, 2002 to December 31, 2002. During this period the CSRI hosted 172 visitors representing 95 universities, companies or laboratories. Of these 56 were summer students or faculty. The CSRI also organized and hosted five workshops with 171 participants. Of these 94 attendees were from 64 universities, companies or laboratories, and 77 were from Sandia. Finally, the CSRI sponsored 14 long-term collaborative research projects.

Contents

Chapter 1. CSRI Overview

1.1. Introduction

The Computer Science Research Institute (CSRI) at Sandia National Laboratories brings together researchers from universities and the national laboratories to solve problems in computer science, computational science and mathematics and to provide new capabilities in modeling and simulation. Participants are also encouraged to develop long-term relationships with laboratory scientists and researchers.

Through the inclusion of university faculty, the CSRI expands the range of expertise and research capabilities that can be applied to problems in modeling and simulation at the national laboratories. Through the interactions with laboratory scientists, researchers from universities and industry are exposed to computational problems that arise at the laboratories in connection with their DOE stockpile stewardship mission.

The Computer Science Research Institute also includes both graduate and undergraduate student programs. These include post-doctoral positions, summer jobs and graduate fellowships. The CSRI encourages students to choose careers in computer science, computational science and mathematics that support directly the challenges of national security programs.

The Computer Science Research Institute complements existing laboratory research programs and university alliances. It provides both a physical and technical focal point for identifying problems, for conducting research and for developing and strengthening interactions between the university and laboratory researchers.

This report presents an overview of the CSRI and describes the projects, visitor programs, and other activities conducted by the CSRI during the period January 1, 2001 to December 31, 2001.

1.2. Technical Focus of the CSRI

A number of potential long-term focus areas for the Sandia program in FY02 are listed and described below. This list represents key technologies in the high-performance massively parallel computing area with the potential to provide substantial benefit in efficiency and accuracy for the Sandia engineering analysis codes and other codes used or being developed for stockpile stewardship applications. Indeed, in some cases the focus areas include breakthrough technology which, when developed, will enable entirely new classes of simulations to be performed. The focus areas are divided into a small number of overarching technical areas, specifically, Algorithms and Computational Mathematics, Enabling Technologies, and System Software.

1.2.1 Algorithms and Computational Mathematics Focus Area:

1.2.1.1 Design and Optimization: As the ability to do "forward" simulations increases, the ability to do the "inverse" problem needs to be developed, e.g., parameter identification and system design, as well as the traditional inverse problems of applied mathematics. Optimization tends to be very application-specific, although some toolkits have been developed that can be generally applied. Current research efforts include work on large-scale optimization, global optimization, and discrete optimization.

1.2.1.2 Linear Solvers: Linear solvers are at the heart of many engineering simulations. There are many algorithms available; however, significant challenges remain. These challenges include the development of scalable preconditioners and preconditioners designed for the specific needs of various applications. Much attention is currently focused on "multiscale" methods and preconditioners as the hope for truly scalable solvers, but a lot of work remains to be done,

especially for unstructured adaptive grids, systems of equations, and complex boundary conditions. Additional work is also needed in many other related areas, including algebraic preconditioners, coupling direct methods for better or more robust convergence, ways to improve performance for machines with deep memory hierarchies, and developing solvers for matrices without the traditional finite-element structure, e.g., in circuit simulation.

1.2.1.3 **Nonlinear Solvers:** Nonlinear solvers often depend on repeated linear solvers, but there are additional research questions. For example, it will be necessary to solve systems with hundreds of variables for 3-D high-fidelity simulations. Present technology is expected to achieve tens of variables within the next year, falling far short of the ultimate requirement. Newton methods and their use in conjunction with preconditioned Krylov methods for specific problems, are of particular interest.

1.2.1.4 **Eigensolvers:** Many scientific and engineering problems require the eigenvalues and eigenvectors of extremely large matrices. Examples of particular interest include modal analysis for structural dynamics, minimum energy eigenfunction calculations in quantum chemistry models, and detecting the onset of turbulence in fluid flow. A common feature of these eigenvalue problems is that the number of eigenvalues required is small relative to the size of the matrices, the matrix systems are often very sparse, and only the action of the matrix on a vector (or several of them) is available. Standard techniques that involve directly factoring the matrix (including sparse direct methods) are often impractical for these problems because of excessive memory and computational requirements. Algorithmic work is needed on scalable eigensolvers, reduced accuracy algorithms, parallel implementations and application-focused algorithmic research.

1.2.1.5 **Algorithms for Differential and Integral Equations:** Differential or integral equations lie at the heart of most engineering simulations. A mathematical analysis of these equations can often reduce the amount of computing needed by simplifying or improving models, choosing better algorithms, or designing better computational experiments. Research topics of interest include coupling or de-coupling of scales, subgrid modeling, asymptotics, bifurcation, and stability analysis.

1.2.1.6 **Complex Phenomena:** This is a very large area, but general goals include identifying and quantifying the effects of uncertainty, developing a predictive capability for complex systems and processes based on computational "experiments," and algorithms that reduce fundamental computational complexity. Topics of interest include stochastic finite elements, sensitivity analysis, experimental design, stability analysis, summability methods, and general methods for handling multiscale (time and space) phenomena.

1.2.1.7 **Adaptivity:** The purpose of the adaptivity area is to develop the methodologies and algorithms for finite element error estimation and adaptive computing, with the general goal being to reduce the cost of computing by increasing the mesh resolution only in areas where needed. Finite element error estimation addresses the discretization error of the finite element solution for some (local) quantity of interest. The goal is to obtain tight bounds or estimates of the error in a way that is relatively cheap to compute (compared to the cost of solving the original problem).

1.2.2 Enabling Technologies Focus Area:

1.2.2.1 **Meshing:** Meshing is a time consuming and difficult part of any engineering simulation, yet the quality of the simulation is highly dependent on the quality of the mesh. Of particular interest are hexahedral meshes and high-quality hex-tet meshes. Research questions here include mesh connectivity, mesh optimization, and mesh refinement. Fully automatic methods and the ability to mesh large complex geometries are of particular interest. The general issue of a robust parallel meshing toolkit remains a high-priority goal of the high-performance computing (HPC) programs at the laboratories.

1.2.2.2 **Automatic Mesh Refinement and Dynamic Load Balancing:** More and more simulation codes include the ability to handle multiple meshes or to automatically refine meshes, and the efficient parallel implementation of these codes will require dynamic load balancing algorithms. Much of the current work is on design and implementation, but as the implementations become available, many new research questions will be raised. The need for dynamic load balancing will be more acute in heterogeneous environments such as will be developed under DisCom2. There will also be the need for "online" load balancing algorithms.

1.2.2.3 **Visualization:** The visualization needs at Sandia have outstripped the abilities of the commercially available tools. New algorithms are needed, and there are many questions to be answered about the appropriate and optimal visualization algorithms that should be used for SSP applications. Also, there is the question of where and when to do the visualization in a large simulation, e.g., as a post-processing operation or as a runtime process, on a stand-alone platform or on the MP machine directly, etc. The answer to these questions will have a major impact on the type of algorithms that are developed for visualization applications. Emphasis in this area will be on scalable visualization tools and algorithms for very large data sets. Distributed, commodity visualization platforms are being developed as an alternative to the costly, nonscalable platforms currently available.

1.2.3 System Software Focus Area:

1.2.3.1 **Operating Systems:** The operating system is a critical component in the effective and efficient use of massively parallel processing (MPP) computers. Current research topics include the use of commodity operating systems (primarily Linux) with modifications and extensions for MPP computers and distributed, cluster-based, virtual MPP computers. As in other areas, a key focus is on scalability. Projects include adding simple memory management and process management to Linux to improve performance while preserving Linux's portability and expandability, improving communication and connectivity, and fault tolerance. The efficient use of SMP nodes within the MPP computing environment is also being considered; this includes the development and implementation of efficient thread and virtual node capabilities and the efficient utilization of resources that are un-partitionable, such as the network interface.

1.2.3.2 **Environments:** An effective environment must address several issues. First, it must provide a fast and "user friendly" environment that allows designers to access easily all of the modeling tools, the data comprehension tools, the problem setup tools and the resources required. Second, it must provide a robust and efficient environment for developers to prototype new methods, algorithms and physics, without redoing major portions of the existing codes. Examples exist of application problem-solving-environments aimed at designers, but these are all "one-of-a-kind" products that are developed for a specific physics code. Examples also exist of component interfaces that allow specific methods to be rapidly prototyped, but again these are not general-purpose, nor are they in common use. Finally, new software tools are needed to model and predict the performance of code and algorithms on MPP computers. The development of tools that combine object-based, Web-centric, client-server technology with high-performance parallel server technology, made available on demand, will also be pursued.

1.2.3.3 **I/O:** Large-scale, simulation-based analysis requires efficient transfer of data among simulation, visualization, and data management applications. Current efforts seek to improve I/O performance of parallel codes by facilitating I/O operations from multiple nodes in parallel through highly portable user-level programming interfaces. This work will involve design, implementation, and testing of a portable parallel file system. Ideally, the parallel file system should include a server side, which may require a particular hardware configuration, and a client side, which is appropriate for use on any ASCI platform. This is not a replacement for MPI-IO. Just as the MPI data movement standard relies on an underlying message-passing or remote memory access protocol, the MPI-IO standard relies on an underlying file system. The goal is to produce at least a

prototype of such a system and, if possible, a product that is appropriate for any future (or current) machine.

1.2.3.4 **Heterogeneous and Distributed Systems:** Parallel computers based on heterogeneous clusters of commodity workstations are starting to appear and will become common. Yet the effective use of these machines presents many research problems. For example, resources such as processors must be scheduled and managed, systems must be fault-tolerant, operating systems must be compatible, protocols for communication must be established, environments must be developed, and the integrated system must be latency-tolerant. The distinguishing feature for work in this area will be scalability to terascale and larger distributed systems.

1.2.3.5 **Architecture:** Our basic architecture is influenced by the highly successful ASCI Red. Cplant™ follows this architecture in spirit if not in details. This project will consider new architectures that will scale to 100 TF, petaflops, and beyond. Among other things is the need for research into interconnect technologies (hardware and software). In addition, for many current and future supercomputing applications, the enormity of the data in processing or post-processing for visualization is a major consideration. This project will consider such questions as how this should affect the architecture of future machines.

1.3. Research opportunities at the CSRI

The CSRI presents many opportunities for collaborations between university researchers and laboratory scientists in the areas of computer science, computational science and mathematics. These include the following

1.3.1 **Collaborative research projects.** The CSRI accepts proposals for collaborative research projects lasting from one to three years. Projects must have a principle investigator and a Sandia collaborator. Projects should address one of the technical areas listed above and the work must be performed on-site at Sandia. Proposals may be submitted to the CSRI director at any time and must be approved by the CSRI executive board.

1.3.2 **Postdoctoral appointments**. The CSRI offers several postdoctoral positions each year. Postdoctoral positions are for one year and are renewable for one additional year. Applications should include a statement of research interests, a resume, and a list of references.

1.3.3 **Summer faculty positions and long-term research visits.** Faculty are invited to consider the CSRI for summer employment or for extended visits. Salaries are generally commensurate with academic year salaries. Proposals to hire research groups including both faculty and graduate students for the summer are also encouraged.

1.3.4 **Faculty sabbaticals.** Faculty may spend all or part of a sabbatical year at the CSRI. Proposals for sabbatical visits are accepted at any time and the salary depends on the normal academic year salary and the sabbatical salary.

1.3.5 **Summer student positions**. Students are encouraged to apply for summer positions at the CSRI. Employment is generally for eight to twelve weeks. Students may be associated with a research group (see Summer Faculty Positions above) or may apply independently.

1.3.6 **Graduate Fellowships**. The CSRI sponsors graduate fellowships through the Krell Institute and the National Physical Sciences Consortium (NPSC). For more information, students can contact the Krell Institute or the NPSC directly, or they may contact the CSRI.

1.3.7 **Short term visits.** The CSRI hosts approximately 100 research visits lasting from several days to weeks. The CSRI generally reimburses visitors for travel expenses.

1.3.8 Workshops. The CSRI hosts one to five workshops per year. Workshops are generally co-organized by Sandia staff and university researchers. Workshop proposals are accepted at any time.

Chapter 2. Research Projects

This chapter summarizes the major projects funded by the Computer Science Research Institute. Project proposals can be submitted to the CSRI at any time and are considered for funding by the Executive Board. Criteria for funding projects include technical excellence and quality of the work, impact on the ASCI program, and the strength and breadth of collaborations with Sandia technical staff. Generally these projects are intended to develop new collaborations with university researchers. Also, the work described in these proposals is performed on-site at Sandia National Laboratories, which improves both the strength and breadth of the collaborations. This distinguishes these projects (and the CSRI) from the normal contracting process at Sandia National Laboratories. The projects are listed in alphabetical order by the PI.

1) Scalable Fault Tolerance through Compiler-Driven Communication-Induced Checkpointing
   Lorenzo Alvisi, The University of Texas at Austin
2) EMU code for shape optimization
   Florin Bobaru, University of Nebraska-Lincoln
3) Efficient Implementation for Overlapping File Access in MPI-IO
   Alok Choudhary, Northwestern University
4) Surrogate-based Optimization
   John Dennis Jr., Rice University
5) User Friendly, Cache Aware, Parallel PDE Solvers
   Craig C. Douglas, University of Kentucky and Yale University
6) Multiscale-Multiphysics Computational Framework
   Professor Jacob Fish, Rensselaer Polytechnic Institute
7) Optimization Algorithms
   Lisa Fleischer, Columbia University
8) Terascale Simulation-constrained Optimization
   Omar Ghattas, Carnegie Mellon University
9) Research in Finite Element Simulations
   Max D. Gunzburger, Florida State University
10) Development of a Vorticity Method for the Solution of Nonlinear Viscous Shear Flows
    Marc S. Ingber, University of New Mexico
11) Supporting MPI Collective Communication Operations with Application Bypass
    D. K. Panda, Ohio State University
12) Large Scale Eigenvalue Methods and Model Reduction of Second Order Systems
    D.C. Sorensen, Rice University
13) PIM Prototype Testbed
    Dr. Thomas Sterling, Caltech
14) Pattern Search for Nonlinear Optimization
    Virginia Torczon, College of William and Mary

Title: Scalable Fault Tolerance through Compiler-Driven Communication-Induced Checkpointing

PI: Lorenzo Alvisi, The University of Texas at Austin

Investigators: Calvin Lin, The University of Texas at Austin
Jeff Napper, The University of Texas at Austin
Thomas Bressoud, Bell Laboratories

Dates: April 25, 2002 – April 24, 2004

CSRI POC: Patty Hough

Project Summary:

We propose to investigate new algorithms for supporting low-overhead fault tolerance for long-running, distributed scientific applications. In particular, we propose to derive algorithms that exploit a combination of compiler analysis and program annotations to determine an efficient schedule for checkpointing the state of distributed applications.

Motivation: The goal of this research is to develop effective techniques for providing fault tolerance to very large distributed computing systems such as those employed by the ASCI program. The need for fault tolerance in these systems is becoming increasingly clear. For instance, measurements taken for the ASCI Blue Mountain cluster at Los Alamos National Labs indicate that the 6000-node system has a mean-time-to-failure (MTTF) of about 1.6 hours. This figure, which is consistent with the prediction that the MTTF for an off-the-shelf single processor system is about 10,000 hours, is already sufficiently low to jeopardize the usability of the cluster. Barring an unforeseen improvement in the reliability of off-the-shelf processors, it will become practically impossible to exploit the full computing potential of the larger ASCI systems, such as Sandia's Red Storm, expected in the near future. Still, the cpu-hungry users of ASCI machines would find unacceptable any solution that requires dedicating a considerable chunk of processors to achieving fault tolerance (for instance, by using active replication). These users are also unwilling to adopt fault tolerance techniques that require them to rewrite their highly tuned code.

Background: Rollback recovery protocols are ideally suited to provide transparent fault tolerance in a distributed system that needs to tolerate common failures with little overhead in terms of dedicated resources and execution time. Briefly, these protocols record information during process execution that is then used during recovery to restore a crashed process to a state consistent with that of the other application processes. One of the objectives of our research agenda over the past five years has been to increase our understanding of these techniques, using a combination of algorithmic work, system building, and experimental analysis.

In this proposal, we focus on a specific rollback recovery technique, called *communication-induced checkpointing* (CIC). To put the proposed research in context, we first discuss how communication-induced checkpointing compares with other approaches used to save the global state of distributed computations. We then describe the results of some of our previous work, which has provided the first experimental evaluation of the performance of CIC protocols.

1 Checkpoint-Based Rollback-Recovery The simplest form of rollback recovery is based on checkpointing the state of the distributed application. Checkpoint-based rollback-recovery techniques fall into three categories.

Uncoordinated checkpointing allows each process to decide independently when it is most convenient to take a checkpoint. For example, a process may reduce overhead by taking checkpoints when the amount of state information to be saved is small [9,12]. However, there are several disadvantages to uncoordinated checkpointing. First, during recovery, processes must search for a consistent global checkpoint out of the set of individual checkpoints. If this search is unsuccessful, it can ignite a series of cascading rollbacks,

known as the *domino effect* [14], which may cause a substantial loss of useful work. Second, a process may take a *useless* checkpoint, i.e. one that will never be part of a global consistent state, thus incurring overhead without advancing the global state from which recovery can be initiated. Third, each process must maintain multiple checkpoints and periodically invoke a garbage collection algorithm.

Coordinated checkpointing [5,10] requires processes to orchestrate their checkpoints in order to form a consistent global state. Every process recovers from its most recent checkpoint, thereby simplifying recovery and avoiding the domino effect. Also, coordinated checkpointing requires only one permanent checkpoint per process on stable storage, reducing storage overhead and eliminating garbage collection. However, processes must incur the overhead of achieving coordination, and all processes send their checkpoints to storage at approximately the same time. The latter characteristics raise serious doubts about the scalability to thousands of processors.

Communication-induced checkpointing (CIC) protocols [2,3,6] avoid the domino effect without requiring all checkpoints to be coordinated. Processes take two kinds of checkpoints: *local* checkpoints can be taken independently, while *forced* checkpoints guarantee the progress of the recovery line by preventing useless checkpoints. Unlike coordinated checkpointing, CIC protocols do not exchange coordination messages but instead piggyback information on each application message. This information is used to determine if communication and checkpoint patterns can lead to the creation of useless checkpoints, and a forced checkpoint is taken to break these patterns. This intuition has been formalized in an elegant theory based on the notion of *Z-cycles* [11]. A key result is that a local checkpoint is useless if it is involved in a Z-cycle [7,11].

CIC protocols are believed to have several advantages. First, they allow processes notable autonomy in deciding when to take checkpoints. Second, they are believed to scale up well with the number of processes since they do not require global checkpoints. Finally, CIC protocols never generate useless checkpoints. Unfortunately, there is a price to pay for these advantages. In order to prevent useless checkpoints, the information piggybacked on application messages occasionally induces the processes to take forced checkpoints before they can process the messages. Second, processes have to pay the overhead of piggybacking information on top of application messages, and they also need to keep several checkpoints on stable storage.

2 An Analysis of Communication-Induced Checkpointing The above discussion suggests that CIC protocols promise to scale to systems with thousands of nodes. Unfortunately, the literature presents only a qualitative analysis of CIC protocols, leaving open the question of how these protocols truly perform in practice. To shed some light on this question, we have performed the first experimental analysis of CIC protocols [1].

Our study compared the performance of three CIC protocols [2,3,6] for a set of applications in the NAS NPB2.3 benchmark suite [4]. Our analysis has shown that the theoretical advantages of CIC protocols do not materialize in practice. In particular, our results show the following:

1. CIC protocols that use an eager approach to preventing Z-cycles by taking forced checkpoints whenever they suspect the formation of a Z-cycle perform worse than lazy protocols that use a time-stamping function to prevent a Z-cycle at the last second.
2. CIC protocols do not scale well with a larger number of processes. We have found that the number of forced checkpoints increases almost linearly with the number of processes.
3. A process takes at least twice as many forced checkpoints as local ones. Therefore, the touted benefit of allowing processes to take independent checkpoints does not materialize.
4. There is a considerable unpredictability in the way CIC protocols behave. The amount of stable storage required, performance overhead, and number of forced checkpoints depend greatly on the number of processes, the application, and the communication pattern. This unpredictability makes of CIC protocols more cumbersome than other alternatives.
5. A successful placement policy of local checkpoints must be dynamic, must account for forced checkpoints, and must adapt to changes in the application behavior.
6. CIC protocols seem to perform best when the communication load is low and the pattern is random. Regular, heavy load communication patterns seem to fare worse.

Proposed Work: We propose to explore a new approach to checkpoint-based rollback recovery that we believe can yield in practice the theoretical advantages that traditional CIC protocols fail to deliver.

The principal reason for this failure is that CIC protocols are prone to generating a large number of forced checkpoints. In these protocols, a process determines whether to take a forced checkpoint on the basis of information piggybacked on the application messages that it receives. This information however is insufficient to provide the recipient with the knowledge necessary to make an optimal decision. In particular, the piggybacked information does not allow the recipient to rule out communication patterns that would lead to Z-cycles. As a result, recipients conservatively take forced checkpoints that would not be otherwise necessary, negatively impacting performance. A second negative consequence of frequent forced checkpoints is that they nullify CIC's promise of maintaining the advantages of uncoordinated checkpointing.

To address these problems, we propose to develop algorithms that allow a source-to-source compiler to determine an efficient checkpointing schedule. This approach can potentially result in significant advantages. For example, the compiler can build an internal representation of the communication patterns exhibited by an application and use it to place checkpoints appropriately in order to build a consistent recovery line and prevent the domino effect. This can yield the same advantages of coordinated checkpointing (easy garbage collection, minimal storage requirements) but without the coordination overhead and reduced scalability. In particular, this approach would allow skewing the time at which different processes checkpoint their state, reducing the risk of overwhelming the I/O to stable storage. Also, the compiler can try to build a "smart" recovery line by scheduling local checkpoints at the most appropriate times in the execution of each the application's processes. To help the compiler, programmers could also annotate their program with hints that indicate times in the execution of a process where checkpointing would be most convenient.

The steps towards realizing this vision are the following:

- develop an internal representation of the *space-time diagrams* [8] used to express the events that occur in the execution of a distributed application,
- use this representation to develop algorithms that build recovery lines by scheduling independent checkpoints in the execution of the application processes,
- define criteria to determine the most preferable recovery line among a set of alternative ones and develop algorithms based on these criteria,
- instantiate these algorithms in a source-to-source compiler that will automatically insert calls to *Egida* [1, 13], a library for low-overhead fault-tolerance developed at UT Austin, into an application and demonstrate on a Sandia application code.

References:

[1] L. Alvisi, E. N. Elnozahy, S. Rao, S. A. Husain, and A. De Mel. An analysis of communication-induced checkpointing. In *Proceedings of the IEEE Fault-Tolerant Computing Symposium (FTCS-29),* pp. 242-249, Madison, WI, June 1999.

[2] R. Baldoni, F. Quaglia, and B. Ciciani. A vp-accordant checkpointing protocol preventing useless checkpoints. In *Proceedings of International Symposium on Reliable Distributed Systems,* pp. 61-67, West Lafayette, IN, October 1998.

[3] D. Briatico, A. Ciuffoletti, and L. Simoncini. A distributed domino-effect free recovery algorithm. In *Proceedings of the IEEE International Symposium on Reliability Distributed Software and Database,* pp. 207-215, December 1984.

[4] NASA Ames Research Center. NAS parallel benchmarks. http://science.nas.nasa.gov/Software/NPB/1997.

[5] E. N. Elnozahy, D. B. Johnson, and W. Zwaenepoel. The performance of consistent checkpointing. In *Proceedings of the Eleventh Symposium on Reliable Distributed Systems,* pp. 39-47, October 1992.

[6] J. M. Hélary, A. Mostefaoui, R. H. B. Netzer, and M. Raynal. Preventing useless checkpoints in distributed computations. In *Proceedings of IEEE International Symposium on Reliable Distributed Systems,* pp. 183-190, 1997.

[7] J. M. Hélary, A. Mostefaoui, and M. Raynal. Virtual precedence in asynchronous systems: concepts and applications. In *Proceedings of the 11th Workshop on Distributed Algorithms.* LNCS press, 1997.

[8] L. Lamport. Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM,* 21(7):558-565, July 1978.

[9] C. C. Li and W. K. Fuchs. Catch: Compiler-assisted techniques for checkpointing. In *Proceedings of the 20th International Symposium on Fault-Tolerant Computing,* pp. 74-81, 1990.

[10] G. Muller, M. Hue, and N. Peyrouz. Performance of consistent checkpointing in a modular operating system: Results of the ftm experiment. In *Proceedings of the First European Dependable Computing Conference (EDDC-1),* pp. 491-508, October 1994.

[11] R. Netzer and J. Xu. Necessary and sufficient conditions for consistent global snapshots. Technical Report 93-32, Department of Computer Sciences, Brown University, July 1993.

[12] J. S. Plank, M. Beck, and G. Kingsley. Compiler-assisted memory exclusion for fast checkpointing. *IEEE Technical Committee on Operating Systems Newsletter, Special Issue on Fault Tolerance,* pp. 62-67, December 1995.

[13] S. Rao, L. Alvisi, and H. Vin. Egida: An Extensible Toolkit For Low-overhead Fault-Tolerance. In *Proceedings of the 29th Fault-tolerant Computing Symposium (FTCS-29)*, Madison, Wisconsin, June 1999, pp. 48-55.

[14] R. B. Strom and S. Yemeni. Optimistic recovery in distributed systems. *ACM Transactions on Computer Systems,* 3(3):204-226, April 1985.

SNL Collaboration: This research will be conducted in collaboration with a number of technical staff members at Sandia. Curtis Janssen and Matt Leininger will participate in coupling software with the Massively Parallel Quantum Chemistry (MPQC) code and developing MPI-based tools. Mike Goldsby and Edward Walsh will participate in characterizing Sandia applications as well as in developing and implementing the software interfaces. Lee Ward will provide hardware and OS expertise for Sandia's ASCI machines. Patty Hough and Vicki Howle will be developing numerical algorithms that take advantage of the tools developed.

Most of the visits will take place at the California site; however, at least one visit will be to the New Mexico site in order to establish new collaborations with others interested in fault tolerance. In particular, we intend to develop working relationships with researchers in 9200 and with the developers of ASCI finite-element applications such as ALEGRA and SIERRA. The work conducted under this proposal will be incorporated into software tools that will be made available to Sandia and others in the ASCI community.

Title: EMU code for shape optimization

PI: Dr. Florin Bobaru, Department of Engineering Mechanics
University of Nebraska-Lincoln

Dates: August 5, 2002 – August 30, 2004

CSRI POC: Steward Silling

Project Summary:

My research interests are in the general area of computational mechanics. More specifically, I am interested in developing numerical algorithms for optimization of materials and structures. I have recently shown the benefits one can obtain in shape optimization of elastic and thermoelastic bodies if, instead of the classical Finite Element Method, a mesh-free approach is used. I am interested in new algorithms that would push the current limits in optimal design of materials and structures. Combining shape and material optimization, or shape and topology optimization are issues I am currently considering. With powerful algorithms for optimization, one would expect breakthroughs in designing individual components of complex systems ranging from new paradigms for MEMS design to improved designs of large scale systems such as better thermal protective shields for space shuttle or damage tolerant armors. Nature has been perfecting and evolving systems of higher and higher complexity to adjust to challenging environments. My goals are to find algorithms that would use the fundamental minimum principles of mechanics and physics to produce designs of comparable efficiency and versatility as nature's answers to complex environments

Research Proposal: This project plans to advance the methods of shape design sensitivity analysis to problems involving regular boundaries subject to singular transformations, and singular boundaries under regular or singular transformations (see Figure 1). The peridynamic theory developed by Dr. S.A. Silling of Sandia will be the analysis tool instead of the classical theory of elasticity.

The boundaries to be considered are either interior boundaries (that may represent cracks, material interfaces, phase changes, etc.) or outer boundaries. Due to this generality of the formulation, the same framework can be used to treat shape optimization problems as well as material optimization of two or multi-species composites (searching for the best shape of inclusions in an elementary volume element by varying the interface).

The problem of calculating sensitivities of a response functional to singular transformations of the boundary (inner or outer, regular or singular) of a body subject to mechanical loading has not been studied so far due to intrinsic difficulties of treating singular boundaries and discontinuous fields in formulations of boundary value problems. In particular, the classical formulation of linear fracture mechanics gives rise to the well-known square-root singularity of the stress field. Clearly, sensitivities of singular boundary (the crack for example) shape variations cannot be addressed in the classical setting.

The problem of calculating sensitivities of a response functional to singular transformations of the boundary (inner or outer, regular or singular) of a body subject to mechanical loading has not been studied so far due to intrinsic difficulties of treating singular boundaries and discontinuous fields in formulations of boundary value problems. In particular, the classical formulation of linear fracture mechanics gives rise to the well-known square-root singularity of the stress field. Clearly, sensitivities of singular boundary (the crack for example) shape variations cannot be addressed in the classical setting.

The peridynamic formulation in elasticity eliminates the difficulties of the classical continuum theories when dealing with spatial derivatives of discontinuous fields. In this formulation, long-range forces are represented by a pairwise force function with compact support acting between two nodes of the discretization. Cusp-like profiles are obtained for the opening of a crack and infinite stresses at the crack tip are eliminated. The ad-hoc assumptions that were proposed in the past to correct the unphysical infinite stresses from the classical linear fracture mechanics theory are now no longer needed. The same model is used in the peridynamic formulation on and off the discontinuity present in the fractured elastic body. The

mesh-free character of the peridynamic theory makes it extremely suitable for numerical treatment of shape optimization problems based on sensitivities.

We propose to develop analytical and numerical sensitivities for singular transformations in elasticity using the peridynamic formulation. Based on these sensitivities it will be possible to develop shape optimization of elastic bodies with singular boundaries including bodies with cracks, notches, cusp-like shapes, etc. At the same time, material optimization of composite (searching for optimal shapes of inclusions) or Functionally Graded Materials (finding the optimal material gradation) will be possible since the boundaries in our formulation are both external and internal. Another possible application would be the sensitivity and uncertainty estimation for elastic bodies with large numbers of random cracks.

Collaborations: The current proposal is seeking collaboration with the author of the peridynamic theory, Sandia's Dr. Stewart A. Silling. Dr. Silling is also the primary developer of the EMU code that implements the peridynamic theory. Dr. Silling is strongly supporting the current proposal and has provided a supporting letter for a proposal submitted by the author of this proposal to the DOE Early Career PI Program. EMU is a mesh-free method designed to model cracks, discontinuities, and other singularities that may exit in a body initially or emerge as a result of deformation. In addition, the EMU code is able to represent interfaces between materials by modifying the bond properties between the discretization points. Modeling complex heterogeneous materials is one of the areas in which the peridynamic theory seems to have a great potential. We propose to couple the sensitivity calculations for singular transformation to the EMU code and build an optimal design code based on the peridynamic formulation. Ideally, when completed, the peridynamic code and the new sensitivity formulations will interface and enhance the capabilities of the DAKOTA system developed at SANDIA.

Length of Participation and Location of visit

The proposed length of participation is for one-month summer faculty for three years at the Sandia National Laboratories, Albuquerque, NM. The starting date for our collaboration is 1st of August 2002. The visit during the 2002 summer will be focused on familiarization with the numerical implementation of the peridynamic theory, the EMU code. The basics ingredients of the sensitivity analysis with singular transformations for domains with singularities will be developed. For the remaining part of 2002 and before the summer of 2003, our collaboration will continue in various forms: Dr. Silling will be invited to present a seminar at the Department of Engineering Mechanics at University of Nebraska-Lincoln in the Fall of 2002. The author of the current proposal has a Master student who will be working to develop the computer programs for a proper coupling of the shape sensitivity analysis into Sandia's EMU code. Dr. Silling will be invited to participate in the graduate committee of the Master student.

The second visit (June 2003) will focus on the development of a series of optimization algorithms for shape and material optimization based on the developed sensitivity programs. We will test peridynamic formulation against some classical problems in shape optimization and, moreover, we will attempt solutions to singular problems never considered before. Parallel versions for the sensitivity calculations will be advanced. The third visit (August 2004) will be dedicated to large-scale simulations involving hundreds and possibly thousands of design variables and the assessment of the benefits of the peridynamic formulation in optimal design of complex systems (complex geometries and complex material behavior). Problems of different scales will be considered. Comparisons with results provided by the DAKOTA system will be made.

Title: Efficient Implementation for Overlapping File Access in MPI-IO

PI: Prof. Alok Choudhary, Northwestern University

Dates: February 19, 2002 – August 15, 2003

CSRI POC: Lee Ward

Project Summary:

Numerous studies of the I/O characteristics of parallel applications have shown that in most cases multiple processors access shared data objects. However, the partitioning and layout of the shared data objects to be stored in the memory can be different from its physical layout on disks, in which case the I/O performance can significantly degrade. In order to solve such problem, collective I/O was proposed in which each participated processor performs I/O on behalf of other processors and, then, all processors use available interconnection network to exchange the data so that each processor obtains the desired data. This technique has been adopted by MPI-IO, the I/O part of the MPI-2 standard, whose goal is to design a high-performance I/O library for parallel and distributed I/O. Collective I/O operations may have the situations that multiple processors issue concurrent read/write requests to overlapped regions in the shared file. The results of writing to the overlapped regions can be defined as written by one of the processors, an aggregation of all processors, or undefined. The mechanism of solving this problem, called atomicity, is implemented differently across file systems, which may involve locking shared files to guarantee the desired results. However, file locking reduces the parallelism of performing concurrent I/O and becomes the bottleneck of the collective operations. We propose to develop techniques to solve this problem. We plan to design a mechanism that automatically detects overlapping region accesses in the collective I/O operations in order to reduce the number of file locking, pass proper parameters to file locking mechanism, or even remove the locking.

ROMIO, a portable MPI-IO implementation, provides uniform parallel I/O APIs to access files on different file systems. Internally, ROMIO is built on top of ADIO which is implemented separately on each file system using its native machine-dependent I/O library. When the underlying file system is the Network File System (NFS), ADIO use the file lock mechanism, fcntl, to perform non-coherent client-side caching on local processor memory by NFS default. This effect can prevent the file consistency problems occurred in other processors. ADIO disables client-side caching by locking the portion of the file being accessed so that the updated data can be flushed to the disk and viewable by other processors. This implementation results a pair of lock and unlock wraparound every native NFS read/write calls, even for those collective I/O performing only nonoverlapping region access.

Task 1. Identify the Overlapping Regions from MPI Derived Datatype: One of the powerful mechanisms provided by MPI that describes the data layout of the messages is called datatype. Datatypes are of two kinds: primitive and derived. Primitive datatypes are those basic datatypes in programming languages, such as integers, floatingpoint, and characters. Derived datatypes are provided by users which is created consisting of multiple primitive datatypes located either contiguously ore non-contiguously. Derived datatypes are used commonly to define individual processor's file views and its I/O buffer layouts for mapping data between file and memory when performing parallel I/O. Since each processor in a parallel I/O operation has an independent file view to the shared data, data access to the overlapped regions in the file may occur. For example, a parallel climate simulation using Laplace equation partitions its problem domain into several regions where each processor is responsible for one region. Practical implementation would have extended the boundary of each region to cover slightly more area, called ghost, which overlaps with the regions responsible by other processors. Given a set of derived datatypes from the processors participated in the I/O, overlapped area should be able to be detected. In this task, we will focus on the study of finding the overlapping access on the MPI-IO derived datatype level.

Task 2. Contiguous Region in File Access: In this task, we will study the case of contiguous file access from all processors. In this case, the file view of each processor in a parallel I/O is a single contiguous chunk of data in the shared file. Once the overlapped regions are identified, file I/O on exclusive regions

can be performed and inter-processor communication will be used to transfer the ghost data to its requiring processors. Therefore, locking file mechanism can be avoided and increase the I/O performance. We plan to implement this task in the unit of bytes in order to take care any possibility of user defined derived datatypes.

Task 3. Non-contiguous Region in File Access: In this task, we will extend the previous task to cover more complex file access. As the example described in Task 1, each processor in a two-dimensional Laplace equation application need to access multiple overlapped regions in order to obtain the responsible two-dimension sub-domain. Each of the multiple file access in a processor is noncontiguous from the other file access. This implementation of detecting overlapped region can be incorporated into MPI's collective I/O so that ghost data can be transferred properly while performing exclusive file access to keep the I/O cost minimum. In order to deal with more general cases of I/O, we will devise a solution in byte level to cover any types of data mapping between file and memory buffer. Our plan is to minimize the use of file locking and, therefore, reduce the file access cost while using inter-processor communication to achieve the desired parallel data partitioning.

Title: Surrogate-based Optimization

PI: John Dennis Jr., Rice University

Dates: August 14 – 16, 2002

CSRI POC: Paul Boggs

Project Summary:

The purpose of this proposal is to fund the PI's interaction with Sandia National Laboratories at NM and CA. These interactions should be of two sorts. First, algorithmic advances and interesting practical experience of the PI and his collaborators in universities and industry would be transferred to the labs, and conversely, lab problems and experience would inform the research of the PI and his collaborators. Bidirectional technology transfer is our goal and firm belief. We seek to be both brokers and developers; to provide support and receive feedback. The point of the visits for which support is requested is to increase the total pool of knowledge about engineering design problems of the particularly nasty sort discussed below.

The class of optimization problems that we target is largely from engineering design and widely regarded by practicing designers as intractable. We have focused for several years not on solving problems better, but on solving problems that can not be solved at all well by existing methods. We have resisted the temptation succumbed to by some university engineering groups of designing elaborate (soft) design support systems based on naive expectations of the spontaneous generation of supporting algorithms. Instead, we have focused on providing the rigorous mathematical infrastructure to underpin design support systems such as Boeing's Design Explorer and Sandia's DAKOTA, though to date, we have had measurable impact only on Design Explorer, where our influence has been considerable. Dr. Greg Shubin, Head of Mathematics and Engineering Analysis, or Dr Evin Cramer, Technical Fellow of Boeing Phantom Works can verify our contributions. <shubin or cramer>@redwood.rt.cs.boeing.com.

Another tenet of our research is that engineers generally understand their problems better than we can, and so while we do provide default choices, all our work has been directed by the rule that engineering intuition should be able to be seamlessly incorporated into our algorithmic framework. We take over only when the user's approach has reached the limit of its resolution, and we proceed to "clean up" the final design or to get the designer out of a rut fallen into by conventional approaches.

Our algorithms are carefully structured so that if a user has a favorite method for solving the surrogate problem, then it can be provided to our FOCUS software as a SEARCH procedure in a direct search method. Thus, we provide direct search methods to act as meta-algorithms supporting user favored heuristics. So far, we have the algorithms and software in place to deal with general nonlinearly constrained optimization for problems of the type we see often in engineering design:

f(x), C(x) are expensive and have few correct digits. This happens because there is an underlying expensive state equations that needs to be solved for a given x to generate the state variables that are used to evaluate the optimization problem functionals. Evaluation may fail (f(x)=infinity) expensively and unexpectedly for feasible x and f and C may be discontinuous even when they are defined. This happens because the state equation solver mentioned above is often a kludgey coupling of table lookups and state equation solvers from different physical disciplines. For example, one might use successive replacements to try to couple a structures solver with a dynamics solver. This procedure may work for some x, but it may not for one nearby. In a Boeing helicopter rotor design problem, the function failed to return a value about 67% of the time. Someone unfamiliar with commercial reality might suggest that the solvers should be tightly coupled to smooth out this behavior. Indeed, a part of the economic justification for making such an effort can derive from a proof of concept we might provide by the techniques suggested here. Before abandoning legacy solvers to reimplement a solver more amenable to SAND type approaches, try this approach to estimate the economic gains. Evaluation will usually fail unless some selected simple constraints hold. This happens when there are simple bounds, like nonnegativity, on some of physical parameters. It means

that we must be able to be always feasible with respect to some constraints, though for efficiency, we would like to allow infeasibilities during the course of the interaction in any constraints for which that is reasonable.

In addition, since we are usually dealing with interpolatory surrogates like polynomials, (kriging) splines, or neural nets, there is a fairly low limit on the number of decision variables we can treat. However, this is a problem with constructing interpolants more than with the underlying approach. Thus, if the surrogates come from a less detailed description of the underlying physics, then we can treat more decision variables. We are aiming for an ExxonMobil problem with a couple hundred each discrete and continuous variables, and for which the surrogate is based on a simplified model of flow-in-porous media. Still, the largest problem we actually have solved was 256 continuous decision variables.

To meet these challenges of our problem class, we have been able, for the very nice Lewis and Torczon barrier algorithm (f(x) is infinite if x is not feasible), to use a crucially different proof to strengthen their convergence results for nonlinear optimization under simple linear constraints by dropping their assumption of continuous differentiability on an open set containing the feasible region. Indeed, if the initial iterate has finite problem function values, then we show the existence of a limit point even when the functions are discontinuous and extended real valued. If the problem is locally smooth at such a point, then we show appropriate optimality conditions from the Bourbaki/Clarke nonsmooth analysis. In addition, the proofs based on our weaker hypotheses are much shorter and simpler.

Our general constrained algorithm FPS adapts Fletcher's new penalty function-free filter step acceptance idea to direct search methods. Thanks to our improved convergence analysis mentioned above, we can allow discontinuities and extended values in the function to which the algorithm is applied, Thus, we can apply our NLP filter method directly to the Lewis and Torczon barrier function and thus enforce every iterate feasibility for simple linear constraints with no need for extra complication in the theory. This FPS algorithm is being used at Boeing in the current release of the Boeing Design Explorer package. It had a great success in wing planform design.

The focus of our efforts supported by research funds will be to:

- Implement our mixed discrete/continuous variable algorithm MVP in FOCUS during the next 3 years. This is the longest-term goal of the proposal. The difficulty lies in the fact that our main focus is on categorical discrete variables and nonlinear nonconvex problems. These are very common in engineering design, and they are usually handled by heuristic parameter studies. For our purposes, a categorical variable is a variable, which is an input to the simulation driving the optimization, but it has the property that if its input value is not from a prescribed finite set, then the simulation will not run and the function or the constraints cannot be evaluated. This property precludes the use of continuous relaxations for discrete variables, and hence branch and bound. Our MATLAB MVP has been quite successful on some artificial test problems as well as on a problem from cryogenic engineering design from a mechanical engineer at the Automotive Research Center at Ann Arbor. In that problem, MVP increased or decreased the number of heat intercepts and changed the neighboring insulation material of as the iteration proceeded. This means that number of heat intercepts, which determines the number of optimization variables was itself an optimization variable. We were able to obtain a 60% reduction in the require power by our approach over the best published result that chose the categorical variables by a parametric study. We regard this as highly promising behavior. We continue work on the algorithm. The software difficulty is in designing a user interface to capture the user's domain specific knowledge concerning the relationships between variables like different insulation materials.

- Extend our MVP work to generally constrained mixed variable problems. We have long advocated using the Ferris-Mangasarian variable distribution techniques to extend our algorithms to higher dimensions. Those techniques can be thought of as an adaptive approach to the linking variables long used by design engineers and explained in the book by Vanderplaats.

- Continue development of multiobjective filter based tools to support engineering decisions on trading off various competing objectives. This is the holy grail. Our preliminary MATLAB experiments indicate that the surrogate/filter approach may extend current capabilities, but there is much to do here. We would like to provide a surrogate based representation of the trade-off surfaces, which we would refine as the user homes in on the interesting region.

The visits supported by this funding would aim to:

- Work with Bill Hart on incorporation of evolutionary approaches to SEARCH for all our algorithms. It is important to provide effective default SEARCH procedures for casual and evaluative use or for naive users. Bill Hart's work seems very well suited in robustness and general applicability to this application. The way our framework is designed, we give the user the opportunity to embed their own optimization procedures to be applied to the surrogate to identify promising candidates for improved designs. This is an important feature of our approach. Many designers have ad hoc approaches.

- Work with Mike Eldred on algorithmic enhancements to DAKOTA.

- Work with Paul Boggs, Patty Hough, and Tammy Kolda on enhancements to the PDS/trust-region algorithm. Specifically, we advocate the use of the local quasi-Newton quadratic model as a surrogate, and we suggest the restriction of the SEARCH based on this surrogate to a lower dimensional subspace as in the work of Byrd, Schnabel and Schultz or in the work of Boggs, Kearsley, and Tolle.

We propose to visit each Sandia installation for at least 2 weeks each year starting in 2002 and continuing until 2004. It would seem to be useful to the Institute that there is no special reason why these visits need to be in the crush of summer visits since the PI will have no teaching duties after the end of the 2000-2001 academic year. At that time, he will retain a Research Professor position at Rice and hold Affiliated Professor status in Industrial Engineering and in Applied Mathematics at the University of Washington in Seattle. All of these positions are functionally equivalent and allow the holder to PI grants and direct graduate students.

Title: User Friendly, Cache Aware, Parallel PDE Solvers

PI: Craig C. Douglas, University of Kentucky and Yale University

Dates: January 1, 2002 – September 30, 2002

CSRI POC: Jonathan Hu

Project Summary:

This project is joint work with Jonathan Hu, Ray Tuminaro, and Jaideep Ray of Sandia/California. Danny Thorne, a graduate student at the University of Kentucky, is also involved. Danny is supported for 10 months at Kentucky through ASCI funds and also spends two months a year as a Sandia summer student. The codes we are producing are being tested on combustion problems that Jaideep Ray is providing.

Many cache aware algorithms have several shortcomings. Implementing them is intrusive to already existing codes and usually leads to ugly, hard to maintain spaghetti style programs. In addition, the modified codes are usually not portable, even to upgraded CPUs in the same processor line. Even new codes are hard to construct and maintain. All of these bad features are normally independent of the type of mesh structure that is used.

A great deal of progress has been made in eliminating these shortcomings for typical academic research codes. Part of the work leads to codes that are two to three times faster than traditional codes for a variety of architectures, but are not optimal for any particular processor/cache combination. The approach is to minimize the number of parameters to one (usable cache size) and to live with a code that is much faster, i.e., one that is "good enough, but portable" instead of perfect. In short, a small amount of inefficiency goes far in making a cache aware code useful and easy to tune.

Producing cache aware codes for adaptively refined meshes is hard work. Producing one that is user friendly is much harder. Having it work on p processors, p=1,2,..., is harder still. There are many issues that have to be resolved: data structures, recognizing and taking advantage of local structure in a grid, domain substructuring, dynamic load balancing, grid management, error estimation, parallel I/O, and (lastly) interacting with a real simulation code.

Luckily Sandia already has a number of the pieces that are needed (e.g., the Zoltan package for dynamic load balancing). What we are producing is not necessarily a standalone code, but one that can be used by many different software packages and applications inside of Sandia.

At the suggestion of Jaideep Ray, we have used the software package Grace to do the grid generation and load balancing. The code is very effective when the number of elements in a grid patch is a power of two. Unfortunately, the code runs poorly when this assumption is not met. Our primary test examples do not fit this assumption. Portions of Grace may have to be replaced, either with code from a Sandia package or replacement code that we write. Using an existing Sandia code is the first choice.

During the four weeks that Craig Douglas spends at Sandia, we expect to start with running codes for both 2D and 3D parallel PDE solvers for adaptively refined grids. We hope to integrate the codes into Jaideep Ray's combustion code and gain experience from how effective it is. If possible, we would like to integrate the code into Ray Tuminaro's ML package.

One of the possible outcomes of the four weeks will be a new version of Jonathan Hu's academic code from his graduate school career. This is an unstructured and quasi-unstructured grid cache aware multilevel solver for two and three dimensions. The original code works with either scalar or coupled elliptic PDEs. The academic code requires Hu in order to set it up, which is not optimal. We will look into producing a new code in C and C++ that takes advantage of the work that is ongoing.

Title: Multiscale-Multiphysics Computational Framework

PI: Professor Jacob Fish, Rensselaer Polytechnic Institute

Investigator: Haim Waisman, Ph.D student
Zhijie Xu, Ph.D student
Peng Cheng. Ph.D student

Dates: May 20, 2002 – August 20, 2002

CSRI POC: John Aidun

Project Summary:

The primary goal of the proposed summer program is to initiate a long-term collaboration between multiscale modeling and simulation group lead by Professor Fish from Rensselaer and ongoing research efforts in this area at Sandia National Laboratories. Research efforts are proposed in three sub-areas of multiscale modeling and simulation:

1. A unified multiscale modeling and simulation approach at the nanoscale. This research effort will be lead by Zhijie Xu. As a part of this effort Mr. Xu will be interacting with Computational Materials & Molecular Biology division lead by Dr. John Aidun.

2. Multiple physical processes interacting at multiple spatial and temporal scales. This effort will be lead by Peng Cheng. The potential collaborators at Sandia on this subject include: Allen Robinson, Josh Robbins, Peter Schultz, Arne Gullerud, Jim Stewart, Mike Wong, Rebecca Brannon and Martin Heinstein.

3. Multilevel solvers: (i) generic aggregation based solvers for large scale systems and (ii) special purpose multilevel solvers for heterogeneous dynamic systems. Haim Waisman will be responsible for the development of these solvers. Both John Shadid and Mark Christon expressed an interest during my visit last year. Also, Sam Key might be interested, but I did not have a chance to meet with him during my last visit.

The final responsibility for the technology development and transition to Sandia National Laboratories will be the domain of Professor Fish. Dr. Fish will arrive with his PhD students on May 20, 2002 and will work with appropriate Sandia researchers and his students to ensure that appropriate interactions are put in place.

Task 1: A unified multiscale modeling approach at the nanoscale

Since no publication has been published on the subject matter I will describe this task in more detail. Zhijie Xu will be working on this task. The proposed multiscale modeling approach at the nanoscale is based on the following three guiding principles: generality, reliability and computational efficiency. We intend to develop a mathematical framework, including the use of both existing and new methods, that will be able to link a wide variety of continuum-to-continuum, continuum-to-discrete and discrete-to-discrete models. The generality of the proposed methodology will be accomplished by generalizing the continuum-based mathematical homogenization theory to account for discrete systems. Mathematical homogenization theory provides a rich mathematical foundation for bridging multiple spatial and temporal scales for continuum systems. Extensions of this theory to be developed as a part of the proposed work will provide a unified mathematical framework bridging not only multiple continuum scales in space and time, but also multiple continuum and discrete scales.

Our approach described below is quite general and can be directly applied to a great variety of complex multiscale problems. To be more specific, we describe our method for the case dealing with atomic

displacements, u_i. Applications to other problems, for example, involving electronic states, will be fully developed during this program.

Multiple scale asymptotic expansion in space (x, y, z, χ) and time (t, τ, η) serves as a starting point of the formulation:

$$u_i = u_i^o(x,y,z,t) + \varepsilon u_i^1(x,y,z,\chi,t,\tau,\eta) + O(\varepsilon^2) \qquad \varepsilon << 1 \qquad (1)$$

where x is the macro spatial scale (reactor, robotic arm, aircraft structural element), $y = x/\varepsilon$ the meso-spatial scale (thin film, quantum well), $z = y/\varepsilon$ the micro-spatial scale (grain, quantum dots), $\chi = z/\varepsilon$ the nano-spatial scale (atoms, nanotubes, nanocrystallites), t the usual macro-chronological time scale, $\tau = t/\varepsilon$ the fast nanochronological time scale, and $\eta = \varepsilon t$ the slow time scale used to eliminate the secularity of the asymptotic expansion.

Asymptotic analysis at multiple continuum spatial scales is well understood. Project PIs have an extensive experience on the subject matter. The use of slow time scale, $\eta = \varepsilon t$, gives rise to nonlocal homogenized equations [1, 2, 3]. For simplicity of presentation, we will focus on two temporal and two spatial scales, which includes: two continuum scales, micro-spatial (z) and micro-chronological (t), and two discrete scales, nano-patial (χ) and nanochronological (τ). The corresponding asymptotic expansion is given by

$$u_i = u_i^o(z,t) + \varepsilon H_{ijk}(\chi,\tau)\, u^o_{(j,k)}(z,t) + O(\varepsilon^2) \quad (2)$$

where $u^o_{(j,k)}$ denotes the symmetric gradient of u^o_j; The first term in (2) represents the microscopic solution, whereas the second term is the contribution from the nanoscale given in terms of the atomistic concentration factor, $H_{ijk}(\chi,\tau)$, and the gradients of the microscopic solution.

We now outline the construction of the mathematical model for one of the members in the hierarchical sequence. The model under consideration assumes that a periodic Atomistic Representative Unit Volume (ARVE) exist, in which the dynamics of each individual atom is explicitly represented. We refer to this model as the Periodic Molecular Dynamics (PMD) model. Discussion of the derivation of the mathematical models for other members in the hierarchical sequence is left to Section 2.2. The displacement field of each atom in ARVE is approximated as:

$$\boldsymbol{u}_i = \boldsymbol{u}_i^o(z,t) + \varepsilon \boldsymbol{H}_{ijk}(\tau)\, u^o_{(j,k)}(z,t) \qquad (3)$$

where $\boldsymbol{u}, \boldsymbol{u}^o, \boldsymbol{H} \in \boldsymbol{R}^n$ and n is the number of atoms in ARVE. The velocities of vibrating atoms directly follow from the chain rule, $\dot{\boldsymbol{u}}_i = \boldsymbol{u}_{i,t} + (1/\varepsilon)\boldsymbol{u}_{i,\tau}$. Having defined the positions and velocities of atoms one can compute the Lagrangian, L , of the discrete system:

$$L \equiv \frac{1}{|\theta|}\int_{\Omega}\int_{\theta}(T^o - \pi^o)\, d\theta' d\Omega' \quad (4)$$

where $|\theta|$ is the volume of ARVE; Ω the domain of the continuum system; $T^o(\boldsymbol{u}^o, \boldsymbol{H})$ and $\pi^o(\boldsymbol{u}^o, \boldsymbol{H})$ are the leading order kinetic and potential energies of all the atoms in the ARVE, respectively. The mathematical model is then obtained from Hamilton's principle with multiple temporal and spatial scales

$$\int_{t_1}^{t_2} \frac{1}{\tau_1 - \tau_2} \int_{\tau_1}^{\tau_2} \delta(L + F)\, d\tau\, dt = 0 \qquad (5)$$

where δ represents the variation and F is the contribution from the external source, and $\tau_1 - \tau_2$ is the period of the fast time variable in the ARVE. The solution of (5) gives the system of ordinary differential equations for the atomistic concentration factors $\boldsymbol{H}_{ijk}$

$$Discrete\ prob.: \frac{\partial \pi^o}{\partial \boldsymbol{H}_{ijk}} + \frac{\rho}{n} u^o_{(j,k)} u^o_{(m,n)} \boldsymbol{H}_{imn,\tau\tau} = 0 \qquad (6)$$

and partial differential equations for the continuum scale solution $u^o(z)$

$$Continuum\ prob.: \int_\Omega \delta u^o_i \rho u^o_{i,tt} d\Omega' - \int_\Omega \delta u^o_{(i,j)} \left(\frac{\partial \tilde{\pi}^o}{\partial u^o_{(i,j)}} - \rho \Psi_{ijmn} u^o_{(m,n)}\right) d\Omega' + \delta f^{ext} = 0 \qquad (7)$$

δf^{ext} is the variation of the external source, ρ is the density, and

$$\tilde{\pi}^o = \frac{1}{\tau_1 - \tau_2} \int_{\tau_1}^{\tau_2} \pi^o d\tau \qquad \Psi_{jkmn} = \frac{1}{n(\tau_1 - \tau_2)} \int_{\tau_1}^{\tau_2} \boldsymbol{H}^T_{ijk} \boldsymbol{H}_{imn} d\tau \qquad (8)$$

It is important to note that the discrete and continuum problems are fully coupled. The discrete problem has to be solved at each evaluation point of the discretized continuum problem. The coefficients of the continuum problem depend on the solution of the discrete problem. Continuum problems will solved using finite element.

Task 2: Multiple physical processes interacting at multiple spatial and temporal scales.

As an initial effort of this task Peng Cheng will be assisting Sandia researchers in implementing the classical mathematical homogenization procedures into one of the Sandia codes. This will involve a development of the library of the Representative Volume Elements including implementation of periodic boundary conditions.

With the completion of this task Peng will focus on development of temporal homogenization procedures. These capabilities will allow for linking distinct temporal scales such as slow diffusion and fatigue processes with fast thermo-mechanical and wave propagation processes.

There are numerous articles dealing with multiple spatial scales and many important technologies have been developed to bridge between the discrete and continuum scale and amongst the continuum spatial scales. It is surprising that the literature and the technologies for bridging multiple temporal scales are practically nonexistent. Multiple temporal scales procedures will be developed for three cases:

i. Multiple temporal scales within a single physical process on a single spatial scale. In many engineering problems multiple temporal scales arise in a single physical process. For example, slow degradation of materials properties due to creep, relaxation and fatigue, subjected to rapidly oscillatory loading exhibit multiple temporal scales. This category of problems possess an intrinsic slow time scale, which may significantly differ from the frequency of external input. Mr. Cheng was involved in development of temporal homogenization theory by which the initial-boundary value problem has been decomposed into: (i) the global initial-boundary value problem with smooth loading for the entire loading history, and (ii) the local initial-boundary value problem with the remaining (oscillatory) portion of loading for a single load period in selected region(s) of the time domain. Large time increments can be used for the global problem, while the integration of the local initial-

boundary value problem requires a significantly smaller time step. This approach allowed for the first time a rigorous modeling of fatigue and life time predictions of aerospace components, which are not based on S-N curves or Goodman-type diagrams. Currently Peng is involved in extending the theory to slow oxidation processes.

ii. Coupling of multiple physical processes. Different physical processes, such as mechanical, thermal, diffusion, and chemical reaction, may evolve along different time frames. Interaction between multiple physical processes requires consideration of relevant time frames within a single reference time coordinate. Environmental effects, such as diffusion of oxygen and eventual corrosion of structures is a slow process, whereas the deformation and in particular heat transfer are fast processes. Multiple scale homogenization approach is space and time will allow us to accurately resolve each physical process and to transfer appropriate information between the scales.

iii. Coupling of spatial and temporal scales. Spatial heterogeneities may cause dispersion of high frequency waves traveling in heterogeneous media. The time frame corresponding to the successive reflection and refraction of waves between the interfaces in microstructure could be significantly different from the time frame of the macroscopic wave motion. In recent work Fish, Chen and Nagai introduced multiple slow temporal scales to alleviate the problem of secularity caused by high order terms in the asymptotic analysis of wave propagation in heterogeneous solids and established a nonlocal continuum approach to capture dispersion effects. This is an ongoing effort at Rensselaer lead by Dr. Wen Chen partially supported by Sandia. This topic will not be part of the proposed summer program at Sandia.

Task 3: Multilevel Solver for Heterogeneous Medium

The key goal is to develop a scalable aggregation based solver for solving transient problems in heterogeneous media. Recently, we have developed an aggregation based method for two-phase heterogeneous medium whose rate of convergence is governed by:

$$\left\| {}^{i}\boldsymbol{e} \right\|_2 = \gamma \left\| {}^{i-1}\boldsymbol{e} \right\|_2 \qquad \gamma = \left[1 + k + 1/k + 2m^2 \left(\Delta t^{cr} / \Delta t \right)^2 \right]^{-1} \qquad (9)$$

where $\left\| {}^{i}\boldsymbol{e} \right\|_2$, $\left\| {}^{i-1}\boldsymbol{e} \right\|_2$ denote the norms of error in cycles i and i-1; k is the ratio of stiffnesses of the microconstituents. Δt^{cr}, Δt are the critical time step required for explicit computations and the implicit time integration step, respectively; m is the number of quasi-uniformly distributed nodes in the representative volume element. It can be seen that when a significant mismatch of stiffnesses exist the method converges very fast. Inertia effects may further speed up the convergence. Since discrete systems possess similar mismatch of properties (e.g., stiffness of polymer chains in axial direction is much higher than in the torsional direction, nanotubes order of magnitude stiffer that the polymer chains) we expect to obtain a convergence estimate similar to equation (9). Haim Waisman will be working with Sandia researchers on transitioning this important technology into one of Sandia's codes.

References

1. J. Fish and W. Chen, "Uniformly Valid Multiple Spatial-Temporal Scale Modeling for Wave Propagation in Heterogeneous Media," Mechanics of Composite Materials and Structures, Vol.8, pp.81-99, (2001).
2. J. Fish, W. Chen and G. Nagai, "Nonlocal dispersive model for wave propagation in heterogeneous media. Part 1: One-Dimensional Case," to appear in International Journal for Numerical Methods in Engineering, (2001).
3. J. Fish, W. Chen and G. Nagai, "Nonlocal dispersive model for wave propagation in heterogeneous media. Part 2: Multi-Dimensional Case," to appear in International Journal for Numerical Methods in Engineering, (2001).

4. J. Fish and W. Chen, "Higher-Order Homogenization of Initial/Boundary-Value Problem" accepted in Journal of Engineering Mechanics, (2001).
5. W.Chen and J.Fish, "A Dispersive Model For Wave Propagation in Periodic Heterogeneous Media Based on Homogenization With Multiple Spatial and Temporal Scales," Journal of Applied Mechanics, Vol.68, No.2, pp.153-161 (2001).
6. J.Fish and A.Wagiman, "Multiscale finite element method for heterogeneous medium," Vol. 12, pp. 1-17, Computational Mechanics: The International Journal, (1993).
7. V.Belsky, M.W. Beall, J. Fish, M.S. Shephard, S. Gomaa, "Computer-Aided Multiscale Modeling Tools for Composite Materials and Structures,"International Journal of Computing Systems in Engineering, Vol.6, No.3, pp.213-223, (1995).
8. J. Fish, K. Shek, M. Pandheeradi, and M.S. Shephard, "Computational Plasticity for Composite Structures Based on Mathematical Homogenization: Theory and Practice," Comp. Meth. Appl. Mech. Engng., Vol.148, No.1-2, pp.53-73 (1997)
9. J. Fish and K. L. Shek, "Computational Plasticity and Viscoplasticity for Composite Materials and Structures," Journal of Composites Part B, Vol.29, No.5, pp.613-619 (1998).
10. J. Fish and K. L. Shek, "Finite Deformation Plasticity of Composite Structures: Computational Models and Adaptive Strategies," Comp. Meth. Appl. Mech. Engng., Vol.172, No.1, pp.145-174 (1999).
11. J. Fish, Q. Yu and K. L. Shek, "Computational Damage Mechanics for Composite Materials Based on Mathematical Homogenization," International Journal for Numerical Methods in Engineering, Vol.45, No.11, pp.1657-1679 (1999).
12. J. Fish and Q.Yu, "Computational Mechanics of Fatigue and Life Predictions for Composite Materials and structures," to appear in Comp. Meth. Appl. Mech. Engng., (2001).
13. Q. Yu and J.Fish, "Multiple Temporal Scale Analysis of Viscoelastic and Viscoplastic Solids under Locally Periodic Loading, to appear in International Journal for Numerical Methods in Engineering, (2001).
14. J.Fish and Q.Yu, "Asymtotic Homogenization of Coupled Thermo-Viscoelastic Composites with Multiple Spatial and Temporal Scales," accepted in Composites Science and Technology: An International Journal, (2001).
15. J. Fish and W. Chen, "Multilevel solution techniques for wave propagation in heterogeneous solids," submitted to Journal of Engineering Mathematics, (2002).

Title: Optimization Algorithms

PI: Lisa Fleischer, Columbia University

Dates: June 3, 2000 – September 30, 2002

CSRI POC: Cindy Phillips

Project Summary:

Our current projects involve core research in the design and analysis of algorithms for discrete optimization problems. They have application to computer and infrastructure surety and logistics.

One of the projects I am involved with is a problem related to network security and scheduling. The question we are interested in is the following. We have a network of users, and for each user pair, a permissible communication level. Some users are not allowed to communicate, while others can do so with limited bandwidth. The communication levels are maintained by routers located at nodes of the network. When communication permissions change, it is necessary to reprogram the routers to enforce the new permissions levels. The question we seek to address is how to reprogram the routers in an efficient manner so that the new permissions levels are enforced as quickly as possible.

Even special cases of this problem are hard to solve exactly on large networks, so we investigate fast algorithms that and solutions that are close to optimal. This problem has connections to basic scheduling problems with precedence constraints, and we also plan on examining these connections.

A second problem we are considering is a network reinforcement problem. Given a network, we have choices for adding security on links of the network at a given cost. We seek the minimum cost set of choices so that an adversary without sufficiently high resources cannot disrupt essential connections in the network. This problem is also hard to solve on large networks. With Carr, Leung, and Phillips, we have an approximation algorithm for this problem. While our approximation guarantee improves the best known guarantee, it is still fairly weak. We seek to improve on this previous work in two directions. One, to improve the algorithm and the guarantee; and two, to further develop the techniques we have introduced in our original algorithm to attack related problems such as facility location. Facility location algorithms can be used to plan supply depot locations for military logistics. Spare parts management is an area of current need for the Department of Defense and military logistics is a current strategic focus area for Sandia Labs as a whole.

Title: Terascale Simulation-constrained Optimization

PIs: Omar Ghattas, Carnegie Mellon University
Larry Biegler, Carnegie Mellon University
Anthony Kearsley, Carnegie Mellon University

Dates: October 1, 2001 – September 30, 2002

CSRI POC: Bart van Bloemen Waanders

Project Summary:

PDE simulation-constrained optimization is a frontier problem in computational science and engineering. Often, the ultimate goal of simulation is an optimal design, optimal control, or parameter estimation problem, in which the PDE simulation is just an inner loop within the optimization iteration. Thus, the optimization (or "inverse") problem is significantly more difficult to solve than the simulation (or "forward") problem. When the simulation problem requires multi-gigaflop computing, as is often the case with complex 3D PDE systems, the optimization problem is of teraflop scale.

In contrast to the large body of research on parallel PDE simulation, very little work has been done on parallel PDE-constrained optimization. This is expected: it makes little sense to address the inverse problem until the forward problem is well understood. However, with the hardening and maturation of many large-scale PDE simulation codes, and with the recent rapid expansion of capability computing into the teraflop territory, we believe the time is right to mount a collaborative initiative with Sandia researchers aimed at creating an algorithmic framework for, and overcoming many of the central barriers to, very large scale simulation-constrained optimization. This is a difficult problem, but its solution offers a large payoff. Finally, since many simulation codes are being restructured and modularized to take advantage of highly parallel supercomputers, we feel that now is the time to catalyze the PDE solver community to make their codes optimization-ready. This window of opportunity may not reappear until the next major architectural change in high end computing.

The class of PDE-constrained optimization problems that we target have the following characteristics;

- The constraints usually result from complex discretization schemes and are often nonlinear.
- The state variables can number in the millions.
- The state constraints can number in the millions.
- The decision (i.e. design/control/inversion) variables range from hundreds to millions.
- The design/control/inversion inequality constraints range from the hundreds to millions.

Such problems are intractable with current optimization technology. Since the PDE constraints play such a central role in the optimization problem, it is clear that exploitation of their special structure is essential, both mathematically and computationally.

Title: Research in Finite Element Simulations

PI: Max D. Gunzburger, Florida State University

Dates: May 1, 2001 – December 31, 2004

CSRI POC: David Womble

Project Summary:

There are a variety of potential research projects and collaborations that I would very much to establish with CSRI and other Sandia Laboratory personnel. The list that follows is not an inclusive one. For example, I would be interested in getting involved, if CSRI and Sandia personnel feel that it would be beneficial, in ongoing projects such as finite element methods for MHD flows and finite element methods for compressible flows.

Stochastic PDE's and design under uncertainty – I believe it is fortuitous that john Red-Horse and I have a common interest in numberical methods for stochastic partial differential equations and their use for design under uncertainty, and that both of us are also in contact with engineers at the Air Force lab in Dayton about such problems. This is fortuitous because these areas of research are emerging as very important to numerous applications and at the same time remain very difficult to treat. I believe interacting with John will enable us to make significant progress in the theoretical and practical treatment of these problems. There is also the possibility of interaction with the Thermal Sciences Department on their response surface techniques for design under uncertainty; here, my work on centroidal Voronoi tessellations may have a beneficial role to play.

Least-squares finite element methods – Pavel Bochev (a CSRI visitor) and I are developing a theory of least-squares finite element methods which unifies all known variants of this class of methods under a single abstract setting. This effort could have considerable practical consequences in addition to the perhaps its more obvious theoretical importance. In particular, the unified theory is very likely to facilitate in the selection of the best method to apply in specific situations. Pavel and I are also in the midst of writing a book on least-squares finite element methods which will collect known theoretical, algorithmic, and implementation results about this emerging methodology.

Shape and value control of complex systems – CSRI already has some excellent collaborators, e.g., Matthias Heinkenschloss and Omar Ghattas, that certainly can provide much of what might be needed in the area of control for complex systems. However, I see a number of opportunities, perhaps a little down the road, for my interaction with CSRI and other Sandia Laboratory personnel on flow control problems and other control problems involving complex systems. For example, my expertise on shape control problems may come in handy when such efforts are needed in the future. Furthermore, it is my understanding that Martin Berggren will possibly be making an extended visit to CSRI. Martin and I have, on a number of occasions, discussed some work that we would like to do together; these projects, I believe, would be of benefit to Sandia and I would certainly welcome the opportunity of working with Martin.

Centroidal Voronoi tessellations – I have already committed myself to a project which will be funded separately by Rebecca Brasson; this is a very focused project in which I will produce, using centroidal Voronoi tessellations, an initial set of points for use in her group's particle-in-cell codes. I believe there are a number of other opportunities for interacting with CSRI and other Sandia Laboratory personnel on research projects based on my previous work on controidal Voronoi tessellations. This includes unstructured, adaptive mesh generation and meshless computing methods. I would very much like to continue the development of PDE computing for solving Sandia problems based on using centroidal Voronoi tessellations in concert with Sandia geometry and discretization software.

I will also participate in a number of other activities as part of my association with CSRI. A partial list of these include the following.

Students – I would very much like for some of my Ph.D. students to spend time at CSRI or other units of the Lab. Ideally, these students would find problems for their dissertations that are of interest to them, to Lab personnel, and to me. In this case, lab personnel would be actively involved in advising the students. The process of matching students to Lab problems and Lab personnel could begin with summer visits. Once firm connections are established, students would spend additional time at the lab. I have already encouraged two of my current students to apply for visits to CSRI for next summer.

Short courses or lecture series – I would be willing to deliver short courses or lecture series on any topic within my areas of expertise; I have already participated in such activities a number of times both in the Us and abroad. Notably, I have given short courses on subjects such as flow control, mathematical and computational aspects of superconductivity, finite element methods for fluid problems, and basic aspects of finite element methods.

Workshop organization – I would be willing to co-organize, with CSRI or other Sandia personnel, workshops on emerging topics in computational and applied mathematics and scientific computing. I have lots of expertise in the organization of both small and large workshops and conferences.

Title: Development of a Vorticity Method for the Solution of Nonlinear Viscous Shear Flows

PI: Marc S. Ingber, Department of Mechanical Engineering
University of New Mexico

Dates: March 19, 2002—March 18, 2003

CSRI POC: Steve Kempka

Project Summary:

The Principal Investigator proposes to develop a highly scalable, three dimensional computer code implementing the vorticity formulation and apply the code to application areas of interest for Sandia National Laboratories. The computational analysis will be based on an innovative massively-parallel vorticity formulation. The vorticity formulation is based on two kinematic steps and one kinetic step. Each step requires an essentially different strategy for parallelization. The bottleneck for the 2D vorticity code was the second kinematic step which required the evaluation of a velocity integral. Traditionally, the operation count for this step scales as $0(N^2)$ where N is the number of points at which the velocity integral is evaluated. Further, the memory requirements for this step also scale as $0(N^2)$. In the proposed research, a new multipole evaluation of the velocity integral will be implemented which will reduce the operation count and the memory requirements to 0(N). Because of the efficiencies of the proposed implementation and the natural advantages of vorticity formulations over primitive-variable formulations, it is anticipated that analyses of incompressible fluid flow problems with unprecedented fidelity will be enabled.

The proposed research will be leveraged by a tight collaboration with Sandia National Laboratories (SNL). SNL will provide expertise in the area of three-dimensional finite element modeling and high-performance computing. The major goal of this research is to develop a state-of-the-art computational tool to study three-dimensional nonlinear viscous fluid flows for discretizations containing on the order of 100,000,000 nodes. There are several research issues that will need to be resolved in order to develop a robust, efficient, and accurate code. Some of these research issues are discussed below.

Background: Vorticity methods for the solution of incompressible viscous flow problems have certain advantages over primitive-variable formulations including a reduction in the number of equations to be solved through the elimination of the pressure variable, identical satisfaction of the compressibility constraint and the continuity equation, an implicitly higher-order approximation of the velocity components, and, for exterior flows, a reduced computational domain. The Principal Investigator along with Dr. Kempka have previously developed a 2D vorticity code based on a Galerkin implementation of the generalized Helmholtz decomposition (GHD) for flow kinematics (Ingber and Kempka, 2001) as proof of concept. The 2D code demonstrated the effectiveness of the vorticity formulation for a variety of applications including interior, isothermal cavity flow, exterior flow about bluff bodies, and natural convection in an 8x1 enclosure at a critical Rayleigh number.

Proposed Research: The proposed research will entail the development of a scalable, parallel, three-dimensional vorticity code to solve the incompressible Navier-Stokes equations. The three-dimensional code will follow the general algorithm used in the two-dimensional code developed by the Principal Investigators (Ingber and Kempka, 2001; Brown and Ingber, 2001).

The numerical algorithm: The numerical algorithm for the proposed vorticity formulation is subdivided into a kinetic and two kinematic steps. The kinetics are governed by the three-dimensional vorticity equation given by

$$\frac{\partial \omega}{\partial t} + (u \cdot V)\omega = (\omega \cdot V)\, u + v\, V^2 \omega \qquad (1)$$

where u is the velocity field, $\omega = V \times u$ is the vorticity field, t is time, and v is the kinematic fluid viscosity. The kinematics are governed by the generalized Helmholtz decomposition (GHD) given by

$$\alpha(x\,[u(x) - \gamma(x) \times n(x)] = = \quad \frac{\int_{\Omega} \omega(y) \times r(\chi, y)}{r^3(x,y) \qquad d\Omega(y)}$$

$$+ \quad \int_{\Gamma} \frac{[(u(y) - \gamma(y) \times n(y)) \times n(y)] \times r(x, y)}{r^3(x, y) \qquad d\Gamma(y)}$$

$$- \quad \int_{\Gamma} \frac{[(u(y) \bullet n(y)]\, r(x, y)}{r^3(x, y) \qquad d\Gamma(y)} \qquad (2)$$

where Γ is the boundary of the domain Ω, γ represents the vortex sheet along the boundary Γ, and n is the unit normal vector on the boundary (pointing away from the fluid domain).

The numerical algorithm for solving the vorticity form of the Navier-Stokes equation is outlined below. First, the vortex sheet strengths are calculated using the tangential components of the Galerkin form of the GHD to determine the vortex sheet strengths. That is, in this kinematic step, Neumann boundary conditions are determined for the vorticity equation by relating the vortex sheet strengths γ to the vorticity flux on the boundary. Second, the internal velocities are evaluated from the regular form of the GHD. This kinematic step is used to linearize the vorticity equation by supplying the interior velocities to be used in the convective acceleration terms in the vorticity equation. Third, the vorticity field is transported by solving the Galerkin weak form of the vorticity equation. This final kinetic step will result in a kinematically incompatible velocity and vorticity field. To restore kinematic compatibility, the creation of new vorticity is determined by solving for newly formed vortex sheet strengths by returning to the first kinematic step for the new time step.

Parallelization: The vorticity algorithm is interesting from a parallelization point of view since each of the three steps require an essentially different parallelization strategy. In the following subsections, the proposed parallel strategy for each step is outlined. The overall parallel program paradigm will be based on the single program, multiple data (SPMD) model with MPI (message passing interface) for parallel communication.

Step 1, Parallelization of the Galerkin GHD: The Galerkin GHD is a boundary integral equation (BIE). The parallelization of this step will be based on a block-block (torus-wrap) data distribution (Ingber, Womble, Mondy, 1994).

In this approach, rectangular blocks of the coefficient matrix are assigned to processors in a wrapped fashion. It can be shown (Hendrickson and Womble, 1994) that this approach minimizes communication volume during the solution phase of the algorithm. Since the boundary element problem is small compared to the finite element problem of step 3, a direct solver will be used to provide a robust and reliable means of determining the vortex sheet strengths.

Step 2, Parallelization of the Velocity Integral: All velocity integral evaluations are independent, and hence, traditional approaches for these evaluations can be made embarrassingly parallel by simply assigning evaluation points (nodes) to processors. Nevertheless, for large problems, this step can consume well over 90% of the CPU and memory requirements. To reduce this bottleneck, an innovative fast multipole method (see Section 3.2) will be developed. Evaluation points can still be assigned to processors and the multipole evaluations are still independent. However, there are some additional complications required in the parallelization of the multipole method. In particular, the multipole scheme is based on a boxing strategy and evaluation of moment tensors. These evaluations will be performed in parallel on the processor array. The evaluation of the velocity integral is performed by scaling an inverted binary tree associated with the box list for a given evaluation point. These box lists and associated moment tensors can reside on different processors. The Principal Investigator has developed a methodology to limit the communication volume between processors for any given evaluation point to 4 point to point sends and receives (Mammoli and Ingber, 2000). This strategy will be implemented in the current research.

Step 3, Parallelization of the Galerkin FEM: The parallelization of the Galerkin FEM will be based on the total-summed-row approach (Shadid et al., 1994). In this approach, a row-wrap data distribution is assumed. Typically, this strategy is based on domain partitioning in which nodal data and connectivity for both the interior elements and border elements required to generate a complete row of the discretized matrix equation reside on a given processor. However, in the current implementation, since the nodal data is required for the calculation of the interior velocities in Step 1, this information is broadcast to all processors outside of the time loop after it is read in from the input file. Hence, no domain partitioning is necessary and rows in the matrix equation can simply be assigned evenly to processors. An iterative equation-solving package such as AZTEC will be used to solve the discrete set of equations. A variety of iterative methods such as GMRES, BiCGSTAB, CGS, QMR as well as some multilevel solvers will be evaluated for the vorticity formulation.

Time integrators and stability analysis: The original 2D version of the vorticity code used an explicit Euler time integration scheme. It has been shown that, even for explicit codes which are very fast per time step, overall computation time is often much longer because of stability constraints requiring very small time steps (Christon, 2001). A thorough stability analysis is proposed for the explicit implementation of the vorticity code. Further, a variety of implicit and semi-implicit time integrators will also be evaluated for the vorticity code. The relative CPU costs and accuracy of the various methods will be compared.

Improvements to the fast multipole method: Just recently, the Principal Investigator has implemented a fast multipole method (FMM) into the 2D vorticity code to evaluate the velocity integral. This second kinematic step of the 3 step algorithm is by far the most time consuming and also requires the most memory. The current implementation of the FMM is based on the Barnes-Hut (1986) algorithm. This implementation has a 0(N log N) scaling where N is the number of finite element nodes at which the velocity integral is evaluated. The memory requirements also scale as 0(N). It is possible with a second multipole center shift to further reduce the operation count to 0(N) (Greengard and Rokhlin, 1988). This second shift will be implemented as part of the proposed research.

Code optimization and scalability: Considerable effort will be exerted to improve the serial and parallel efficiency of the vorticity code. The code will be written in Fortran 90. All large arrays will be dynamically allocated, and if necessary, deallocated. There are several recursive algorithms in performing the FMM boxing strategies which most likely can be significantly improved. Studies will be performed to determine efficient block sizes with the goal of optimizing matrix-vector multiplications and reducing cache misses. The parallel algorithms will be thoroughly tested and evaluated to minimize communication volume and to maximize the scalability of the overall algorithm. The code will be profiled and bottlenecks will be identified. Algorithmic changes and coding alternatives will be implemented to alleviate the bottlenecks.

Input, output, and visualization: The last area of code development will deal with input, output, and visualization of the parallel vorticity formulation. The code will adhere to the EXODUS II input/output format to maintain compatibility with SNL grid generating and visualization tools. Further use will be made of the NEMESIS family of tools for domain decomposition and other parallel constructs. Finally, 2D, 3D, and immersive 3D visualization packages will be adopted to develop better insights into the various flow phenomena.

Application research: There are several application areas of interest to Sandia National Laboratories for the developed vorticity code. Although the vorticity formulation can be extended to compressible flows, the Principal Investigator believes the formulation is most advantageous for incompressible codes since the incompressibility constraint is identically satisfied. (This constraint causes a lot of problems for primitive-variable formulations.)

The first application area will be the safety analysis for nuclear weapons subjected to abnormal environments as a result of fire accidents. Metals and foams within the weapons can liquefy and result in an incompressible flow. It is of great concern to determine the temperature field, extent of melting, and flow of the component materials. In particular, the foam acts as both a mechanical and thermal barrier between components. The analysis will attempt to determine how long a particular weapon can be subjected to an

abnormal environment before safety issues become a concern. Additionally, the analysis of thermal environments experience by weapons during normal transport will also be analyzed where convection heat transfer is important.

The second application area will be the nonacoustic detection of submerged bodies from the wake signature caused by vorticity/free-surface interactions. Although this is not a current research focus of SNL, it would enable SNL to compete for future DoD funding.

The third application area will be numerous manufacturing flows. These flows include encapsulation (a current ASCI concern), injection molding, and braze furnace free convection processes where combined buoyant and forced convection is important.

Collaboration with Sandia National Laboratories: A tight collaboration is proposed with Sandia National Laboratories (SNL). SNL will provide expertise in the dynamical aspects of the vorticity simulation through their expertise in advanced scalable, 3D finite element methods. UNM will provide expertise in the kinematic aspects of the vorticity simulation through their expertise in advanced scalable, 3D integral equation methods. SNL and UNM have worked together previously to unravel several theoretical issues associated with vorticity formulations, and the proposed research will build upon this collaboration to develop a 3D vorticity formulation for efficient engineering analysis.

The key personnel at SNL will include Steve Kempka and Mark Christon. The key personnel at UNM will include Marc Ingber. A postdoctorate research assistant will be hired 3/4 time to work at the CSRI.

Title: Supporting MPI Collective Communication Operations with Application Bypass

PI: D. K. Panda, Ohio State University

Investigators: P. Wyckoff, , Ohio State University
P. Sadayappan, Ohio State University

Dates: July 1, 2002 – June 30, 2003

CSRI POC: Ron Brightwell

Project Summary:

Overview: For large-scale parallel systems supporting MPI, it is desirable that the MPI implementation ensures "progress" in order to achieve good performance and scalability [1]. Currently, collective communication operations in MPI are implemented by explicit send/recv calls by the processes. However, if a single node gets delayed (say an intermediate node of a broadcast, reduction, or barrier operation), the whole operation gets delayed. This leads to increased execution time for applications and limited scalability.

Modern interconnects are supporting new communication mechanisms such as remote memory operations (RDMA Read and RDMA write). Similarly, modern NICs are providing programmable interface and memory to support collective communication operations with minimal interaction from processors [2,3,4]. These advances allow communication operations to be implemented without explicit use of send/recv calls by the processes. This leads to the following two open challenges:

Can we implement MPI collective operations with application bypass by taking advantage of RDMA operations and NIC-level support?

How much performance benefit can be delivered to applications with this bypass property?

In this research, we propose to investigate both these challenges.

Proposed Work: The proposed work is divided into four broad directions: developing suitable mechanisms and semantics to implement collective communication with application bypass with host-level support, taking advantage of RDMA operations, taking advantage of NIC-level support, and evaluating the performance benefits at the application level. These four directions are briefly discussed below.

1) Mechanisms/Semantics to support application bypass with host-level support: Currently collective operations are implemented with explicit send/recv calls by the participating processes. In order to implement these operations with application bypass, new mechanisms and semantics are needed to be developed and evaluated. For example, for supporting broadcast at an intermediate node without the support of the application processes, do we need to have a dedicated thread or interrupt handler to take care of sending the messages to other nodes? How should this dedicated thread or interrupt handler be designed with minimal overhead? How the collective communication library needs to be interfaced with the dedicated thread or interrupt handler? What additional information a collective communication message should be carrying to invoke the appropriate thread or interrupt handler?

For supporting collective operations where an intermediate node has to supply data (such as barrier, reduction, or complete exchange), there is a need to define a clear semantic to support such operations at the application level. In current MPI implementations, all collective operations are blocking. In order to support collective operations with application bypass, we need to design and develop a new semantic. One possibility is to design a split-phase collective operation semantic which can support non-blocking/fuzzy collective operation. Under this semantic, a process can supply data to the dedicated thread/handler while

entering the collective operation, then leave for performing computation, and come back later to check for the completion of the collective operation.

We plan to study and analyze the above mechanisms and semantic to determine the right combination to support collective operations with application bypass and with minimal overhead. We will implement and evaluate these alternatives under MPICH/GM on Myrinet, MVICH/VIA over GigaNet cLAN, and MPICH/EMP over Gigabit Ethernet.

2) Taking advantage of RDMA operations: Modern interconnects support Remote DMA operations. Two kinds of RDMA operations are possible: RDMA Read and RDMA write. Myrinet and GigaNet cLAN implement RDMA write only. Compared to send/recv, RDMA write allows data to be written from a sender to a receiver without the exclusive involvement of the receiver. This one-sided operation is currently being exploited in many MPI implementations to support efficient point-to-point communication. We plan to study the benefit of such one-sided operation for supporting collective communication operations with application bypass.

The main issues here are as follows: How to define a region of pinned memory in a system to support collective operations with RDMA writes? How to organize this pinned area into different buffers and allocate these buffers for different collective operations in a unique manner across processes? How to integrate such pinning and buffer organization with the dedicated host-level thread/handler for collective operations so that the send operations at intermediate nodes can be replaced with appropriate RDMA write operations?

In our preliminary work along this direction with MVICH/VIA on cLAN, we have observed that barrier operation on an 8-node can be implemented with a 29% improvement with RDMA write operations compared to the traditional send/recv operations (31 microsec vs. 44 microsec). For larger systems, we hope to see much more benefit. In this evaluation, the benefits associated with RDMA write operation is only explored. The intermediate nodes still perform send/recv operations. We plan to extend this framework with appropriate pinning and buffering schemes for all collective operations and integrate it with appropriate mechanisms from step (1) to support collective operations with application bypass.

3) Taking advantage of NIC-level support: Modern NICs are becoming smarter with programmable processors, DMAs, and larger amount of memory. Potential benefits of such smarter NICs are being exploited to support fast point-to-point communication operations [6,7], collective communication operations [3,4], and remote atomic operations [2]. For collective operations, we have explored schemes where the NIC processors at intermediate nodes try to take care of most of the send/recv operations instead of the host nodes. For example, for a barrier operation, the host processors inform their respective NIC processors about reaching their barrier points. Then the NIC processors independently perform the required pair-wise exchange operations without the help of host processors. At the end of the exchange operations, the NIC processors inform the host processors to exit the barrier.

Such NIC-level support provides a natural mechanism to support a collective operation in an application bypass manner. For the time being, all our implementations still support blocking collective operations. We plan to extend this framework to support non-blocking collective operations. Also, on current generation NICs, NIC processors do not have sophisticated floating point units. Thus, it is very difficult to design fast implementation of reduction operation with complete NIC-level support. We plan to use a hybrid scheme where the NIC processors can perform the required communication operations and gather the data to some pinned memory and then the dedicated thread/handler at the host can perform the required computation steps. Such a hybrid scheme will deliver good performance and scalability to support collective communication operations in an application bypass manner.

4) Evaluating performance benefits at the application level: The designs proposed above have multiple trade-offs, such as RDMA support vs. NIC-level support, dedicated thread vs. interrupt handler support at the host, different buffering and pinning schemes, and different algorithms for collective operations. While designing and developing these schemes we first plan to use performance evaluations with micro-benchmarks to select the stronger and effective ones. As the entire application bypass collective operation

framework is developed, we also plan to use in-depth application-level evaluations to study the overall benefits of this research.

At OSU/CIS, we have access to the standard NAS benchmarks as well as some parallel visualization applications. At Ohio Supercomputer Center (OSC), we have access to many production-level application codes. As the schemes for application bypass collective operations get developed, we will compare their benefits with the existing (conventional) schemes for a range of applications. Such evaluations will demonstrate the benefits of application bypass collective operations and provide guidelines for designing next generation parallel systems and applications with high performance and scalability.

Milestones: The term of the proposed work will be one year starting July 1, 2002. The following list of milestones represents the approximate completion date of the research directions as described above.

Sept. 2002 - Design and Development of separate thread/interrupt handler to support application bypass and their impact on performance

Design, development, and evaluation of non-blocking semantics for collective operations to support application bypass

Dec. 2002 - Design of suitable pinning and buffering schemes to support collective operations with RDMA operations

Developing algorithms for frequently used collective communication operations (barrier, broadcast, and reduction) with RDMA support

Integrating RDMA support with separate thread/interrupt handler at the host to support application bypass and studying the performance benefits

Mar. 2003 - Design of NIC-based support for collective operations

Integrating the schemes with separate thread/interrupt handler at the host to support application bypass and studying the performance benefits

June 2003 - Performance evaluation of both schemes for a range of applications and systems

Deriving guidelines about which schemes work best to support collective operations with application bypass

Deliverables: The following items will be delivered during the course of the contract:

Different mechanisms (thread/interrupt handler, RDMA support, NIC-based support), semantics (non-blocking collectives), and their implementations on contemporary communication subsystems (Myrinet/GM, GigaNet/VIA, and Gigabit Ethernet/EMP) to support collective operations with application bypass.

Performance and scalability analysis of the new scheme for a range of applications.

References:

[1] Portals 3.0: Protocol Building Blocks for Low Overhead Communication, Ron Brightwell, Bill Lawry, Arthur B. Maccabe, Rolf Riesen, CAC '02 Workshop, April 2002.

[2] D. Buntinas, D.K. Panda, and W. Gropp, NIC-Based Atomic Operations on Myrinet/GM, SAN-1 Workshop, held in conjunction with High Performance Computer Architecture (HPCA) Conference, Feb 2002.

[3] D. Buntinas, D. K. Panda and P. Sadayappan, Fast NIC-Based Barrier over Myrinet/GM, Int'l Parallel and Distributed Processing Symposium (IPDPS), April 2001.

[4] D. Buntinas, D. K. Panda, J. Duato, and P. Sadayappan, Broadcast/Multicast over Myrinet using NIC-Assisted Multidestination Messages, Fourth Int'l Workshop on Communication, Architecture, and Applications for Network-Based Parallel Computing (CANPC '00), Jan 2000, pp. 115-129.

[5] R. Gupta, The Fuzzy Barrier: A Mechanism for the High Speed Synchronization of Processors, Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems, 1989, pp. 54-63.

[6] P. Shivam, P. Wyckoff, and D. K. Panda, EMP: Zero-copy OS-bypass NIC-driven Gigabit Ethernet Message Passing, Supercomputing (SC '01), November 2001.

[7] P. Shivam, P. Wyckoff and D.K. Panda, Can User Level Protocols Take Advantage of Multi-CPU NICs?, Int'l Parallel and Distributed Processing Symposium (IPDPS '02), April 2002.

Title: Large Scale Eigenvalue Methods and Model Reduction of Second Order Systems

PI: D.C. Sorensen, Rice University

Dates: August 1, 2002 – September 30, 2003

CSRI POC: David Womble

Project Summary:

Large Scale Eigenvalue Problems We shall continue to develop techniques for improving the performance of ARPACK. This software for large eigenvalue problems is in wide use at Sandia and is based upon our implicitly restarted Arnoldi method. We hope to develop pre-conditioning techniques appropriate for stability and bifurcation analysis of dynamical systems. These will be closely related to the use of iterative methods for solving the equations required to implement a Cayley transformation. However, they will construct a fixed (preconditioned) polynomial operator as an approximation to the shift-invert operator.

Model Reduction of Second Order Dynamical Systems Direct numerical simulation of dynamical systems has been an extremely successful means for studying complex physical phenomena. However, as more detail is included, the dimensionality of such simulations may increase to unmanageable levels of storage and computational requirements. One approach to overcoming this is through model reduction. The goal is to produce a low dimensional system that has the same response characteristics as the original system with far less storage requirements and much lower evaluation time. The resulting reduced model might be used to replace the original system as a component in a larger simulation or it might be used to develop a low dimensional controller suitable for real time applications.

In the past year, we have made considerable progress on the fundamental model reduction of systems of the form

$$\dot{x} = Ax + Bu,\ y = Cx,$$

where A, B, C are real $n \times n$, $n \times m$ and $p \times n$matrices, while u, y, x are vector valued functions of time. Large systems of this form arise in many applications, for example in circuit simulation and in the simulation of PDEs. We have developed balanced model reduction techniques for large scale systems through low rank approximation of certain system Grammians. These techniques are matrix-free in the same sense as Krylov methods for eigenvalue computation and solution of linear systems. Balanced reduction is an excellent candidate for the development of robust and widely applicable software because of the existence of *a-priori* error bounds and the preservation of important system properties. We intend to extend these results and also develop new techniques for model reduction of second order systems $M\ddot{x} + G\dot{x} + Kx = Bu,\ y = P\dot{x} + Qx$.

Such systems are far more challenging to work with but they have many more applications. Moreover, even though a second order system can be reduced to a first order system through standard techniques, a reduced model obtained from the first order formulation is usually not valid when converted back to the second order setting. We hope to develop balanced model reduction techniques that work directly with the second order system.

Title: PIM Prototype Testbed

PI: Dr. Thomas Sterling, Caltech

Investigator: Maciej Brodowicz

Dates: 7/31/2001 – 7/31/2003

CSRI POC: Erik DeBenedictis or Rolf Riesen

Project Summary: An opportunity exists to extend the impact and implications of research being performed by Caltech and Sandia National Laboratory under the sponsorship of CSRI related to advanced architecture for high performance computing. A tool developed, in part, through CSRI support to enable the rapid prototyping of processor in memory architectures has been implemented and tested. This unique testbed integrates a combination of four FPGA chips and two banks of SRAMs with 256 bit wide row to emulate the structure of a dual node PIM chip. In addition, a dozen firewire interfaces provide the transport layer to test various external communication interface protocols. An additional interface provides the means to rapidly download different logic designs to test and evaluate alternative PIM processor architectures. Along with Caltech, the University of Notre Dame has been pushing the state of the art in advanced PIM architectures. Together, they are developing the combined concepts of multithreading, virtual memory address translation, and parcel message driven computation in a single PIM processor architecture. Their recent architecture, PIM-lite, has been simulated and a full layout of the design has been created but their has not been the means to test this design. While they eventually wish to implement their PIM-lite design in silicon, this does not provide a flexible medium for iteration and optimization. The FPGA testbed does and the University of Notre Dame, under the direction of Professor Peter Kogge and his colleague Jay Brockman had stated their desire for one of the Caltech prototype testbeds. With a number of graduate and undergraduate students at Notre Dame to develop software tools and environment, the consequence of establishing a team of SNL, Caltech, and UND will be that each institution will provide important capabilities to the overall project. It is proposed that SNL CSRI sponsor the fabrication of a new prototype testbed at Caltech/JPL to be delivered to UND in the first Quarter of FY03. The software developed by Notre Dame will be shared with SNL and Caltech to accelerate progress on their research project. In addition, this new effort will establish intellectual ties between Notre Dame and SNL CSRI.

Title: Pattern Search for Nonlinear Optimization

PI: Virginia Torczon, College of William and Mary
Robert Michael Lewis, College of William and Mary

Dates: March 6, 2001 – August 31, 2002

CSRI POC: Tammy Kolda

Project Summary:

This is a proposal to continue *the* development of pattern search algorithms for nonlinear optimization problems of the form:

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } x \in S \subseteq \Re^n, \end{aligned}$$

where f: $\Re^n \rightarrow \Re$.

This work will proceed along two fronts. First, the co-PI's will meet with W. E. Hart, Department of Applied Mathematics, Sandia National Laboratories, Albuquerque, and T. G. Kolda, Computational Science and Mathematics Research Department, Sandia National Laboratories, Livermore, to discuss efforts to implement pattern search methods for problems with constraints. This investigation will pursue the implementation of algorithms that have already been analyzed by the co-PI's [4, 6, 5] All four investigators will meet at Sandia National Laboratories, Albuquerque, March 7-9, 2001 to discuss how to coordinate this undertaking.

Second, Torczon will continue work with Kolda to complete the analysis of Asynchronous Parallel Pattern Search (APPS). The results of our prior collaboration, which included P. D. Hough, Computational Science and Mathematics Research Department, Sandia National Laboratories, Livermore, to develop APPS is reported in [2].

Kolda and Torczon will show that in the unconstrained case (i.e., $S \equiv \Re^n$), APPS converges (even in the case of faults) under the same assumptions as pattern search [3, 1]. The result will give analytic rigor to the extant implementation of APPS, which is now being distributed (see http://csmr.ca.sandia.gov/projects/apps.html). Further, the analysis may also suggest additional modifications or improvements that can be made to either expand the applicability or increase the flexibility of APPS. The investigators already have a preliminary draft of the paper; further work will involve verifying the results and preparing the manuscript for submission to *SIAM Journal on Optimization.* Torczon is tentatively scheduled to visit the Computational Science and Mathematics Research Department, Sandia National Laboratories, Livermore, May 14-25, 2001.

[1] Elizabeth D. Dolan, Robert Michael Lewis, and Virginia J. Torczon. On the local convergence properties of pattern search. Technical Report 2000-36, ICASE, September 2000. In revision for *SIAM Journal on Optimization.*

[2] P. D. Hough, T. G. Kolda, and V. J. Torczon. Asynchronous parallel pattern search for nonlinear optimization. *SIAM Journal on Scientific Computing,* to appear. Originally released as Technical Report SAND2000-8213, Sandia National Laboratories, Livermore, California, January 2000.

[3] Robert Michael Lewis and Virginia Torczon. Rank ordering and positive bases in pattern search algorithms. Technical Report 96-71, ICASE, Mail Stop 132C, NASA Langley Research Center, Hampton, Virginia 236189-2199, 1996.

[4] Robert Michael Lewis and Virginia Torczon. Pattern search algorithms for bound constrained minimization. *SIAM Journal on Optimization,* 9(4):1082-1099, 1999.

[5] Robert Michael Lewis and Virginia J. Torczon. A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds. Technical Report 98-31, Institute for Computer Applications in Science and Engineering, Mail Stop 132C, NASA Langley Research Center, Hampton, VA 23681-2199, July 1998. In review for *SIAM Journal on Optimization.*

[6] Robert Michael Lewis and Virginia J. Torczon. Pattern search methods for linearly constrained minimization. *SIAM Journal on Optimization,* 10(3):917-941, 2000.

Chapter 3. Sabbaticals

The following university faculty did all or part of a sabbatical stay at Sandia National Laboratories during calendar year 2001.

Dr. Martin Berggren,
Aeronautical Research Institute of Sweden, and
Uppsala University

Dr. Alex Pothen
Old Dominion University, and
ICASE, NASA Langley Research Center

Dr. Robert Preis
University of Paderborn, Germany

Title: Numerical Methods For Partial Differential Equations with Emphasis on the Equations of Fluid Mechanics

PI: Martin Berggren,
Aeronautical Research Institute of Sweden, and
Uppsala University

Dates: February 1, 2002 – June 30, 2003

Project Summary:

My research interests are centered around numerical methods for partial differential equations with emphasis on the equations of fluid mechanics. This includes modeling and discretization issues as well as algorithms for solving the resulting algebraic problem and software aspects.

Numerical methods for fluid-mechanics problem have reached a certain maturity (even if there is still a long way to go to achieve good prediction of, say, massively separated flow at high Reynolds numbers). Therefore, I believe that many of the scientific breakthroughs in the future will appear when the numerical solution is part of a larger interdisciplinary effort. This is a reason why my specific research activities have been directed to such "compound" problems; fluid-structure interaction, analysis of flow instabilities, flow control, and aerodynamic shape optimization. The complexity of interdisciplinary problems makes it essential to interact closely with experts in different disciplines. A visit to the CSRI would allow me a unique opportunity to interact with scientists associated with Sandia. Collaboration involving some of the following Sandia staff members would in particular be of interest: Mike Eldred, Richard Lehoucq, Walter Rutledge, Kambiz Salari, Andy Salinger, and Bart van Bloemen Waanders. I have also been informed that Professor Max Gunzburger is making arrangements for regular visits. My research interests intersects those of Gunzburger, so it would certainly be of interest to cooperate. I also hope and believe that my specific expertise will be of value for Sandia.

Shape Optimization

Computational Fluid Dynamics (CFD) is increasingly used for analysis purposes, for instance in the design stage of components in which fluid-mechanical properties are important, such as for vehicles, turbines, or pumps. As the turnaround time for an analysis cycle decreases with increasing hard- and software performance, it is natural to aim for using CFD in an automatic search of the best designs. This is the purpose of shape optimization. Structural optimization methods, such as minimizing weight or maximizing stiffness under prescribed loads, appear to be more developed than shape optimization in a fluid-dynamics context; there are even several commercial packages, from Altair engineering for instance, performing shape or topology optimization of structures. Although shape optimization in a fluid-dynamics context is a less mature field, some substantial results already exists, perhaps best exemplified by the work of Jameson and coworkers [11]. The industrial interest for the technique can be exemplified by the EU-sponsored project *Aeroshape*, a three-year project devoted to aerodynamic shape optimization that started January 2000, in which most of the leading European aeronautical industries and research institutes participate (Parts of my work at FFA is within this project.)

One important, fundamental issue is well-posedness. Oscillatory shapes often appear in shape optimization computations, as discussed by Pironneau [12], for instance. This can have several reasons, physical as well as numerical. For instance, it is known that so-called riblets, small grooves in the streamwise direction on the surface of embedded solids, may decrease the viscous drag [2]. Also, the sensitivity of the objective function to oscillatory changes in the shape may be low and this may cause numerical instabilities for the discrete problems, similarly to what happens in mixed-methods for incompressible flow (checker-boarding, inf-sup, LBB conditions, etc.) it is thus important to find a sound formulation of the problem, covering what is wanted out of the optimization (for instance, do we want those riblets or not?). The formulation should yield a well-posed mathematical problem, not prone to numerical instabilities, but one that still covers a large span of designs; the easy way of achieving well-posedness is otherwise to shrink the design

space sufficiently. Some results in this direction can be found in the literature; two recent contributions are by Bedivan [4] and Gunzburger & Kim [9], but many issues still have to be addressed.

Many aspects of these more fundamental issues can be studied on a model-problem level, say finite-element approximations of the impressible Navier-Stokes equations in 2D. However, it is also important to consider some type of "real-world" analysis codes.

A numerical technique that is gaining in popularity in industrial-type codes for solving the Navier-Stokes equations is cell-vertex finite-volume discretizations in an edge-based formulation combined with unstructured meshes [3]. The use of unstructured meshes simplifies automatic mesh generation and adaptation for complex geometries. The edge-based formulation has the advantage that the implementation is basically unchanged when going from 2D to 3D, and it is also independent of the shape of the control volumes. This facilitates the use of hybrid meshes, where different control-volume shapes are used in the same mesh, typically so that regular hexahedrals with high aspect ratios are used to resolve the boundary layer close to solid boundaries, whereas tetrahedrals fill the area outside the boundary layer.

The applications in mind are of the final-design type. That is, at least a rough preliminary design of the component is assumed to be known. It is the purpose of the optimization to find the "best" design through moderate modifications of the preliminary design. It thus seems appropriate to use a local, gradient-based optimization technique combined with adjoint-based methods for the calculations of gradients, to allow for efficient gradient evaluation when the number of design variables is large and the objective-function evaluations are expensive. My own experience from similar applications strongly indicate that a fast convergence of the optimization requires in general extremely accurate gradient evaluations [5, 8, 10]. It would be interesting to investigate this approach in the context of edge-based unstructured finite-volume solvers, which particular emphasis on very accurate, "discrete" implementation of the adjoint equations. Practical implementation and testing in an industrial-type code should be a part of this.

To summarize, two possible projects could be the following:

- Investigations regarding problem formulations, well-posedness, approximation properties, and numerical algorithms. Studies on model problems not more complicated that incompressible Navier-Stokes equations at low Reynolds numbers in 2D.

- The use of gradient-based optimization for shape optimization using an industrial-style, edge-based finite-volume code.

Alternate Approaches to Linear Stability Analysis

Stability analysis for fluid flows is concerned with deciding if a particular flow is stable for infinitesimal or finite perturbations. Linearized for perturbations u around a laminar flow state U and performing eigenvalue analysis of an operator $A(U)$ appearing in an equation of evolution, such as $u_t + A(U)u = 0$, for the disturbance u shows whether it will be linearly *asymptotically* stable. For a long time it has been recognized that this analysis does not explain the transition to turbulence for certain flows, particularly shear flows such as pipe Poiseuille flow. When the laminar state U is strongly sheared, as in a jet, in the flow over a surface, or in a pipe, the operator $A(U)$ will be highly nonnormal and thus susceptible to significant transient effects, regardless whether the eigenvalues of $A(U)$ indicate stability or not. The importance of this mathematically well-known fact for so-called subcritical (or bypass) transition to turbulence has only surprisingly recently been recognized [6, 7, 13, 14]. In these articles, transient effects have been analysed using methods that is only applicable to paralled flow cases. Recently, methods very similar to the ones used for optimal-control problems and shape optimization problems have been applied to the study of transient effects [1]. This approach is quite general and applicable to transient studies of almost any flow. However, it has only been applied to the growth of steady disturbances in boundary layers. It would be interesting to study transients in more general situations using this approach.

Title: Combinatorial Methods in Scientific Computing

PI: Alex Pothen
Old Dominion University, and
ICASE, NASA Langley Research Center

Dates: September 10, 2001 – September 9, 2002

Project Summary:

I will pursue three topic during a sabbatical visit to Sandia National Laboratories and its Computer Science Research Institute. In addition to these issues I am interested in interacting with SNL scientists to learn of other problems in the broad areas of scientific computing and combinatorial algorithms and to collaborate with them in solving them.

1. Combinatorial algorithms for preconditioning

Developing effective preconditioners for solving sparse systems of equations on Tera-op parallel computers is an important challenge. The problems in this field are primarily algebraic in nature and traditionally have been solved with the tools of linear algebra. However, recently, combinatorial techniques have been shown to lead to fresh insights and new algorithms for creating and computing.

David Hysom (my PhD student who is scheduled to graduate) and I have developed scalable, parallel algorithms for computing incomplete factor preconditioners for symmetric indefinite and unsymmetric problems. These algorithms have been shown to be scalable both experimentally (for several PDEs) and by analysis (for model problems); The PILU preconditioners also have the desirable property that the number of iterations of the preconditioned Krylov solver do not increase with the number of processors (subdomains).

The parallel algorithm has three major steps. In the first step, we create parallelism by partitioning the problem into subdomains, and then map each subdomain to a processor. In the second step, we preserve the parallelism within a subdomain by ordering the interior vertices in a subdomain before its boundary vertices. In the third step, we maximize the parallelism in computing the boundary nodes of all the subdomains by ordering the subdomains by means of a graph coloring.

We wish to extend this work in new directions and to interact with significant recent developments in computing preconditioners. One is the development of multilevel ILU preconditioners that have linear time complexity. A second is a significant extension of Vaidya's method for computing provably good preconditioners for symmetric problems via graph embedding (support theory-based preconditioners) by Erik Boman and Bruce Hendrickson at Sandia. Experimental work by Dror Chen and Sivan Toledo at Tel-Aviv shows that these preconditioners are superior to incomplete Cholesky (IC) preconditioners with comparable fill for 2-D problems. However, for 3-D problems, they are inferior to IC, and hence new ideas are needed to make them better. We will interact with these authors to understand the insights to preconditioning that support theory provides and then exploit this understanding for improving current preconditioners.

One challenge in this area is to extend the domain of applicability of the support theory results by decomposing a matrix as a sum of low-rank matrices, such that the support theory can be applied to each of the low-rank matrices in a straightforward manner. One theoretical goal we have is to use support theory to understand the nature of "fill" paths (paths in the graph of the initial coefficient matrix between two vertices whose interior vertices are numbered lower than the endpoints) in graph models of incomplete factorization. Hysom and I have proved an incomplete fill path theorem that characterizes level values of edges in terms of lengths of fill paths. We have also characterized relations between various rules for computing level values of fill (the "sum" rule and the "maximum" rule) and the structure of fill paths. We plan to apply support theory to incomplete Cholesky factorization to develop bounds on the condition number of the preconditioned matrix as a function of the level value. This will lead to improved

understanding of the convergence behavior of preconditioned Krylov space solvers. Another important goal is to study the role played by different orderings of the initial matrix (these affect the structure of fill paths) on convergence of the iterative solver. We will also study the influence of various partitioning objectives on parallel performance and convergence of incomplete factor preconditioners.

2. Parallel graph coloring algorithms for estimating Jacobians and Hessians in parallel

We will develop parallel algorithms and software for estimating Jacobians and Hessians. This is done via finite differencing (FD) or by automatic differentiation (AD). The problems of minimizing the number of (expensive) function evaluations needed to compute the Jacobian or Hessian can be formulated as graph coloring problems. Earlier work by Coleman, More, and others on this problem require several variant coloring problems to be solved: the problems for FD and AD are quite different while the problems (indeed even the graphs considered) are different for Jacobian estimation and Hessian estimation. Currently parallel software for graph coloring is not available and Lois McInnes, Paul Hovland and colleagues at Argonne (associated with the ASCI level -1 FLASH project at the University of Chicago), who are integrating parallel optimization software into PETSc, have expressed interest in our work.

Our first contribution to this problem is a unified perspective of the several variant problems that simplifies the development of parallel algorithms and software: We show that a *single* algorithm for a more general coloring problem, the distance-two coloring problem, can solve all of the variant problems that arise in optimization. (In a distance-two coloring all vertices within a distance two of a given vertex must receive different colors. The familiar graph coloring problem is a distance-one coloring. In a bicoloring both rows and columns are colored with different sets of colors.) For AD, Jacobian estimation involves a distance-two bicoloring; for FD, Jacobian estimation involves a distance-two coloring; and both work with the bipartite graph of the Jacobian. This saves the space to store the symmetric graph of the matrix $J^T J$ (the column or row intersection graph of the Jacobian J). Computing the graph of $J^T J$ in parallel is also cumbersome if J is distributed among the processors. Hessian, estimation involves the adjacency graph of the symmetric Hessian but the coloring needed is a path coloring a relaxation of the distance-two coloring.

As a first step we have designed a new parallel algorithm for distance two coloring for shared-memory multiprocessors that exhibits good speed-ups. The expected run time of this algorithm is $O(\Delta^2 n/p)$, where n is the number of nodes in the graph, p is the number of processors and is the maximum degree of a node This result holds when n is sufficiently large relative to p the expression becomes more complicated if this condition does not hold. (This is an extension of a parallel algorithm for distance-one coloring proposed by Gebremedhin and Manne at the University of Bergen. Gebremedhin is currently working with Pothen.) As a first step, we propose to develop this implementation further and provide a portable implementation that uses OpenMP. Then we will adapt this parallel implementation to solve the variant coloring problems in AD and FD.

The challenge here is to extend and modify the ideas in our shared-memory parallel algorithm for scalable distributed memory machines. The functionality of parallel graph coloring is important for parallel optimization, and if the coloring phase is relatively inexpensive compared to the numerical computations low speed-ups in the coloring algorithms would be tolerable. (Currently available algorithms cause a speed -down rather than speed -up when the number of processors increases.)

3. Problems in computational proteomics

The proteome of an organism is the set of all the proteins made by its cells. Current best estimates are that there are roughly 30 K to 40 K genes in the human genome; each gene is responsible for making multiple proteins (through the use of different reading frames, and by chemical modification). An important task is to characterize the proteomes of various organisms in order to understand the molecular mechanisms of life better. This is a fundamental goal of biology. For this goal to be achieved we need rapid, inexpensive, automated methods to identify proteins similar to that recently developed for genomics.

A number of companies are currently involved in developing such technologies to rapidly identify proteins. One such technology is Surface Enhanced Laser Dissociation Ionization (SELDI) mass spectrometry. This

technology has been developed by Ciphergen Inc., and is being used at Eastern Virginia Medical School (EVMS) for prostate and breast cancer diagnosis and classification.

We have been collaborating with molecular biologists and physicians at EVMS to develop algorithms for classifying prostate cancer benign prostate hyperplasia (BPH), and healthy prostate tissue from SELDI mass spectral data of mixtures of proteins isolated from serum.

We have applied statistical techniques of discriminant analysis to classify the data, approximated in a subspace of the leading singular vectors (principal components). The computational techniques involved are the computation of singular values and vectors of sparse matrices, and the solution of sparse least squares problems. Currently this technique looks promising, with error rates less than 2%. It is successful in identifying healthy prostate from BPH and cancer, while the distinction between BPH and cancer is currently not as good. We are continuing to develop this method.

Our goal is to more generally identify interesting research problems in computational proteomics. The first issue in April of *The Scientist* (www.thescientist.com) and the first issue of the new Journal *Proteomics* in 2001 [1] discuss some of these problems. The challenge is to study the biological literature to understand the major computational problems, formulate them as problems in computer science and computational mathematics, and then develop effective algorithms. I propose to give several lectures to the Discrete Algorithms Group to describe the computational problems in proteomics.

Title: Graph Partitioning Algorithms

PI: Dr. Robert Preis
University of Paderborn, Germany

Dates: October 1, 2001 – September 30, 2002

Project Summary:

There are several research topics I will work on at the CSRI.

- Analyses of the Multilevel Paradigm for Graph partitioning: The multilevel graph partitioning paradigm has been proven to be a very powerful approach to efficient graph-partitioning. Several algorithms have been developed for the single tasks of the paradigm. However, the quality of the multilevel process as a whole has only been proven on an experimental basis. First attempts to analyze the multilevel approach are very weak due to massive assumptions. The goal is to give an overall analyses of the multilevel approach with no or with only small assumptions.

- Graph partitioning objectives for parallel computing: The efficient use of a parallel computer system is a major motivation for the analyses and development of graph partitioning methods and tools. However, the basic graph partitioning problem re the computation and communication of the applications only to a certain extent. Several new approaches to address the multiple and complex objectives more precisely have been developed recently (also by researchers of the CSRI). Nevertheless, there are still many open questions in this area and a high demand for further analyses, development of advanced methods and more experimental evaluation and comparison between the approaches.

- Development of diffusion load balancing schemes: Several diffusion load balancing schemes have been developed in the past. A large class of them computes a balancing which is minimal in the l 2 norm. This leads to a small communication volume in most cost models. However, there is a lack of load balancing schemes that calculate balancing that are minimal in other norms (like e.g. l 1 norm) which are the optimization goal of certain communication systems. The goal is to develop new load balancing schemes and to compare them to existing schemes analytically as well as experimentally.

- Identification of almost invariant sets in chaotic dynamical systems: The analyses of dynamical systems is important in several areas like e. g. molecular dynamics or energy-efficient trajectories for spacecraft missions. Almost invariant sets of a dynamical systems are areas of the system that are almost disconnected from the rest of the system. The identification of almost invariant sets can be performed by deriving a transformation matrix of the dynamical system. This transformation matrix can be viewed as a weighted, directed graph and almost invariant sets are subsets with only a small number of external edges. Existing data clustering and graph partitioning methods can be used to determine the invariant sets. The goal is to firstly attack the identification of almost invariant sets with existing methods and tools and, secondly, to derive new methods for solving the problem of this application more precisely.

Chapter 4. Workshops

The following workshops were either hosted by the CSRI or partially supported by the CSRI during 2002.

Title: 4th Algorithm Engineering and Experiments
The ALENEX Conference

PI: Clifford Stein, Columbia University

Dates: January 4-5, 2002

CSRI POC: Cynthia Phillips

Conference Summary:

The annual ALENEX workshops provide the only U.S. forum for the presentation of original research in the implementation and experimental evaluation of algorithms and data structures. It is one of two premier conferences in this area, the other being the European-based Workshop on Algorithm Engineering.

Algorithm engineering and experimental algorithmics attempt to answer questions such as "How does one make specific algorithms run extremely fast in practice (e.g. what are the right data structures, how can one code to effectively use the memory hierarchy, etc)? What are the proper methods to use in computer experimentation so that one has confidence in the results of the experiments? The conference scope includes evaluation of algorithms for realistic environments and scenarios, that is, distinguishing the performance of candidate algorithms for particular data characteristics and side constraints. The conference also includes results that show synergy between mathematical and experimental analysis. This conference encourages the design and dissemination of test suites. Authors frequently make codes and test data available to the research community.

Alenex is colocated with the premier U.S. discrete algorithms conference (the ACM/SIAM Symposium on Discrete Algorithms). Thus many of the best algorithms designers and analyzers (theoretically and experimentally), both professors and students, will be in attendance.

RELATIONSHIP TO SANDIA'S RESEARCH:

Any Sandia project that uses sophisticated data structures and/or algorithms can benefit from the research presented at this conference. Starting from a careful previous study can save a lot of time testing candidate structures/algorithms. To date, many of the algorithms tested have been discrete optimization problems, such as those used for logistics (scheduling, transportation), network design and routing, etc. The Optimization and Uncertainty Estimation Department is doing research on discrete optimization applied to problems of national security such as logistics for stockpile stewardship, infrastructure and site surety, computational biology, and manufacturing design. These sorts of algorithms have also been used in scientific computing at Sandia (e.g. in domain decomposition, radiation transport codes, preconditioning, etc).

The Optimization and Uncertainty Estimation Department has an active interest in experimental algorithmics. In particular, when adapting theoretically-strong algorithms to applications environments, e.g. where additional DOE regulations must be enforced, we must use experimental analysis to provide confidence in our solutions. This is also necessary for benchmarking heuristic codes developed elsewhere at Sandia. Our massively-parallel computing environment presents special algorithm engineering and testing challenges (e.g. multiple, sometimes uncontrollable, nondeterminism). We hope to engage this community in addressing some of these issues.

Title: Bay Area Scientific Computing Day 2002

PI: Juan Meza, Sandia National Laboratories

Dates: February 15, 2002

CSRI POC: Juan Meza

Conference Summary

We propose to hold a one-day workshop, Bay Area Scientific Computing Day 2002, with the goal of fostering interaction and collaboration with scientific computing researchers from the Bay Area. The Bay Area Scientific Computing Day is an annual gathering to encourage the interaction and collaboration of researchers in the field of scientific computing from the Bay Area. It also provides new researchers an opportunity to present their work to the local community. This year will be the third year of the workshop. The workshop was sponsored by Lawrence Berkeley National Laboratory in its inaugural year in 2000, and by Lawrence Livermore National Laboratory in 2001. The targeted research communities are Sandia National Laboratories (SNL), Lawrence Livermore National Laboratories (LLNL), Lawrence Berkeley National Laboratories (LBNL), Stanford Linear Accelerator Center (SLAC), NASA-Ames, University of California-Davis, University of California-Berkeley, Stanford University, and out-of-town visitors in the area. Moreover, in order to foster interaction with local industry, we will invite computational scientists from local companies such as HP, IBM, Chevron, Silvaco International, SRI International, Molecular Dynamics, and Phillips.

Format

The one-day workshop will tentatively take place on February 15, 2001 at the Crowne Plaza Hotel in Pleasanton. CA located 14 miles west of Sandia National Laboratories.

Speakers

Speakers will be selected from the three major DOE labs in the Bay Area, as well as from Stanford University, University of California at Berkeley, University of California at Davis, SLAC, and NASA-Ames. The organizers are particularly interested in providing a forum for young or beginning researchers in the field to talk about their work.

Title: Third International Conference on Discrete Element Methods

PI: Ben Cook, Sandia National Laboratories
Richard Jensen, Sandia National Laboratories

Dates: September 23-25, 2002

CSRI POC: Ben Cook

Conference Summary:

Discrete Element Methods (DEM) include a suite of numerical techniques developed over the past 25 years to model granular materials, geologic media, and other non-continuum materials (e.g. ceramic powders, etc.). DEM technology is poised for great advances in the development and application of the technique. With the advent of increased computing power, specifically massively parallel computing, significantly larger and more realistic simulations of physical systems have become achievable. External researchers have recently implemented improved numerical algorithms ranging from faster contact detection schemes to parallelized, 3-D versions of DEM. These advances combined with greatly improved computing resources are enabling, for the first time, physically realistic simulations of systems at the macro-scale. The

last dedicated DEM conference was held eight years ago. Given the recent advances in the field, a gathering of the DEM community to review recent progress and to map future research directions is overdue.

Several groups at Sandia have also made significant advances recently in DEM numerical techniques that have enabled applications to modeling of blast effects for the mining industry and the modeling of near-wellbore phenomena for the petroleum industry. By leveraging this existing expertise, in combination with Sandia's pre-eminence in the field of massively parallel computing, it is anticipated that significant progress can be made in applying DEM to problems of current national security interest, namely rock penetration and weapons effects on buried targets. This conference will bring together the community of researchers and practitioners in the area of DEM modeling to review current work in the area and to discuss promising new approaches and applications of the technology.

Sandia is leading the organization of the 3rd International Conference on Discrete Element Methods. The conference will be held in Santa Fe in September 2002. Our motivation in co-sponsoring this event is to gain timely assessment of the state-of-the-art in discrete element methods, for the purpose of better supporting anticipated DOE/DP programs that will draw on this technology, and to promote Sandia's technical leadership in this area. In addition, conference promotional materials will enhance Sandia's visibility at top technical universities and will strengthen our recruiting efforts in the broader area of computer and computational science.

Title: Cluster Performance Enhancements High-Intensity Retreat (CPEHIR)

PI: Rolf Riesen, Sandia National Laboratories

Dates: September 3 – 5, 2002

CSRI POC: Rolf Riesen

Conference Summary:

We will generate and discuss ideas to enhance performance of scientific clusters. We are concentrating on enhancements to systems software that improve the performance for our users. Improvements can be in the operating system, the libraries, the runtime system, or the system management software. We are looking for ideas that improve the execution time of applications, let a system process more jobs in a given time, make system management and maintenance easier and increase uptime.

Title: The Sandia Workshop on Quantum Mechanical Techniques: Exchange-Correlation Functionals in Density-Functional Theory

PI: Ann Mattsson, Sandia National Laboratories
Alan Wright, Sandia National Laboratories

Dates: August 14 – 16, 2002

CSRI POC: John Aidun

Conference Summary:

Density-Functional Theory (DFT) was presented in two seminal papers by P. Hohenberg and W. Kohn (1964) and by W. Kohn and L. Sham (1965). DFT is an *exact* reformulation of the Schrödinger equation, which, in its time-independent form, governs the quantum mechanical state of a system of N electrons bound by M stationary ion cores. The DFT theorems establish that the lowest energy state (the "ground

state") of the N-electron system can be correctly described by a functional of the electron density, the latter being a function of 3 spatial variables. DFT thus provided a practical route to quantum calculations for large systems of atoms by obviating a formal solution to the Schrödinger equation in terms of a many-particle wave function, which depends on the 3N spatial variables of all the electrons. But while Hohenberg and Kohn proved the existence of an exact functional that describes "exchange", *i.e.*, consistency with the Pauli exclusion principle that only one electron can occupy a given quantum state, and the spatial correlation of electron positions that arises from their Columbic interactions with the other electrons and the ion cores, the form of the exact exchange-correlation functional remains unknown. Physical arguments and an appeal to simplicity have led to the two types of approximate functionals currently in use: The local density approximation (LDA), in which the exchange-correlation functional is computed using local values of the electron density (Kohn and Sham, 1965), and the generalized gradient approximation (GGA), which uses both the local density and its gradient (Perdew and Wang, 1986). During the 25 years following its discovery, DFT had only a minor impact on physics and chemistry research because of the daunting computational demands of calculations for realistic systems of interest and because DFT itself was a research topic. During the past decade, advances in algorithms and massively parallel computers have greatly expanded the size and complexity of systems that can be treated so that DFT calculations are now widely used in research programs at Sandia and most other government, industry, and academic materials research laboratories. The current prevalence of DFT calculations employing the various approximate exchange-correlation functionals has better revealed their deficiencies. We ourselves are aware of deficiencies in treating systems involving surfaces, which impacts problems of great technological interest including surface chemistry, the work of adhesion of interfaces, and the characteristics of vacancy defects. At the same time, the increasing capability and importance of DFT calculations for providing reliable material property values and understanding of basic material behaviors has created greater urgency for developing a more accurate exchange-correlation functional. Given the opportunity to increase the value and impact of DFT calculations on a growing range of Sandia programs, we propose hosting a workshop aimed at spurring the development of more accurate exchange-correlation functionals.

DFT is contributing to several Sandia projects and stands to expand its impact to others in the near term. It is currently being used in ASCI programs at Sandia to characterize the properties and performance of ferroelectric PZT ceramics and desiccants, and the process of metal corrosion and the low dose rate radiation effect in transistors. Defects in gallium nitride are being studied under NW Campaigns and a Grand Challenge LDRD program. This material has numerous applications of interest to Sandia including solid-state lighting, solar-blind ultraviolet detectors, and next-generation synthetic aperture radars. Because of its firm theoretical foundation, DFT also provides an essential underpinning for current Sandia developments in multiscale simulation methods for more complex materials. DFT, in the form of Sandia's SeqQuest code, is currently being used in this capacity at Caltech under their ASCI ASAP project. In addition, DFT will no doubt play an important role in understanding the physics and chemistry inherent in microscale (MESA) and nanoscale (CINT) systems.

High performance computing at Sandia will directly and quickly benefit from improvements in the accuracy of DFT calculations with implementation of better functionals into the two DFT quantum codes currently under development: SeqQuest (PI: Peter Schultz, 9235); Socorro (PI: Alan Wright, 1112). Both of these code development projects are partially funded by the ASCI M&PM project "First Principles Methods" (Alan Wright, PI). SeqQuest's predecessor, MPQuest, consumed the majority of cycles on the Paragon in days past. DFT quantum calculations continue to consume many cycles on CPlant and also some on Janus. This is bound to increase as these codes will be extensively used in multiscale material modeling and simulation, where DFT calculations provide the first principles results that serve as the underpinnings of a leading materials multiscale approach.

While deficiencies of the current DFT functionals are well known to exist, unsuccessful calculations are not usually published in journal articles. This circumstance makes a workshop setting a requirement for spurring progress on this important problem. The proposed workshop will bring together the right people, leading external theorists and practioners together with Sandia experts, to sit down together to detail DFT failings that they each are individually familiar with and then explore strategies for developing improved functionals. From exploratory inquiries we have confirmed commitments from Walter Kohn (UCSB), John

Perdew (Tulane), Michael Teter (Cornell), Kieron Burke (Chemistry, Rutgers), and E.K.U. "Hardy" Gross (Free University, Berlin) to attend should we be able to go forward with this workshop. David Langreth (Physics, Rutgers) has tentatively accepted, pending success in rearranging his schedule. Up to 14 additional external participants and up to 10 Sandians would be invited. The tentative list of Sandia participants includes Peter Feibelman, Dwight Jennison , Kevin Leung , Ann Mattsson , Thomas Mattsson , Normand Modine , Peter Schultz , David Teter , and Alan Wright .

Sandia will host the *Sandia Workshop on Quantum Mechanical Techniques: Exchange-Correlation Functionals in Density-Functional Theory* in Albuquerque over a period of three days during the week of August 12th. Attendance will be by invitation, only, so that the gathering can be kept small enough to facilitate efficient technical exchange. (Up to ten participants may be foreign nationals.) The format will be designed so as to promote in-depth discussions of inadequacies in current exchange-correlation functionals and to explore strategies for developing improved functionals. Formal invitations will be sent out and the venue will be selected once funding is secured. We estimate 30 participants will attend.

The draft agenda for the meeting follows. There will be two main technical sessions, one focused on cataloging and discussing the deficiencies of current approximate functionals and the other aimed at developing strategies for creating better functionals. The first evening will include a poster session showcasing the Sandia DFT codes, Socorro and SeqQuest, the Sandia projects currently making use of them, and Sandia research organizations including CSRI, MESA, and CINT. We will solicit corporate sponsorship for the two evening receptions. We will enlist the services of a technical writer to document the proceedings of the workshop for our own and the participants' future reference. We will also look into arranging for the workshop to be covered by a science writer from *Physics Today*.

Title: CSRI Workshop on Fault Tolerance

PI: Patty Hough, Sandia National Laboratories

Investigator: Tom Bressoud, Bell Laboratories
Lee Ward, Sandia National Laboratories

Dates: June 10-11-2002

CSRI POC: Patty Hough

Conference Summary:

Background: Large-scale ASCI codes such as ALEGRA, Xyce, and the SIERRA applications are expected to run for days or even weeks on thousands of processors in order to solve the problems of interest to the ASCI program. Because of the large number of processors, the complexity of the computing platforms, and the long execution times, it is reasonable to expect at least one failure to occur during any single run. More specifically, existing statistics indicate that failures can be as frequent as one every two days on 2000 processors. This number becomes progressively worse as the number of processors increases. Traditionally, the applications of interest have not been equipped to efficiently deal with failures, so when one occurs, the application crashes. While many applications write restart files to allow for warm starting after a failure, this is cumbersome and incurs an unacceptable amount of overhead. As a result, there has been movement in both the research and applications communities toward more efficient and more sophisticated strategies for fault tolerance.

In response to the growing interest in scalable techniques for fault tolerance, we are proposing to host a workshop that will bring together Sandians interested in fault tolerance and external experts on fault tolerance. The primary goals of the workshop are to:

- educate external experts about Sandia problems,
- raise Sandia awareness of state-of-the-art fault tolerance technology, and

- identify gaps between Sandia's needs and current research trends.

In order to accomplish these goals, the agenda has been designed to include the following:

- 4 technical presentations on Sandia ASCI applications and platforms,
- 6 technical presentations, by external fault tolerance experts, covering failure detection, communication infrastructures, logging/checkpointing, and compiler technology
- 1 formal discussion session.

Expected Outcome: Upon completion of the workshop, we will document the accomplishment of workshop goals in the form of a web page containing the following:

- a description of the workshop,
- electronic versions of workshop presentations,
- a summary of high-priority research directions, and
- a list of new collaborative efforts.

Title: The LCI Third International Conference on Linux Clusters: The Linux HPC Revolution

PI: Bob Balance,
High Performance Computing, Education, and Research Center (HPCERC)
The University of New Mexico

Dates: October 23-25, 2002

CSRI POC: Neil Pundit

Conference Summary:

The Linux HPC Revolution is a conference organized by the Linux Clusters Institute (LCI) for users and administrators of high- performance Linux clusters.

The Linux HPC Revolution is the premier international forum for Linux clusters users and system administrators to share information on management, administration, and scientific computing techniques on Linux clusters. The conference will feature speakers from academia, research labs, and industry involved in Linux cluster-based high-performance computing. These speakers will address efforts to integrate and develop science and engineering applications for Linux clusters, in order to achieve maximum performance and scalability.

The conference program committee is soliciting novel papers on a broad range of topics related to systems integration, operation and support, end user applications, tools, and experiences. Topics of interest include (but are not limited to) the following:

- Performance Evaluation, Analysis, and Optimization Performance Tools, debuggers, and Environments
- High Performance Applications and Libraries Experiences in Development of Highly
- Parallel Applications
- System Management and Administration
- Tools for Building and Administrating Clusters
- Scheduling and Load Balancing
- Resource Management
- Network, Interconnects, and Protocols
- Middleware for Clusters
- Meta- and Grid-Computing
- Parallel I/O, File Systems, and Storage

- Autonomic computing
- Security
- Compilers
- New Experimental and Commercial Linux Clusters

Key dates:

- Extended abstract submission deadline: May 31, 2002
- Authors notification: July 20, 2002
- Final paper submission deadline: August 26, 2002
- Early registration deadline: August 26, 2002
- Close of registration: September 30, 2002

Title: Numerical Aspects of Circuit and Device Simulation Workshop

PI: Tamara Kolda, Sandia National Laboratories
Scott Hutchinson, Sandia National Laboratories
Robert Melville, Agere Systems

CSRI POC: Tamara Kolda and Scott Hutchinson

Dates: April 4-6, 2002

Conference Summary:

We will hold a 3-day workshop with the goal of forging connections between Sandia and leading world experts in numerical analysis and numerical aspects of circuit and device simulation. As part of the ASCI program, the Xyce™ Parallel Electronic Simulator is being developed in support of electrical designers of Sandia National Laboratory. We will simulate circuits of unprecedented size (i.e., millions of unknowns) and, therefore, require the use of state-of-the-art and even new approaches for improving and parallelizing critical numerical kernels. For example, most circuit simulators use direct methods to solve the linear equations that arise each iteration of the nonlinear solve, but we must use iterative methods in order to solve problems of the size we are contemplating. The matrices themselves have unusual characteristics which make them difficult to partition and to precondition. The nonlinear solver is plagued with problems due to the highly nonlinear behavior of devices such as diodes and bipolar junction transistors (BJTs). Solutions to these problems will contribute to the linear and nonlinear solvers projects as well as to the Xyce project. Overall, the topics we are concerned with are the solution of:

- Stiff coupled time-dependent ODE/DAEs,
- Large-scale nonlinear equations,
- Sparse, ill-conditioned linear systems, and
- PDE solutions for semiconductor devices.

Goals. The specific goals of the workshop are to:

- Foster and strengthen collaborations between Sandia and leaders in circuit and device modeling and in numerical analysis.

- Exchange information on the state-of-the-art methods for linear systems, nonlinear equations, ODE/DAEs and PDEs as applied to circuit and device simulation.

- Assess barriers and challenges to creating a scalable parallel circuit simulator and identify promising algorithmic and software approaches for overcoming these barriers.

Format. The workshop will be help March 27{29, 2002 at The Bishops Lodge in Santa Fe, NM. Each day will consist of two presentation sessions and one discussion session. The presentation sessions will consist of 3{4 30-minute presentations. The discussion session will be 1-2 hours and consist of break-out groups focused on specific discussion topics which summaries from each break-out group at the end of the session. Approximately two-thirds of the sessions will focus on the numerical aspects of circuit simulation (linear systems, nonlinear equations, ODE/DAEs), and the remaining one-third will focus on the numerical aspects of device modeling (PDEs). The presentations will be collected and posted online after the workshop.

Title: Petaflops Supercomputer Workshop

PI: Erik P. DeBenedictis, Sandia National Laboratories

Dates: June 17 – 18 2002

CSRI POC: Erik DeBenedictis

Conference Summary:

Questions to be posed and discussed at the workshop include:

> **USAGE.** How will the requirements that DOE's applications make of the underlying hardware change in the next 5-10 years? How will the applications mix change when considering new applications?
>
> **HARDWARE.** What will continued evolution of existing Teraflops clusters and Massively Parallel Processors (MPPs) yield in 5-10 years? How suitable will it be for the applications expected to be in use at that time? What can DOE do to influence the course of evolution so that machines better suit DOE's applications?
>
> **OPERATING SYSTEMS.** If we project current evolutionary directions for software like Linux, Light Weight Kernels (LWK), batch schedulers, etc. 5-10 years into the future, how well will it work with projected hardware and applications?
>
> **CALL TO ACTION**. Discuss methods to:
>
> 1. Reduce uncertainty by characterizing applications and workload and projecting hardware and software to the future
> 2. Identify weak technology, considering the appropriateness of creating prototypes to show problems/solutions and recommending the funding research in critical areas?
> 3. Discuss methods of influencing industry

Title: Workshop on OS and Runtime issues for High-Performance Computing Systems

PI: Arthur B. (Barney) Maccabe, UNM,

Investigator: Ron Brightwell, SNL
Al Geist, ORNL
Ron Minnich, LANL
Rusty Lusk, ANL

Dates: February 28, 2002

CSRI POC: Neil Pundit

Conference Summary:

ASCI applications were showing 75% parallel efficiency when scaled to thousands of processors on ASCI Red. On the ASCI Blue machines (Blue Mountain and Blue Pacific) the parallel efficiency of these applications dropped below 30%. Currently they are showing parallel efficiencies of approximately 40% on ASCI White. While many factors may impact this dramatic drop in parallel efficiency, one obvious difference between ASCI Red and the other ASCI machines is the design of the OS and runtime environments. The OS/Runtime environment for ASCI Red is based on Puma, a light weight message passing kernel, in contrast to the full-featured operating systems, IRIX and AIX, that provide the basis for the ASCI Blue and White machines.

In this workshop, we will explore the issues that must be addressed in developing operating systems and runtime environments for the next generation of ASCI class systems.

Title: Workshop on the Implementation of Multi-PIM Systems (WIMPS)

PI: Thomas Sterling
California Institute of Technology

Investigator: Bill Gropp, Argonne National Laboratory
Peter Kogge, University of Notre Dame
Bob Lucas, USC ISI
Sheila Vaydia, LLNL
Hans Zima, University of Vienna

Dates: February 4 – 6, 2002

CSRI POC: Neil Pundit

Project Summary:

The challenge of achieving multiple Petaflops performance for a broad range of applications while operating within practical constraints of size, power, cost, and reliability may be realized within the next five years through a synthesis of advanced semiconductor fabrication technology, innovative architecture, and new runtime and compiler software by exploiting an emerging class of computer structure referred to as *Processor in Memory* or PIM. The enabling foundation for PIM is the recent ability to fabricate both DRAM cells and CMOS logic on the same semiconductor die, permitting new architectures based on close physical and logical relationships between these two historically separate primary components. High bandwidth, low latency, low overhead, low power, small size, and other beneficial properties may result from PIM technology. If coupled with 3-D packaging and advanced cooling techniques, high-density systems of unprecedented performance per unit area may be possible. It is estimated that a Petaflops system (with 20 Terabytes of memory) could be implemented in a cubic meter. In the near future, a Petabyte of main memory could be packaged in a comparable size. This opportunity in ultra high end computing is accompanied by potential impact in other fields such as embedded, mobile, autonomous, and space-borne computing systems. Finally, PIM may be employed to augment the capabilities of otherwise conventional systems to facilitate and accelerate certain classes of processing such as data intensive computation. Nonetheless, even after ten years of research in to this potentially powerful strategy, PIM has had little actual impact on real world systems and applications. The reasons for this are complex and reflect technical, economic, and sociological problems. In spite of such obstacles the potential value of such systems is of sufficient scale that a concerted initiative may be justified to address the challenges and bring PIM to practical realization.

A workshop is being organized to consider the opportunities, challenges, and strategies governing the future evolution and application of PIM technology, architecture, software, and applications. The Workshop on the Implementation of Multi-PIM Systems (or WIMPS) is structured as an intense three-day forum for an in-depth inquiry of the technical and other steps that must be taken to realize the promise of this broad class of system structures. In particular, scalable ensembles of PIMs and the necessary logical and physical relationships among them will be explored.

The **Goal** of the workshop is to:

Advance the field of PIM system architecture in order to accelerate their implementation, availability, and application for a broad range of real-world roles including scalable embedded systems, integrated accelerators, and high performance computing.

The **Objectives** of the workshop are to:

- Establish the potential value and opportunities of enabled by PIM
- Capture the state of the art, experience, and practices with PIM
- Identify the critical technical challenges to the achievement of advanced PIM-based systems
- Develop tenable strategies to fully exploit PIM potential
- Devise a roadmap and recommendations for future directions, collaborations, and research sponsorship

Topics to be discussed will include, but not be restricted to:

- existing projects (e.g. IRAM, Blue Gene/Cyclops, Gilgamesh),
- engineering (e.g. packaging, cooling, interconnects),
- semiconductor technology roadmap,
- architecture,
- communication,
- multithreading,
- execution models,
- low level system software,
- programming models,
- applications,
- scalability,
- markets,
- interoperability, and
- schedules.

The agenda will be a mix of formal briefings, plenary discussions, and breakout groups targeted at specific questions, as well as adequate time for unstructured discussions. The product of this meeting will be compendium of briefing and referenced background material, an extended summary of the findings and recommendations, and a publishable white paper on the state-of-the-art of PIM device technology and system architecture. Participants will number approximately 36 and be invited from industry, academia, and national labs and centers. Sponsorship is provided by Federal agencies with additional contributions from industrial sponsors.

Chapter 5. Seminar Abstracts

The CSRI hosts an active short-term visitor program, which is closely aligned with a seminar schedule. Short term visit typically last between two days and two weeks and include a seminar as well as a broad range of meetings and collaborations with Sandia staff. The longer visits are encouraged to increase the likelihood that a significant collaboration will develop. The abstracts for the CSRI seminars are listed below in alphabetical order by speaker.

Title: Parallel Applications for Quantum Physics and Chemistry

Speaker: Paul M. Alsing,
Senior Research Physicist, Albuquerque High Performance Computing Center,
Center for Advanced Studies, Department of Physics and Astronomy, UNM

Date/Time: Wednesday, October 30, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: In this talk I will discuss some recent problems in quantum physics and chemistry that I have been working on at the Albuquerque High Performance Computing Center and the parallelism I have employed to speed up execution time and increase problem size. These projects have been collaborative research efforts with other members of the University of New Mexico, as well as with researchers at Sandia and Los Alamos National Laboratories. I will discuss an implicit split operator Schrodinger propagator used for investigating two body centered atom-ion collisions, the recovery of classical dynamics from coupled quantum systems through continuous measurements, the use of a parallel force decomposition for the simulation of water with both translation and rotational degrees of freedom, and a simple parallelization of strong beam-beam collisions in cyclotron particle accelerators combining the force decomposition algorithm with fast multipole methods. I will also discuss an interactive visualization tool we developed based on OpenDX TCP/IP modules, which we have used to interrogate and interact with real-time executing parallel codes.

CSRI POC: Bruce Hendrickson, 845-7599

Title: Approximation by quadrilateral finite elements

Speaker: Professor Doug Arnold
University of Minnesota Math department and director of the IMA

Date/Time: Tuesday, February 19, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Finite element spaces are usually defined by a three step process. First, a finite dimensional space of functions, often polynomial, is given on a reference element, such as the unit simplex or cube. Second, this space is transferred to arbitrary images of the reference element via transformations of a given class, e.g., affine or multilinear. Finally, the finite element space is pieced together on a mesh consisting of transformed images of the reference elements. A fundamental issue is to determine the rate of approximation afforded by the resulting finite element space as the mesh is refined. Assuming that the space given on the reference element contains all polynomials of a certain degree and the transformations from the reference element are affine, the theory establishing optimal rates of approximation is one of the pillars of the mathematical foundations of finite elements. However the situation is more complicated for non-affine transformations, such as are needed for quadrilateral meshes, and the results in this case are--

suprisingly--not widely known. In this talk I will review the classical theory of finite element approximation and prove a new result establishing necessary and sufficient conditions for optimal order quadrilateral approximation. One consequence is that many popular finite element schemes attain lower rates of convergence on quadrilateral meshes than has been generally believed.

CSRI POC: Richard Lehoucq, 845-8929

Title: The Institute for Mathematics and its Applications

Speaker: Professor Doug Arnold
University of Minnesota Math dept and director of the IMA

Date/Time: Tuesday, February 19, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: The Institute for Mathematics and its Applications (IMA) in Minneapolis is a leading research center, founded by the National Science Foundation in 1982. The primary mission of the IMA is to increase the impact of mathematics by fostering research of a truly interdisciplinary nature, establishing links between mathematics of the highest caliber and important scientific and technological problems from other disciplines and industry. Through a variety of programs, it provides opportunities for scientists, mathematicians, and engineers from academia and government labs and industry to make contact and interact with each other, to learn about new developments, and to stimulate the study of interesting and relevant problems and their solution. In this informal presentation the new director of the IMA will discuss IMA operations and upcoming programs in order to promote participation and input from Sandia researchers.

CSRI POC: Richard Lehoucq, 845-8929

Title: Statistical Characterization and Probabilistic Modelling of Material Microstructures

Speaker: Sanjay Arwade, Dept. of Civil Engineering
Johns Hopkins University

Date/Time: Thursday, November 7, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: Microstructural heterogeneity is a significant contributor to randomness in the behavior of a variety of physical systems. Examples include crack initiation in solids, electrical conductivity, and damage in biological systems. The ability to generate realizations of random microstructures which match, in a statistical sense, physical material samples, will aid in the investigation of such phenomena through numerical simulation. Through the example of a polycrystalline metal, an approach is illustrated for statistically characterizing microstructures, developing probabilistic models for features of the microstructure, and generating digital realizations of the material. Some simple examples are also given which demonstrate the utility of such realizations in the simulation of the mechanical behavior of materials. This approach is thought to be particularly useful for application to problems which do not necessarily involve novel mechanics, but which, as a result of their small scale, are highly sensitive to details of the microstructure.

CSRI POC: Steve Wojtkiewicz, 284-5482

Title: Tensor Methods for Solving Large-scale Systems of Nonlinear Equations

Speaker: Brett W. Bader and Prof. Robert B. Schnabel
University of Colorado, Boulder

Date/Time: Tuesday, June 4, 2002, 10:00-11:00 am (PT)

Location: Building 921 Room 137 – SNL CA
Building 980 Room 95 – SNL NM (Videolinked)

Brief Abstract: This talk will present an iterative, Krylov-subspace based approach to solving large-scale systems of nonlinear equations using tensor methods. Tensor methods are a class of methods that base each iteration on a model that augments the standard linear models with a limited second-order term. Standard tensor methods for solving small, dense systems of nonlinear equations have performed especially well on problems where the Jacobian matrix is rank deficient or ill-conditioned at the solution. This success originates from the special, restricted form of the quadratic term included in the local tensor model that provides information lacking in a (near) singular Jacobian. This talk will present two variants of an iterative tensor method for solving large-scale systems of nonlinear equations. A third method by Feng and Pulliam will also be included for comparison. These three iterative methods are all based on Krylov subspace projection techniques and intend to capitalize on the success of direct tensor methods, even when only approximately solving the local model. The proposed tensor-Krylov methods require hardly more storage or arithmetic per iteration than a Newton-Krylov method yet still exhibit the superlinear convergence behavior of traditional tensor methods on certain ill-conditioned and singular systems. The ability of the proposed methods to solve the local tensor model to a specified accuracy with very little extra arithmetic beyond standard GMRES is an advantage over existing methods. Preliminary numerical results show promise that these methods will outperform Newton-GMRES on certain problems.

CSRI POC: Tammy Kolda, (925) 294-4769

Title: Parallel Sparse Preconditioning Techniques for Circuit Simulation

Speaker: Achim Basermann
NEC Europe Ltd., C & C Research Laboratories
Sankt Augustin, Germany

Date/Time: Monday, April 8, 2002, 11:00-12:00 noon (MST)

Location: Building 980 Room 95
Building 921 Room 137 (Videolink to CA)

Brief Abstract: The simulation of large, highly integrated circuits leads to non-linear differential algebraic equations. For integration of these equations, accurate solution methods for sparse linear systems are required within the non-linear iterations. The corresponding matrices are real, non-symmetric, very ill-conditioned, have an irregular sparsity pattern, and often include a few dense rows and columns.

When the systems become large, iterative solvers are very likely to outperform direct linear solvers. For convergence acceleration of iterative solvers, parallelization and appropriate preconditioning are suited techniques to reduce the execution time.

A parallel BiCGstab algorithm with distributed Schur complement (DSC) preconditioning is presented which achieves an accuracy of the solution similar to a direct solver but is distinctly faster for large problems. The efficiency of the method is increased by transforming the equation system into a problem without dense rows and columns as well as by exploitation of parallel graph partitioning methods. The cost of local, incomplete *LU* decompositions are decreased by fill-in reducing re-ordering methods of the matrix

and a threshold strategy for the factorization. Moreover, it is shown how the DSC preconditioner can be applied to the symmetric case.

The efficiency of the parallel solvers is demonstrated with real problems from circuit simulation and structural mechanics on a PC cluster. In particular, the effect of partitioning and re-ordering algorithms on the iterative solution are presented, and the computational overhead of these methods in comparison with the costs of the DSC solver are discussed.

CSRI POC: David Day, 844-1868

Title: Consistent stabilization techniques in finite element practice

Speaker: Marek Behr
Rice University

Date/Time: Monday, March 4, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Finite element formulations of physical flow problems frequently require stabilization techniques in order to provide necessary robustness. Distinct from ad hoc fixes, these techniques possess firm theoretical foundations, and in some cases are known to produce optimal approximations. They circumvent, or compensate for, undesirable features of the standard Galerkin formulation, yet they retain its accuracy and consistency characteristics.

This overview considers Streamline-Upwind Petrov-Galerkin (SUPG) stabilization for advection-dominated flows, Pressure-Stabilizing Petrov-Galerkin (PSPG) stabilization that enables the use of equal-order interpolation functions, and Least-Squares (LS) stabilization which combines SUPG and PSPG in a single elegant form. The LS concept is then naturally extended to stress-velocity-pressure formulation, which is used to simulate incompressible flows of viscoelastic fluids, and to hyperbolic kinematic boundary conditions. which are encountered in free-surface flows. Example applications of some of these methods are presented, including simulations of free-surface flows in waterways.

CSRI POC: Pavel Bochev, 844-1990

Title: Discrete Mathematical Applications of LINK

Speaker: Jonathan Berry
Lafayette College

Date/Time: Wednesday, June 26, 2002, 9:00 – 10:00 am

Location: Building 980 Room 95

Brief Abstract: The LINK project is an NSF-sponsored effort to create a software system for discrete mathematics analogous to Mathematica, Maple, etc. The resulting system is much smaller, of course, but also more flexible in certain ways. In this talk, the project leader will cover the history of the creation of the system, summarize its architecture, and demonstrate several research applications that demonstrate its flexibility. These include graph partitioning heuristics, ear decomposition, and questions about tournaments.

CSRI POC: Cynthia Phillips, 845-7296

Title: Multiresolution Techniques on a Multiresolution Multidisplay Parallel Image Compositing System

Speaker: William J. Blanke
University of Texas at Austin

Date/Time: Thursday, May 23, 2002, 11:00 am-12:00 noon

Location: Building 980 Room 95

Brief Abstract: In most computer graphics applications resolution is a tradeoff. Using low resolution images provide a low quality display, but typically allow higher frame rates because less data needs to be computed. High resolution images, on the other hand, give the best display, yet are hindered by slower refresh times and thus limit user interactivity. Low image quality and low user interactivity are both detriments to computer graphics visualization applications. The question is then what can be done to minimize this impact.

The aim of this project is to explore how to use multiresolution in order to provide the best balance between image quality and user interactivity on a parallel multidisplay multiresolution image compositing system with antialiasing called the Metabuffer. The architecture of the Metabuffer, a simulator written in C++, and a Beowulf cluster based emulator have been developed in this research. Additional supporting hardware and software needed in the project include an algorithm to partition datasets into Metabuffer viewports and a wireless visualization control device.

CSRI POC: Philip D. Heermann, 844-1527

Title: Technology Trends and Teraflop Chips for Petaflop Systems

Speaker: Doug Burger and Stephen W. Keckler
Department of Computer Sciences
The University of Texas at Austin

Date/Time: Tuesday, March 12, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: Modern microprocessors are reaching an inflection point, where the traditional methods of extracting performance are reaching fundamental limits. We will show that three trends will change future high-performance designs significantly. First, growing wire delays will prevent conventional out-of-order issue processors from exploiting much more instruction-level parallelism than today. Second, hard limits on pipeline depth will prevent the continuation of the decade-long, 40% annual acceleration of processor clocks. Finally, limited off-chip bandwidth will prevent future chips from solving performance limits solely by providing large numbers of processing cores on a die.

In the second half of this talk, we will describe the UT-Austin TRIPS processor, which addresses the problems enumerated above. The TRIPS processor consists of a small number of large processing cores on a single die. Each core contains large numbers of ALUs, which can either be used to accelerate sequential codes through aggressive speculation and dataflow execution, or can accelerate scientific codes by providing streaming and vector-like support. The TRIPS processor, when fabricated in 35nm technology, will have a single-chip peak performance of 4 Teraflops. Finally, a major research goal for the TRIPS architecture is to achieve significant fractions of peak performance on high-performance computing workloads. We will describe a set of techniques that we are exploring to mitigate package-level bandwidth restrictions and allow high chip-level utilization.

CSRI POC: Erik DeBenedictis, 284-4017

Title: Spatially Adaptive MultiWavelet Representations on Unstructured Grids with Applications to Multidimensional Computational Modeling

Speaker: Julio E. Castrillon

Date/Time: Tuesday, April 2, 2002, 10:00-11:00 am

Location: Building 980 Room 95 (NM)
Building 921 Room 137 (CA video link)

Brief Abstract: In this talk we discuss the development of wavelet representations for complex surfaces, with the goal of demonstrating their potential for 3D scientific and engineering computing applications. Surface wavelets were originally developed for representing geometric objects in a multiresolution format in computer graphics. However, we further extend the construction of surface wavelets and prove the existence of a large class of multiwavelets in R^n with vanishing moments around corners that are well suited for *complex geometries*. These wavelets share all of the major advantages of conventional wavelets, in that they provide an analysis tool for studying data, functions and operators at different scales. However, unlike conventional wavelets, which are restricted to uniform grids, surface wavelets have the power to perform signal processing operations on *complex meshes*, such as those encountered in finite element modeling. This motivates the study of surface wavelets as an efficient representation for the modeling and simulation of physical processes. We show how surface wavelets can be applied to partial differential equations, cast in the integral or weak form. We analyze and implement the wavelet approach for a model 3D potential problem using a surface wavelet basis with linear interpolating properties. We show both theoretically and experimentally that an *O(h^2_n)* convergence rate, h_n being the mesh size, can be obtained by retaining only $O(N(\log N)^a)$ entries, where a is a small positive constant, in the discrete operator matrix, where N is the number of unknowns. Moreover our theoretical proof of accuracy vs. compression is applicable to a large class of Calderon-Zygmund integral operators. In principle, this convergence analysis may be extended to higher order wavelets with greater vanishing moments. This results in higher convergence and greater compression.

CSRI POC: Pavel B. Bochev, 844-1990

Title: Linear-Scaling Quantum Calculations: MondoSCF

Speaker: Matt Challacombe
Los Alamos National Laboratory

Date/Time: Thursday, April 4, 2002, 1:30-2:30 pm

Location: Building 980 Room 95

Brief Abstract: The near sighted principle, that quantum interactions in insulating systems are local, has over the last decade guided the development of numerous "fast" numerical methods for ab initio quantum chemistry that achieve a reduced, O(N) cost with system size N. In addition to linear scaling, the near sighted principle also promises O(1) communication costs with respect to N/p and hence scalability if locality can be exploited in a careful way.

In this talk I will briefly overview the MondoSCF project for O(N) quantum chemistry and its current and future capabilities. I will then outline a global strategy for achieving scalable parallelism that extends current methods of irregular parallel computation. These involve ordering schemes based on a novel space

filling curve to achieve data locality, new data structures to support early onset linear scaling and fine grained decompositions, and methods for distributed dynamic load balancing.

CSRI POC: Susan Rempe, 845-0253

Title: Scientific Data Management, I/O, Mining and Analysis

Speaker: Professor Alok Choudhary
Northwestern University

Date/Time: Thursday, August 1, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: Management, storage, efficient access and analysis of 100's of GBs to 100's of TBs of data, that is likely to be generated and/or used in various phases of large-scale scientific experiments and simulations, such as those in ASCI application domains, are extremely challenging tasks. Current data management and analysis techniques do not measure up to the challenges posed by such large-scale requirements in term of performance, scalability, ease of use and interfaces. Tera-scale computing requires newer models and approaches to solving the problems in storing, retrieving, managing, sharing, visualizing, organizing and analyzing data at such a massive scale.

In this talk I will describe the research and development work being performed at Northwestern University to address the above problems. In particular, we will describe the architecture and implementation of a metadata management system which allows the user to store, analyze and use access patterns, relationships amongst data sets, data analysis and I/O optimizations for scientific applications. I will describe the use of automatic I/O optimization techniques that can be incorporated into applications in a seamless fashion. I will also briefly present data mining techniques for processing results from scientific simulations. I will also summarize a system which uses data mining techniques for querying images. Finally, I will introduce some I/O optimization techniques for parallel I/O libraries and file systems. These techniques and the corresponding results will be presented at SNL later during the summer by my collaborators and students in a different presentation.

CSRI POC: Neil Pundit, 845-7601

Title: Generalized Fourier Analysis of Semi-Discretizations of the Advection-Diffusion Equation

Speaker: Mark A. Christon, Thomas E. Voth
Computational Physics R&D (9231)

Mario J. Martinez
Multiphase Transport Processes Department (9114)

Date/Time: Monday, August 5, 2002, 9:00-9:45 am

Location: Building 980 Room 95

Brief Abstract: This talk presents a detailed multi-methods comparison of the spatial errors associated with finite difference, finite element and finite volume semi-discretizations of the scalar advection-diffusion equation. The errors are reported in terms of non-dimensional phase and group speeds, discrete diffusivity, and artificial diffusivity. It is demonstrated that Fourier analysis (aka von Neumann analysis) provides an automatic means for separating the spectral behavior of the discrete advective operator into its

symmetric dissipative and skew-symmetric advective components. Further it is demonstrated that streamline upwind Petrov-Galerkin and its control-volume finite element analogue, streamline upwind control-volume, introduce both an artificial diffusivity and an artificial phase speed. These terms are in addition to the usual semi-discrete artifacts observed in the discrete phase, group and diffusivity. The generalized Fourier analysis may also be used to extract asymptotic estimates of local truncation error independent without resorting to Taylor series analysis. This permits comparison of methods that are not based on Taylor series, e.g., Galerkin's method, with methods that are. Finally, resolution requirements may be obtained in terms of phase and group speeds, and artificial and discrete diffusivity. Ultimately, this work can only be considered a first step in a multi-methods comparison. As such, this work is intended to identify some of the relative strengths and weaknesses of multiple numerical methods in the context of advection-diffusion processes.

CSRI POC: John Shadid, 845-7876

Title: Evaluating the Potential of PIM-based Systems for Radiation Transport Calculations

Speaker: Thomas Christopher
SNL Consultant

Date/Time: Friday, August 16, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: Processor in Memory (PIM) designs are being proposed as future supercomputers due to their abundance of compute power and low, on-chip memory latency. There are concerns that need to be addressed related to the limited quantity of on-chip memory and low bandwidth to external DRAM.

This talk presents the status of on-going work to project the performance of regular-mesh SN radiation transport codes on PIM architectures.

Counting the cost of the hardware multiplied by the time it is held, we can calculate the cost-effectiveness of a system. We find that PIM-based systems can be as cost-effective as SMP clusters and in some cases an order of magnitude more cost effective, but it is highly dependent on the parameters of the problem. Minimizing the effects of off-chip memory access is indeed important.

CSRI POC: Erik DeBenedictis, 284-4017

Title: Transient and Pseudo-Transient Integration Methods

Speaker: Todd S. Coffey
North Carolina State University

Date/Time: Monday, April 8, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: There are two numerical methods that I will discuss in this talk. The first is pseudo-transient continuation, a steady-state partial differential equation solver which uses variable size time-steps to track transient problem dynamics until switch-over to Newton's method to regain a fast convergence rate. This method is globally convergent and outperforms a line-search method, Newton-Armijo, on a upersonic Euler problem. Pseudo-transient continuation also has application in differential-algebraic equation problems. I will present an updated theory for pseudo-transient continuation which establishes a global convergence result for semi-explicit index 1 differential-algebraic equations. The second method is an

extension to PVODE, a parallel differential equation solver maintained by CASC at LLNL, to reduce the cost of nonlinear solves used at each time step. This cost is reduced by solving the nonlinear problem on a coarser mesh, moving the solution back to the fine mesh, and refining on the fine mesh by linear solves.

CSRI POC: Richard B. Lehoucq, 845-8929

Title: Large Eddy Simulation as a Reduced Order Model for Optimal Control of Turbulence

Speaker: S. Scott Collis
Mechanical Engineering and Materials Science
Rice University

Date/Time: Monday, December 9, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: This talk focuses on the application of optimal control to develop control strategies for wall-bounded turbulent flows. A novel aspect of our research is that we use a reduced order model, namely large-eddy simulation (LES), in order to both predict the state of the turbulent flow as well as the optimal control. By so doing, our LES based methods exploit the success of the dynamic subgrid-scale model to greatly improve the efficiency of optimal control formulations applied to turbulent flows. We begin by briefly presenting techniques for optimal control of turbulent flows based on the dynamic subgrid-scale LES model. The optimal control formulation is implemented using model-based predictive control where the flow sensitivity is computed from the adjoint LES equations. Our LES results for optimal control of terminal turbulent kinetic energy are compared to Direct Numerical Simulation (DNS) at turbulence Reynolds numbers of $Re_\tau = 100$ and 180. These comparisons indicate that, similar to DNS, optimal control based on LES can laminarize low Reynolds number turbulent channel flow $Re_\tau = 100$ but with significantly lower computational expense. At $Re_\tau = 180$ the optimal control yields 40% drag reduction but relaminarization is not achieved. Results are also presented for a novel hybrid LES/DNS scheme in which the optimization iterations are performed using LES while the flow is advanced in time using DNS. These hybrid simulations retain the computational efficiency of LES and the accuracy of DNS. Results from hybrid simulations clearly demonstrate that the controls computed based on LES optimization are also viable in the context of DNS. In all cases, the agreement between LES, DNS, and hybrid LES/DNS indicates that reliable turbulence control strategies can be efficiently developed based on LES.

CSRI POC: Bart van Bloemen Waanders, 284-4676

Title: Discontinuous Galerkin Methods for Turbulence Simulation

Speaker: S. Scott Collis
Mechanical Engineering and Materials Science
Rice University

Date/Time: Wednesday, December 11, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: It is generally recognized that high-order spatial discretizations are advantageous when performing Direct Numerical Simulation (DNS) or Large Eddy Simulation (LES) of turbulent flows. In order to obtain higher-order accuracy, many numerical methods traditionally used for DNS and LES, such as infinite-difference and spectral methods, directly enforce higher regularity in the computed solution through the use of wide stencils or smooth basis functions. Unfortunately, this requirement for higher

regularity often comes at the expense of flexibility. Here, we take a different approach by applying the discontinuous Galerkin method (DGM) to LES of turbulent flows. In DGM, the computed solution is discontinuous between element interfaces with adjacent elements communicating through appropriate numerical fluxes similar to those used infinite-volume methods. High order accuracy is achieved by locally enriching the function space on each element. The flexibility of the DGM can be exploited for LES in a number of ways including: unstructured meshes, local refinement, multiscale models, and hybrid modeling. This talk focuses on the formulation and software implementation of a DGM flow-solver designed for large-scale turbulence simulation. This new flow-code has been implemented in fully modern ANSI/ISO C++ using the Standard Template Library and generic programming concepts. The MPI-2 library (including MPI-IO) is used for the parallel implementation and performance results on our 82 processor Linux cluster will be shown. Numerical results will be shown for a number of model problems including: steady-state inviscid flows, unsteady laminar flows, aeroacoustics, and fully turbulent flows.

CSRI POC: Bart van Bloemen Waanders, 284-6746

Title: File Consistency in a Parallel Environment

Speaker: Kenin Coloma
Northwestern University

Date/Time: Tuesday, September 17, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Concurrent file access on NFS and ENFS requires application developers to meticulously plan IO. MPI-IO can greatly alleviate this burden on application writers, and also provides a means for transparent performance optimizations. ROMIO, a widely used implementation of MPI-IO can still run into problems with file consistency under a set of reasonable conditions - namely client-side file caching. Extending the collective IO routines in ROMIO, this file consistency issue can be resolved with minimal performance degradation.

CSRI POC: Eric Russell, 844-3679

Title: REALITYGRID: An E-Science Testbed for Condensed Matter and Materials

Speaker: Professor Peter Coveney, Queen Mary, University of London

Date/Time: Tuesday, January 8, 2002, 10:00-11:00 am

Location: Building 980 Room 1

Brief Abstract: A major EPSRC grant has just been awarded to a consortium of universities and collaborating institutions under the U.K. government's new initiative on e-science. Called RealityGrid, the project aims to grid-enable the realistic modelling and simulation of complex condensed matter structures at the meso and nanoscale levels as well as the discovery of new materials. The project also involves applications in bioinformatics and its long term ambition is to provide generic technology for grid based scientific, medical and commercial activities.

"Grid computing" refers to an ambitious and exciting global effort to develop an environment in which individual users can access computers, databases and experimental facilities simply and transparently, without having to consider where those facilities are located. The Global Grid Forum is a starting point for

more information; an excellent example of a project dedicated to realising the Grid concept is The Globus Project. The UK Grid Support Centre provides support for UK grid applications.

RealityGrid proposes to extend the concept of a Virtual Reality center across the grid and links it to massive computational resources at high performance computing centres and experimental facilities. Using grid technology to closely couple high throughput experimentation and visualization, RealityGrid will move the current bottleneck out of the hardware and back to the human mind. A twin-track approach will be employed within RealityGrid: a "fast track" will use currently available grid middleware to construct a working grid, while a "deep track" will involve computer science teams in harnessing leading-edge research to create a robust and flexible problem-solving environment in which to embed the RealityGrid.

RealityGrid is a collaboration between distinguished teams of physical scientists, computer scientists and software engineers. To meet its objectives, it will utilise a computing environment built around the UK's most advanced computing technology and infrastructure.

CSRI POC: Steve Plimpton, 845-7873

Title: Fully Automatic hp-Adaptive Simulations

Speaker: Leszek Demkowicz
Texas Institute for Computational and Applied Mathematics (TICAM)
The University of Texas at Austin

Date/Time: Tuesday, November 12, 2002, 9:00-10:00 am (MST)

Location: Building 980 Room 95 (Sandia-NM)
Building 921 Room 137(Sandia-CA)

Brief Abstract: I will present an overview of hp-adaptive finite element (FE) technology for solving a general class of equations involving H1- and H(curl)- conforming FE discretizations. We will begin with a short review of the Rham diagram for hp FE spaces and discussion of projection-based interpolation operators. Next, I will comment on new data structures and implementation of 3D and 2D FE codes supporting ANISOTROPIC h- and p-refinements, for BOTH H1- and H(curl)-conforming (edge) finite elements. Finally, I will present an hp-adaptive strategy that produces automatically a sequence of hp-refined FE meshes that deliver exponential convergence in terms of the problem size (number of d.o.f.). Given a (coarse) mesh, we produce the next optimal (coarse) mesh by minimizing the projection-based interpolation error of the fine mesh solution. The fine mesh is obtained from the coarse mesh through a uniform, global hp-refinement, i.e. each element is broken into four (2D, eight in 3D) smaller elements, and the order of approximation p is raised by one. The fine mesh solution is obtained using a two-grid solver operating on the coarse/fine grid and based on a V-cycle with an optimal relaxation. Time permitting, I will present a number of non-trivial examples including 2D and 3D computations (2D and axisymmetric 3D Maxwell equations, 2D heat conduction, 2D elasticity, 3D elliptic problems), demonstrating the potential of the methodology.

CSRI POC: Pavel Bochev, 844-1990

Title: A Petaflop in every Pot

Speaker: Dr. Monty Denneau
IBM Watson Research Center

Date/Time: Wednesday, April 3, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Relying on the scaling of current microprocessors and parallel machines, it is not possible to reach one petaflop performance in the next few years. Both the cost (a billion or so if you scale ASCI White) and the power (practically a small reactor) are out of reason. It is necessary to rethink the whole enterprise.

We describe the Cyclops architecture as part of the Blue Gene project. Rather than assembling a modest number of very complex machines, we use literally millions of extremely simple processors, each executing a minimalist instruction set. The talk will also address packaging, reliability and yield issues, along with discussion of suitable applications.

CSRI POC: Erik P. DeBenedictis, 284-4017

Title: *Ab Initio* Calculation of Electrical and Optical Properties of Metals in the Warm Dense Matter Regime

Speaker: Michael Desjarlais
Pulsed Power Sciences Center, Dept. 1674

Date/Time: Tuesday, August 13, 2002, 9:30-10:30 am

Location: Building 980 Room 95

Brief Abstract: Accurate computer modeling of electrically or magnetically driven high energy density physics experiments in the vicinity of the metal-insulator transition has, in the past, been hampered by large uncertainties in the electrical properties of these materials, and for many materials of interest this is still the case. We have begun a program to improve the accuracy of our conductivity and opacity tables and algorithms for these materials --- our most complete treatment to date being that of aluminum. A critical regime of interest for many of the pulsed-power driven experiments at Sandia is from moderate compressions over solid density down to one hundred fold expansions from solid, and temperatures from ambient up to several eV. For much of this region, the electrical conductivity has not been experimentally measured, and analytical results are limited by the uncertainties remaining in modern treatments of these strongly coupled and degenerate, or weakly degenerate systems. To improve our understanding of this warm dense matter regime, we have performed numerous *ab initio* calculations of the optical conductivity of aluminum, copper, and tungsten using a combination of DFT based molecular dynamics and the Kubo-Greenwood formula. Through a Kramers-Kronig transformation of the real part of the optical conductivity we obtain the absorption coefficient and reflectivity. Under those conditions for which data is available, the agreement is very good. Using the results from these calculations, in conjunction with data and theory where available and reliable, we have generated wide range SESAME format tables for the dc conductivity and are currently working on improvements to the low energy opacities. Subsequent computer simulations of high energy density experiments, such as magnetically launched flyer plates and exploding wires, give very good agreement with the experiments

CSRI POC: John Aidun, 844-1209

Title: A Bayesian Approach to Expression Level Estimation from Gene Array Data

Speaker: Ron Dror
MIT Department of Electrical Engineering and Computer Science

Date/Time: Monday, January 14, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: By characterizing the activity levels of tens of thousands of genes simultaneously, gene arrays promise to revolutionize many areas of medicine and biology. Unfortunately, array technologies suffer from high levels of measurement noise. I will describe a statistical technique to estimate gene expression levels or expression level ratios from one or more gene array experiments, incorporating a model of measurement noise and prior information about biological expression levels. The Bayesian Estimation of Array Measurements (BEAM) technique provides a principled method to handle issues currently addressed through rough and varying heuristics, such as identification of changes in expression level, combination of repeated measurements, and rectification of negative measurements. The BEAM technique also produces associated measures of estimation uncertainty (e.g., p-values) which serve to determine the statistical significance of reported results. While our method applies to any gene array technology, I will illustrate it using a detailed noise model that we developed for Affymetrix yeast chips. BEAM can be used to extract additional and more statistically rigorous conclusions from existing data. It also allows for the design of experiments that maximize the useful information derived from a minimum number of chips.

Joint work with Jon Murnick and Nicola Rinaldi (MIT and Whitehead Institute).

CSRI POC: Cindy Phillips, 845-7296

Title: Building space-time meshes over arbitrary spatial domains

Speaker: Jeff Erickson
University of Illinois at Urbana-Champaign

Date/Time: Wednesday, August 14, 2002, 10:30-12:00 noon

Location: Building 980 Room 24

Brief Abstract: We present an algorithm to construct meshes suitable for space-time discontinuous Galerkin finite-element methods. Our method generalizes and improves the `Tent Pitcher' algorithm of UngOr and Sheffer. Given an arbitrary simplicially meshed domain D of any dimension and a time interval [0,T], our algorithm builds a simplicial mesh of the space-time domain D x [0,T], in constant time per element. Our algorithm avoids the limitations of previous methods by carefully adapting the durations of space-time elements to the local quality and feature size of the underlying space mesh.

This is joint work with Damrong Guoy, John Sullivan, and Alper Ungor, to be presented at the 11th International Meshing Roundtable. Our paper is available on the web at arXiv:cs.CG/0206002 and http://www.cs.uiuc.edu/~jeffe/pubs/slowpitch.html.

CSRI POC: David Bunde, 844-2932

Title: Algorithms for Making Quantum Monte Carlo Practical

Speaker: Michael T. Feldmann
California Institute of Technology

Date/Time: Tuesday, February 26, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Quantum Monte Carlo (QMC) is quickly gaining popularity as the high accuracy

particle-correlation correction method of choice in computational chemistry and materials science. QMC has an attractive computational complexity (roughly O(n3) while competing methods are O(n6) and higher) yet the prefactor is often prohibitively expensive. QMC is fairly easily parallelized so very large scale calculations on parallel supercomputers are done to overcome this large prefactor. We examine several algorithms which will begin to make QMC a more cost effective method on current and future supercomputers. These algorithms allow an improved parallelization technique, quick correlated statistics analysis, convergence based termination, and inexpensive correlation function optimization.

CSRI POC: Peggy Aragon, 845-3009

Title: The modeling of particulate media and mult-fracturing solids: computational issues

Speaker: Y.T. Feng, Civil & Computational Engineering Centre
School of Engineering, University of Wales Swansea

Date/Time: Thursday, September 26, 2002, 10:30-11:30 am

Location: Building 980 Room 95

Brief Abstract: This talk will review computational strategies in the context of combined finite/discrete element approaches for effective modeling of large scale practical problems involving multiple fracture and discrete phenomena. The computational issues considered include: 1) Fracture criteria and propagation mechanisms within both the finite and discrete elements, together with mesh adaptivity procedure; 2) detection procedures for monitoring contact between large numbers of discrete objects; 3) interaction laws governing the response of contact pairs; 4) parallel implementation; and 5) other issues.

CSRI POC: Rich Lehoucq, 845-8929

Title: Multiscale Mechanics and Nonlinear Homogenization

Speaker: Professor Jacob Fish
Rensselaer Polytechnic Institute

Date/Time: Monday, May 20 through Friday, May 24, 2002

Location: All sessions will be held in Bldg 980, Room 95 from 9:00-10:30 am, EXCEPT for Tuesday, May 21 when the session will be 1:00-2:30 pm, Video link to SNL/CA

Brief Abstract: Mathematical homogenization theory based on multiple scale expansion method will be derived. Homogenization theory will be then extended to nonlinear problems. Applications to damage, plasticity and fatigue are discussed.

Outline:
1.1 Basic principles of homogenization in 1D
1.2 Mathematical homogenization in multiple dimensions
1.3 Unit cell problem and application of periodic boundary conditions
1.4 Nonlinear homogenization (plasticity and damage)
1.5 Temporal homogenization
1.6 Space-time homogenization with application to impact in heterogeneous medium
1.7 Integrated approach to fatigue

CSRI POC: John Aidun, 844-1209

Title: Stabilizing MINOS

Speaker: Michael Friedlander
Stanford University

Date/Time: Wednesday, April 10, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: For optimization problems with nonlinear constraints, linearly constrained Lagrangian (LCL) methods sequentially minimize a Lagrangian subject to linearized constraints. These methods converge rapidly near a solution but may not be reliable from arbitrary starting points. The well-known example MINOS has proven effective on many large problems. Its success motivates us to propose a globally convergent variant. Our stabilized LCL method possesses two important properties: the subproblems are always feasible, and they may be solved inexactly. These features are present in MINOS only as heuristics.

The new algorithm has been implemented in MATLAB, using the SNOPT Fortran code as the subproblem solver. Only first derivatives are required. We present numerical results on a nonlinear subset from the COPS and CUTE test-problem sets, which include many large examples. The results demonstrate the robustness of the stabilized LCL procedure.

CSRI POC: William Hart, 844-2217

Title: On some Efficient Numerical Algorithm for Metacomputing of CFD applications

Speaker: Professor Marc Garbey
Department of Computer Science, University of Houston

Date/Time: Monday, March 11, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: Metacomputing by definition [1] is the coupling of heterogeneous, computational resources by software in such a way that they can be used as easily as a personal computer. A number of attempts have been made to establish this concept in science and engineering among which the GLOBUS project is the most widely developed one [2]. Experiments with large configurations have however shown that the latency of wide area networks is prohibitively high and that substantial bandwidth cannot easily be achieved [3]. Further varying latencies and bandwidths are no longer only a problem of distributed computing. With the introduction of clusters of fat nodes built from commodity parts all modern hardware architectures have already introduced several levels of memory access speed.

Although a cluster is usually referred to as a single system, in reality it is a heterogeneous system of commodity parts with a widely varying quality and speed of data transfer. Linking such systems together to form a metacomputer does not add a new level of complexity but only extends the cluster concept. A metacomputer can thus be seen as the highest level of clustering. This is not just a question of academic classification of systems. It is rather the key problem of software development for high-end systems for the next years. While most codes were originally developed on homogeneous parallel systems (like the Cray T3E) and tend to ignore the heterogeneity of modern cache based and clustered systems, code that was developed in metacomputing environments naturally takes into account this heterogeneity. In order to give the right guiding principle for software development clustering might therefore rather be seen as a special

case of metacomputing.

Furthermore there are special problems that require a level of compute power that is not available on a single system.

The development of practical algorithms that give good performance on supercomputers linked by slow networks is difficult. For Poisson or Helmholtz operators, it is well known that the speed of propagation of information in the spatial domain is infinite. The situation is similar for parabolic equations. However there are two factors that help to design algorithms for slow communication:

1. Even if the information propagates at infinite speed, it can be still damped in space relatively fast.

2. More than 90 per cent of the information that we carry in a practical computation is somewhat noise.

Based on these two observations, we will present three numerical algorithms that can achieve numerically efficiency even in metacomputing situations with slow network:

• C(p,q,j) scheme for the time integration of system of differential equations [4]: an application is reactive flow computation, with a Navier Stokes code and a Reaction-Diffusion code for the chemistry running on DEC(Compaq (HP)) clusters in Paris and Lyon. The bandwidth in this experiment was 10 Mb/s, which is about 80 times slower than the internal network of a DEC Cluster. We were able to reach a parallel efficiency of 80%

• Aitken like acceleration of the Schwarz method in domain decomposition [5] [6]: this is used to construct a fast Poisson solver on a network of several Grays T3e in the US and Europe linked by ordinary network. This new family of domain decomposition solver has been generalized to non linear elliptic problems. The metacomputing experiments have benefit in an essential manner of an international collaboration with the team of M. Resch in Univ. of Stuttgart and T. Rossi and J. Toivanen in Univ. of Jyvaskyla.

• Stabilization of explicit time stepping for Parabolic problems via filtering [7]: this method is currently developed to speed up reaction-convection diffusion solver for air quality application in metacomputing environment.

CSRI POC: Rich Lehoucq, 845-8929

Title: Error Estimation, Multilevel Method and Robust Extrapolation in the Numerical Solution of PDEs

Speaker: Marc Garbey
Department of Computer Science, University of Houston

Date/Time: Tuesday, March 12, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Code verification in Computational Science may take great advantage of modern developments in applied mathematics such as multilevel methods, a posteriori estimator, anti-optimization technique etc... as well as information technology tools such as grid computing or data mining.

We will focus our presentation on a generalization of Richardson extrapolation method. Richardson extrapolation is based on a very simple and elegant mathematical idea that has been successful in several

areas of numerical analysis such as quadrature or time integration of ODEs. In theory, Richardson extrapolation can be used also on PDE approximations when the convergence order of a discrete solution is clearly known. But in practice, the order of a numerical method often depends on space location and is not accurately satisfied on different levels of grids used in the extrapolation formula. We propose in this presentation a more robust and numerically efficient method based on the idea of finding automatically the order of a method as the solution of a least square minimization problem on the residual. We introduce a two-level and three-level least square extrapolation method that works on non matching embedded grid solutions via spline interpolation. Our least square extrapolation method is a post-processing of data produced by existing PDE codes, that is easy to implement and can be a better tool than Richardson extrapolation for code verification. It can be also used to make a cascade of computation more numerically efficient. We can establish a consistent linear combination of coarser grid solutions to produce a better approximation of the PDE solution at a much lower cost than direct computation on a finer grid. To illustrate the performance of the method, examples including two dimensional turning point problem with sharp transition layer and the Navier-Stokes flow inside a lid-driven cavity are adopted.

This is joint work with W. Shyy of the Department of Aerospace Engineering, Mechanics and Engineering Science, University of Florida.

CSRI POC: Bill Oberkampf, 844-3799

Title: Resource Deployment based on Autonomous System Clustering

Speaker: James Gast,
University of Wisconsin

Date/Time: Friday, December 6, 2002, 11:30 am - 12:15 pm

Location: Building 980 Room 95

Brief Abstract: The discussion touches on the some of the difficult problems associated with mapping and modeling the Internet. It is motivated as a problem in placing a small number of video servers in the Internet and maximizing coverage. Studies of Internet topology are frustrated by the lack of an accurate model of the flow of packets across the graph. Business relationships cause actual paths to be radically different than optimal shortest paths and passive techniques to collect topology information miss some of the most important exchange points and links.

This talk shows how we leveraged prior work to infer business relationships from Border Gateway Protocol and a large collection of targeted traceroutes to generate a concise representation of the Internet at the Autonomous System level. We then use a novel dynamic programming approach to validate that the simple model is appropriate for the server placement problem.

CSRI POC: Erik DeBenedictis, 284-4017

Title: Global Optimization Using Parallel Techniques

Speaker: Ed Gatzke
University of South Carolina

Date/Time: Wednesday, June 26, 2002, 1:00-2:00 pm

Location: Building 980 Room 95

Brief Abstract: This presentation describes a variety of deterministic methods for solution of different types of nonconvex optimization problems. Many optimization problems (involving only continuous variables) include nonconvex algebraic constraints. General nonlinear programming techniques often return local optima and offer no guarantee of convergence to the global solution. A convexification method for general factorable nonlinear constraints is presented. This method creates a Linear Programming (LP) lower bounding problem, which may be used in spatial branch-and-bound algorithms. Variable space reduction methods are also presented. The relationships between branch-and-bound methods for global nonlinear optimization, Mixed-Integer Linear Programming (MILP), and Mixed-Integer Nonlinear Programming (MINLP) are introduced. An alternative MINLP method based on decomposition / outer approximation is presented. This method allows for separate decoupled parallel solution of lower and upper bounding problems (MILP and global NLP, respectively) and provides guarantees for convergence in a finite number of iterations. Other unique parallelization techniques are also presented.

CSRI POC: William Hart, 844-2217

Title: Interior-Point Optimization through Trust Regions

Speaker: E. Michael Gertz
Post-doctoral Researcher, University of Wisconsin, Madison

Date/Time: Monday, September 16, 2002, 10:00-11:00 am (PST)

Location: Building 921 Room 137(Sandia-CA)
Building 980 Room 95 (Sandia-NM)

Brief Abstract: We describe a trust-region approach to non-convex, nonlinear optimization. The method, based on a merit function of Forsgren & Gill, handles non-convexity correctly and has excellent convergence properties. The algorithm uses off-the-shelf linear equation solvers, allowing us to leverage years of research in numerical linear algebra.

CSRI POC: Tamara Kolda, (925) 294-4769

Title: Dynamic Meshes, Dynamic Interfaces and Hemodynamics

Speaker: Omar Ghattas
Carnegie Mellon University

Date/Time: Tuesday, January 22, 2002, 10:00-11:00 am (PST)

Location: Building 921 Room 137 (Sandia-CA)
Building 980 Room 24 (Sandia-NM)

Brief Abstract: Many important phenomena in science and engineering, including our motivating problem of microstructural blood flow, can be modeled as flows with dynamic interfaces. The major challenge faced in simulating such flows is resolving the interfacial motion. Lagrangian methods are ideally suited for such problems, since interfaces are naturally represented and propagated. However, the material description of motion results in dynamic meshes, which become hopelessly distorted unless they are regularly regenerated. Lagrangian methods are particularly challenging on parallel computers, because scalable dynamic mesh methods remain elusive.

I will present a parallel dynamic mesh Lagrangian method for flows with dynamic interfaces that we have been developing at CMU. We take an aggressive approach to dynamic meshing by triangulating the propagating grid points at *every* time step using a scalable parallel Delaunay algorithm. Contrary to

conventional wisdom, I will provide evidence that the costs of the geometric components (triangulation, coarsening, refinement, and partitioning) can be made small relative to the flow solver. For example, in a 2D simulation of 10 interacting viscous cells with 500,000 unknowns on 64 processors of a Cray T3E, dynamic meshing consumes less than 5% of a time step. Moreover, our experiments on up to 128 processors show that the computational geometry scales about as well as the flow solver.

I will discuss the application of our dynamic mesh Lagrangian method to microstructural simulation of blood flow, which is essentially a problem in modeling the interaction of fluid-solid mixtures. The model is termed "microstructural" because it distinguishes the fluid (blood plasma and hemoglobin) from the solid (cell membrane) at micron scales, and computes the momentum exchange between them -- in contradistinction to typical macroscopic models that treat blood as a homogeneous viscous medium with phenomenological incorporation of cellular effects. I will conclude with a discussion of the prospects for microstructural modeling of blood flow at scales of interest in the design of artificial heart devices.

This work is joint with former graduate student Ivan Malcevic (now at GE), CMU colleagues Guy Blelloch, Gary Miller, and Noel Walkington, and University of Pittsburgh Medical Center collaborator Jim Antaki.

SNL Contact: Monica Martinez-Canales, (925) 294-3157

Title: Multiscale Newton-Krylov Methods for Inverse Wave Propagation

Speaker: Omar Ghattas
Carnegie Mellon University

Date/Time: Wednesday, February 27, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Toward our goal of modeling strong earthquakes in seismic regions, we are interested in determining mechanical properties of sedimentary basins (such as the greater Los Angeles Basin) from seismograms of past earthquakes. As an intermediate problem, we consider the inverse wave propagation problem of determining the material properties of a heterogeneous acoustic medium, given a source and observations at receiver locations on its boundary. The inverse problem is formulated as a PDE-constrained optimization problem, in which the objective function is an L2 norm misfit between model and observations, the constraint is the acoustic wave equation with appropriate initial and boundary conditions, and the inversion variable is the density field. Both Tikhonov (H1) and total variation regularization are used to render the inverse problem well-posed. The Euler-Lagrange equations representing first-order optimality yield a coupled three-field system composed of the (forward-in-time) acoustic wave equation, the (backward-in-time) adjoint acoustic wave equation, and the integro-differential inversion equation. Newton solution of an appropriate spatio-temporal discretization of this system for state variables, adjoint variables, and the material field yields a linear system with a large indefinite coefficient matrix. Due to the size (proportional to the product of the number of grid points and time steps) and lack of structure of this system, we invoke a block elimination that produces a Schur complement system in the reduced space of the material field. We solve the Schur complement system using conjugate gradients; these inner iterations are embedded within an inexact Newton outer iteration. The Schur complement is too large and expensive to be formed explicitly; instead, each CG matrix-vector product requires the solution of a forward and an adjoint wave propagation problem. CG converges rapidly for the Tikhonov-regularized Schur complement, and no preconditioning is needed. On the other hand, TV-regularized Schur complement systems are much more difficult to converge, and we are currently investigating various strategies for their solution. To address the difficulties posed by numerous local minima of the objective function, we employ a multiscale algorithm that sweeps through the source frequencies, determining successively higher spectral components of the material field, on successively finer grids. Successive material fields remain in the (increasingly narrowing) basin of attraction of the global minimum. Numerical experiments on synthetic 3D inverse wave propagation problems with up to 0.75M variables on up to 256 Cray T3E processors demonstrate

high parallel and algorithmic scalability, and the ability to capture global minima and recover accurate inversion fields. This work is joint with Dr. Volkan Akcelik, who successfully defended his PhD thesis on this topic last week.

CSRI POC: Bart van Bloemen Waanders, 284-6746

Title: Recent developments in (a) shock physics simulations and b) prediction with quantification of uncertainty

Speaker: James Glimm
Department of Applied Mathematics and Statistics
SUNY at Stony Brook

Date/Time: Monday, March 18, 2002, 9:00-10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Two closely related questions will be addressed: how to get better simulations and how to assess their validity and their impact on predictions. In spite of the obvious link between these topics, we will treat them as independent questions. Results joint with coworkers will be presented.

Our approach to better shock physics simulations has been to track interfaces. Numerical errors associated with material boundaries are troublesome, especially for Eulerian simulations. A tracked interface can be viewed as the ultimate ALE method, with the Lagrangian aspect reduced to the interface itself. Tracking simulations in agreement with fluid mixing experiments will be reviewed. A new development is a fully conservative and higher order tracking method. This method will be explained and preliminary results presented.

We use a Bayesian approach to prediction. We have focused on models to express the solution error, needed to specify the Bayesian likelihood, or in less technical terms, needed to determine how tightly measurements linked to physics parameters can be used to constrain uncertainties in the parameters, and once this is done, how tightly the simulations constrain the prediction. Here we illustrate our ideas by applications to prediction of production from petroleum reservoirs.

CSRI POC: Mark A. Christon, 844-8279

Title: Solution of Non-Symmetric, Real Positive Linear Systems

Speaker: Gene H. Golub
Fletcher Jones Professor of Computer Science Stanford University

Date/Time: Wednesday, February 6, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: The methods we discuss use a Hermitian/skew-Hermitian splitting (HSS) iteration and its inexact variant, the inexact Hermitian/skew-Hermitian splitting (IHSS) iteration, which employs inner iteration processes at each step of the outer HSS iteration. Theoretical analyses show that the HSS method converges unconditionally to the unique solution of the system of linear equations. Moreover, we derive an upper bound of the contraction factor of the HSS iteration which is dependent solely on the spectrum of the Hermitian part. Numerical examples are presented to illustrate the effectiveness of both HSS and IHSS iterations. In addition, a model problem of three-dimensional convection-diffusion equation is used to illustrate the advantages of our methods.

CSRI POC: Rich Lehoucq, 845-8929

Title: Theory and Applications of the Polynomial Numerical Hull

Speaker: Professor Anne Greenbaum
University of Washington

Date/Time: Thursday, August 22, 2002, 1:00-2:00 pm (PST)

Location: Building 921 Room 137 – Sandia CA
Building 980 Room 24 – Sandia NM (Videolinked)

Brief Abstract: The asymptotic behavior of a matrix or linear operator is governed by its eigenvalues, but, in general, the transient behavior is not. Yet this transient behavior is often important in physical applications such as convection and fluid flow and in numerical algorithms such as finite difference schemes and iterative methods for solving linear systems. In this talk we consider sets associated with a matrix A that are designed to give more information than the spectrum alone can provide about this transient behavior. Specifically, we discuss the *polynomial numerical hull of degree k*, which is defined as the set of points z in the complex plane for which $|p(A)| >= |p(z)|$ for all polynomials p of degree k or less. Geometrical properties and methods for computing these sets are discussed. As a sample application, we show how these sets can be used to predict cutoff phenomena in Markov chains, such as the much-publicized result that it takes seven riffle shuffles to randomize a deck of cards.

CSRI POC: Victoria Howle, (925) 294-2204

Title: Reduced-order modeling for the simulation, control, design, and Optimization of complex systems

Speaker: Max Gunzburger
Florida State University

Date/Time: Tuesday, December 10, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: Standard discretization techniques (e.g., finite volume and finite element methods) for the computational simulation of complex systems such as the Navier-Stokes equations of fluid mechanics lead to very large discrete systems which are very expensive to deal with regards to both storage and CPU costs. The high costs of computational simulations often makes it impossible to treat design, optimization, and control applications that invariably require multiple and/or real time direct simulations. For these and other reasons, there is much interest in developing reduced-order models, i.e., models involving very few degrees of freedom, for the computational simulation of complex systems. We discuss the necessity for reduced-order models and, in broad terms, when they can be expected to be effective. We then discuss some of the many existing reduced-order models, focusing on the Lagrange and Taylor reduced-basis methods and on the currently popular proper orthogonal decomposition (POD) approach which is equivalent to using well-known statistical techniques. We provide examples, drawn from computational fluid dynamics applications, of successes and failures resulting from the use of these approaches. We then present a new approach which is based on centroidal Voronoi tessellations (CVT) which are also equivalent to applying well-known statistical clustering techniques. We compare CVT with POD with respect to both cost and effectiveness and provide some preliminary computational examples of the use of the CVT-based approach for reduced-order modeling applications.

CSRI POC: Rich Lehoucq, 845-8929

Title: Support Vector Machines and Lung Cancer Classification using array Comparative Genomic Hybridization

Speaker: Doug Hardin
Vanderbilt University

Date/Time: Tuesday, April 10, 2002, 2:30-3:30 pm

Location: Building 980 Room 24

Brief Abstract: Support Vector Machines (SVMs) are a recently developed class of kernel-based machine learning algorithms. Array comparative genomic hybridization (array CGH) is a recently introduced technology that measures gene copy number changes of hundreds of genes in a single experiment. This talk will include a brief introduction to SVMs as well as a description of preliminary work (with C. Aliferis and P. Massion) on using SVMs for the classification of non-small cell lung cancers from array CGH data. We compare with a several other machine learning methods including KNN and Feed-Forward Neural Networks.

CSRI POC: David Womble, 845-7471

Title: Computer Simulations of the Selectivity and Permeation of Ions in Ion Channels in Physiological Membranes

Speaker: Doug Henderson
Brigham Young University

Date/Time: Tuesday, December 17, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Channels are proteins that contain a hole through which an ion can pass through an otherwise impermeable membrane. They may deliver a signal to a cell that can control a biological function. For example, a calcium channel, which selectively passes calcium ions, sends signals to the heart muscles. This selectivity can be quite dramatic; a calcium channel can select calcium ions from a solution in which the calcium ions are present only in extreme dilution. Recent computer simulations of the selectivity and permeation of ions in gramicidin-like, calcium, and sodium channels are discussed for cases where the water molecules are modeled as a dielectric continuum and as a set of discrete polar molecules.

CSRI POC: Laura Frink, 844-1910

Title: Multiple Parameter Continuation/Computing Implicitly Defined Manifolds

Speaker: Michael E. Henderson
IBM Research T. J. Watson Research Center

Date/Time: Tuesday, April 23, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Many structures of interest in Applied Mathematics can be written as the solution of a nonlinear system with a small number of parameters. Geometrically these are pieces of low dimensional manifolds embedded in spaces of dimension up to around 100 (algebraic systems), in the hundreds to thousands (periodic motions, homoclinic and heteroclinic trajectories in dynamical systems), and on up into the millions (PDE's). We would like to be able to compute these structures over some range of parameter values.

I will describe a predictor-corrector continuation method for computing implicitly defined manifolds, and show examples of computations of equilibrium configurations of a clamped twisted rod (a simplified model of DNA) - 15,000 points in R^604, and periodic motions of a pair of spring-coupled driven pendula - 12,500 points in R^203.

Instead of using a mesh, I represent the manifold as the union of overlapping neighborhoods, each centered at a point on the manifold. To extend the manifold, a new point is chosen from the boundary of this union and a neighborhood of it is added to the collection. I will show that when the neighborhoods are spherical, or nearly spherical, the boundary can be expressed very simply in terms of a certain set of convex polyhedra.
Furthermore, a boundary point can be found in terms of the vertices of these polyhedra . The resulting algorithm is very simple, and since the new point need only be near the boundary, surprisingly robust.

There are strong parallels between the continuation and mesh generation on surfaces. The usual continuation methods are roughly equivalent to generating a mesh on an isosurface with marching cubes, and advancing front
methods. The algorithm described in this talk advances a Laguerre Voronoi
diagram by adding new points on infinite edges.

CSRI POC: Andy Salinger, 845-3523

Title: Fictitious Domain Decomposition Methods for a Class of Partially Axisymmetric Problems: Application to the Scattering of Acoustic Waves

Speaker: Ulrich Hetmaniuk
Center for Aerospace Structures
University of Colorado/Boulder

Date/Time: Thursday, August 8, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Many scatterers, particularly in aerospace and military applications, are neither fully axisymmetric, nor arbitrarily shaped. Rather, they are obtained by assembling one or two major axisymmetric components and a few features. Examples include rockets, satellites, guided missiles, and submarines. For such scatterers, an axisymmetric acoustic scattering analysis is not applicable. On the other hand, a straight-forward three-dimensional (3D) analysis is inefficient because the geometrical properties of the axisymmetric components are not exploited.

In the mid-frequency regime, a 3D acoustic scattering analysis may also be unfeasible when the finite element discretization of the exterior Helmholtz problem requires hundreds of millions of grid points.

My proposed methodology exploits such local axisymmetry with a fictitious domain approach. The original computational domain is embedded into an axisymmetric domain. A Helmholtz problem is then formulated inside this domain as well as an auxiliary Helmholtz problem inside each feature. Lagrange multipliers are introduced at the interfaces between the features and the axisymmetric computational domain to enforce a set of carefully constructed constraints. Consequently, this fictitious domain decomposition method transforms the original large-scale 3D problem into a set of two-dimensional Fourier problems connected

by Lagrange multipliers to another set of small-scale auxiliary problems associated with the features of the scatterer.

It is proved that the solution of the constrained problem with Lagrange multipliers is equivalent to the solution of the original problem. The algebraic system is solved directly or with the FETI-H iterative algorithm. Some large-scale computations are reported that highlight the intrinsic parallelism of the methodology and illustrate a reduction in the solution time and memory requirements by more than an order of magnitude.

In my presentation, the error analysis for the Fourier finite element method applied to the Helmholtz equation will also be discussed. This analysis is challenging because the Fourier coefficients belong to weighted Sobolev spaces and because the sesquilinear form for the Helmholtz equation is not elliptic. To the best of my knowledge, these error estimates constitute a new result.

CSRI POC: Richard Lehoucq, 845-8929

Title: Low Stretch Object Location in Wide Area Networks

Speaker: Kris Hildrum
UC Berkeley

Date/Time: Thursday, September 12, 2002, 10:00-11:00 am

Location: Building 980 Room 24

Brief Abstract: Modern networking applications replicate data and services widely, leading to a need for location-independent routing -- the ability to route queries directly to objects using names independent of the objects' physical locations. I will mention one such application, OceanStore, and then talk about how the routing structure it requires can be built.

Two important properties of a routing infrastructure are routing locality and rapid adaptation to arriving and departing nodes. We show how these two properties can be efficiently achieved for certain network topologies. To do this, we present a new distributed algorithm that can solve the nearest-neighbor problem for these networks. We describe our solution in the context of Tapestry, an overlay network infrastructure that employs techniques proposed by Plaxton, Rajaraman, and Richa

CSRI POC: Bruce Hendrickson, 845-7599

Title: Model Validation Methodology: From Validation Experiments to Systems Level Application

Speaker: Professor Richard Hills
New Mexico State University

Date/Time: Monday, June 17, 2002, 10:00 - 11:00 am

Location: Building 980 Room 95

Brief Abstract: Our increased dependence on computer models leads to the natural question - how do we validate a model against experimental data? Models have traditionally been tested against experimental measurements through simple comparisons such as x-y plots, scatter plots, or contour plots. While such qualitative comparisons are appropriate for model building, the use of such comparisons for model validation naturally lead to the questions: when is the agreement between experimental measurements and

model predictions sufficient, and how should we quantify this agreement? Unfortunately, defining rigorous metrics for such comparisons is difficult since there are always uncertainties in the validation experiment measurements and in the model parameters. Because of these uncertainties, we expect there to be differences between experimental observations and model predictions, even for perfect models. In addition, when models produce multivariate data (time histories for example), the differences between the experimental observations and model predictions can be highly correlated, which must be accounted for. Furthermore, we often measure one quantity from validation experiments, but desire to predict another quantity (decision variable) for the target application of our model. Finally, complex multi physics models often require a suite of validation experiments to test the model over the range of parameters and physics addressed by the target application.

CSRI POC: Timothy G. Trucano, 844-8812

Title: Automatic Differentiation: Tools and Techniques

Speaker: Paul Hovland
Argonne National Lab

Date/Time: Wednesday, July 17, 2002, 2:30-3:30 pm

Location: Building 980 Room 95

Brief Abstract: We provide an overview of automatic differentiation tools for derivative computation, discussing their current capabilities and probable advances in the coming years. We discuss the use of automatic differentiation in conjunction with toolkits for scientific computing, with particular emphasis on some of the challenges encountered and possible mechanisms for overcoming these challenges. We conclude with a survey of applications and empirical observations.

CSRI POC: Andy Salinger, 845-3523

Title: Recent Developments in Computational Dynamics Integration Methodologies

Speaker: Gregory Hulbert
University of Michigan

Date/Time: Monday, April 29, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Despite the seeming maturity of the subject, the past five years has seen resurgence in research efforts towards developing improved numerical methods for solving problems in computational dynamics.

The research thrusts have been threefold: (1) development of general variationally-based frameworks for time integration, (2) new time integration methods with enhanced stability, accuracy, efficiency, and numerical dissipation properties, (3) new frameworks and algorithms for so-called co-simulation.

In this talk, I will present an overview of these new developments in computational dynamics integration, including research conducted not only in the computational dynamics laboratory at the University of Michigan, but also other efforts worldwide.

CSRI POC: Mark Christon, 844-8279

Title: Role of Error Quantification in Provably Convergent Optimization under Uncertainty (OUU)

Speaker: Takeru Igusa
Professor of Civil Engineering
Johns Hopkins University

Date/Time: Monday August 12, 2002, 1:30-2:30 pm

Location: Building 980 Room 95

Brief Abstract: Trust region optimization algorithms are provably convergent provided that exact values for the objective function and its gradient are available. When the objective function is in terms of uncertain variables, however, then only variable fidelity estimates for the objective function can be computed. While stochastic approximation algorithms (based on Robbins-Monro or Kiefer-Wolfowitz) are provably convergent for this OUU problem, they share some limitations which often preclude their use in large-scale problems.

In this talk, we examine the conditions under which trust region algorithms, based on variable fidelity estimates of the objective function, are provably convergent. It is shown that these conditions are inseparably tied to the error quantification problem. For sake of clarity, only the most basic mathematical constructs, such as Cauchy sequences, mean value theorems, and second moment characterization of the errors, are used to develop the analytical conditions for provable convergence.

We end the talk by exploring sampling strategies and surrogate models which can provide the error estimates needed in the above algorithm.

CSRI POC: Steve Wojtkiewicz, 284-5482

Title: Ca2+ sparks, Ca2+ waves, Supercomputers, and the Intracellular Origins of Atrial Arrhythmias

Speaker: Leighton Izu
University of Maryland Medical School

Date/Time: Monday, August 5, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: With every heartbeat two separate, but interlinked, dynamical systems in the heart are brought into play: the electrical system of the sarcolemma and the intracellular calcium (Ca2+) control system. In a normal cardiac atrial cell, the Ca2+ control system is slave to the electrical system. That is, sarcolemmal electrical depolarization triggers a traveling wave of Ca2+ release. However, this normal sequence can be broken when Ca2+ waves spontaneously occur. When this occurs the released Ca2+ can alter the sarcolemmal electrical activity and derange the timing of the action potential, and may initiate life-threatening arrhythmias.

This talk describes modeling of the Ca2+ control system in cardiac atrial cells. The Ca2+ control system spans many time and space scales so modeling it poses a computational challenge. I will talk about the molecular properties of the channels that control Ca2+ release and how they affect a Ca2+ spark, which is the fundamental unit of Ca2+ release. We will examine how properties of the spark and cellular architecture determine the probability that the Ca2+ sparks will self-organize into a traveling Ca2+ wave.

Finally, I will discuss our future goal of developing a comprehensive mathematical description of an atrial cell that links the sarcolemmal electrical system and the intracellular Ca2+ control system. This will give us a window into the intracellular basis of cardiac arrhythmias.

CSRI POC: Shawn Means, 844-1699

Title: Effect of Confinement on DNA Dynamics in Microfluidic Devices

Speaker: Richard Jendrejack, Ph.D. candidate
University of Wisconsin, Madison

Date/Time: Wednesday, June 5, 2002, 1:30-2:30 pm

Location: Building 980 Room 95

Brief Abstract: The dynamics of dissolved long-chain macromolecules are very different in highly confined environments, such as microfluidic devices, than in bulk solution. I will begin by presenting a Langevin model of DNA, and demonstrate that the model produces results which are in quantitative agreement with available equilibrium and non-equilibrium experimental data in free-solution. Having validated the model in free-solution, I then present a method for the simulation of macromolecular dynamics in microfluidic geometries. The method self-consistently combines a coarse-grained Langevin model of the polymer dynamics with a numerical solution for the flow generated by the motion of polymer segments. With the method, we examine the behavior of DNA in micron-scale channels. Relaxation time and diffusivity are shown to depend strongly on confinement through hydrodynamic interactions, with a very broad crossover from free-solution behavior. Static equilibrium properties calculated from these dynamic simulations compare well with scaling predictions and previous Monte Carlo simulations.

CSRI POC: Steve Plimpton, 845-7873

Title: Micromechanics of Materials with Pores/Microcracks

Speaker: Mark Kachanov
Tufts University

Date/Time: Wednesday, November 13, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: The presentation focuses on micromechanics of materials with anisotropic heterogeneous microstructures that contain pores and cracks of various shapes and orientations. The main problems of interest are:

1. Microstructure-property relations for anisotropic microstructures;
2. Elastic-conductive cross-property relations;
3. Recovery of information on microstructures from the effective properties (the inverse of the first problem).

In order to address these problems, we first determine the proper parameters to characterize concentration of inhomogeneities. Effective elastic and conductive (thermal and electric) properties are expressed in terms of these parameters. Then, explicit relations between elastic and conductive properties are established. This connection is invaluable for understanding microstructure-property relations. In addition, it can serve as a basis for non-destructive quality control, for example, monitoring the loss of stiffness by

measuring electric resistivity. Another possible application of the connection between elastic and conductive properties is microstructure optimization for combined elastic/conductive performance.

Having established connections between microstructure and effective properties (elastic and conductive) and between several sets of effective properties, we go a step further and pose a question: "What information about the underlying microstructure can / cannot be recovered from the effective properties?" This question is particularly important because
it is much easier to measure effective properties, especially conductive, than to make microstructure measurements. We discuss in detail the extent of uncertainty in the information recovery and its reduction if additional information (for example, defect shapes) becomes available. The developed methodology is general and applies to a wide range of materials, both naturally occurring and man-made. As an example, the application of this methodology to thermal barrier coatings (plasma-sprayed and PVD) will be presented.

CSRI POC: Michael Braginsky, 284-6766

Title: Algorithmic Structure of FEMLAB

Speaker: David Kan (Comsol, Inc.)

Date/Time: Wednesday, October 16, 2002, 10:30-11:00 am

Location: Building 980 Room 95

Brief Abstract: In this talk, we will consider the algorithmic structure of FEMLAB, a multiphysics modeling tool based on partial differential equations (PDEs) using the finite element method (FEM). Beginning with the equation forms implemented in FEMLAB, we will take a tour through the process of setting up equations and boundary conditions, mesh generation, symbolic manipulation, matrix assembly, solver types, and postprocessing. Although some motivating examples will be presented, the seminar will focus on the computational issues and challenges confronted in designing a general, flexible PDE modeling environment. Ample time for questions and answers will be available.

CSRI POC: David Ropp, 845-7431

Title: Nonlinearly Consistent Approximations for Multiple Time Scale Systems

Speaker: Dana A. Knoll
Theoretical Division
Los Alamos National Laboratory

Date/Time: Friday, March 8, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: The effect of various numerical approximations in nonlinear multiple-time-scale problems is studied in the framework of the modified equation analysis (MEA) technique. First, MEA is used to study the effect of linearization and splitting in a simple nonlinear ordinary differential equation (ODE) and linear partial differential equation (PDE). The ODE and PDE are differenced in time in various ways and the resulting truncation errors are compared analytically and numerically. It is demonstrated quantitatively how both linearization and splitting result in accuracy degradation when timesteps larger than competing (fast) time scales are employed. Linearization is particularly dangerous when sharply nonlinear coefficients are present, since it may result in the stalling of the time convergence rate. In contrast, a nonlinear method shows excellent time convergence properties allowing the use of timesteps

orders of magnitude larger without accuracy degradation. These findings have also been confirmed in more realistic applications. Specifically, the various time differencing schemes have been applied to a nonlinear, coupled radiation-diffusion problem, to a 2D magnetohydrodynamics problem, and to the 2D shallow water equations. While timesteps of the order of the dynamical time scale without accuracy degradation are possible with the nonlinearly consistent differencing scheme, linearized and/or split difference schemes require timesteps of the order of the fast time scales in the system for equal accuracy.

CSRI POC: John N. Shadid, 845-7876

Title: Towards an Optimally Preconditioned Eigensolver

Speaker: Professor Andrew Knyazev
University of Colorado, Denver

Date/Time: Thursday, October 17, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: Many applied and engineering problems, e.g., in structural mechanics, design of nuclear reactors, ocean modeling, and quantum chemistry, lead, after simplifications and an appropriate approximation of original partial differential equations, to extremely large and ill-conditioned linear systems with symmetric positive definite matrices of coefficients and similar symmetric eigenvalue problems. The preconditioned conjugate gradient method became the standard solver for such linear systems. Our ultimate goal is developing an analogous optimal method for eigenproblems. Ideally, we want to be able to compute an eigenvector of interest at the same cost as that of solving a linear system of equations, using the same preconditioner. That would allow, in particular, a simple adaptation for eigenproblems of available domain decomposition based and multigrid preconditioners for linear systems.

Searching for the optimal eigensolver, we describe the Locally Optimal Block Preconditioned Conjugate Gradient (LOBPCG) Method for symmetric eigenvalue problems, based on a local Rayleigh-Ritz optimization of a three-term recurrence. LOBPCG can be viewed as a nonlinear conjugate gradient method of minimization of the Rayleigh quotient, which takes advantage of the optimality of the Rayleigh-Ritz procedure.

Numerical results establish that our LOBPCG Method may practically be as efficient as the best possible algorithm on the whole class of preconditioned eigensolvers. We discuss several competitors, namely, some inner-outer iterative preconditioned eigensolvers. Direct numerical comparisons with two of them, the inexact Jacobi-Davidson methods JDCG and JDQR, show that our LOBPCG method is faster.

CSRI POC: Richard Lehoucq, 845-8929

Title: Future of Computing

Speaker: Peter Kogge
University of Notre Dame

Date/Time: Tuesday, October 29, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: This talk focuses on building a coherent picture of where technology may lead us in the area of high end computing over the next decade or so. Because many of the problems we now face are a direct result of design decisions made so long ago that they are never even questioned today, we will start

with a tour of what computing really is, and how did we manage to get to the current set of assumptions about its implementation. From that we will go through some forecasts for the "business as usual" approach, and then look at a range of emerging options that have at least the potential to "upset the applecart."

CSRI POC: Erik DeBenedictis, 284-4017

Title: The regularity of the free boundary in a two-dimensional optimal compliance problem

Speaker: Chris Larsen
Worcester Polytechnic Institute

Date/Time: Tuesday, October 15, 2002, 9:30-10:30 am

Location: Building 980 Room 95

Brief Abstract: A well-known problem in optimal design is the optimal compliance problem. The situation is this: an object is fixed on part of its boundary and a force is applied to another part. The problem is to design the object that best resists this force. Some constraint or penalty has to be added, since usually the larger the object, the lower the compliance. Typically the size of the object is constrained or penalized, the idea being to build the best shape for a given weight.

In general, this problem has no solution. As was observed by Kohn and Strang, minimizing sequences of sets tend to homogenize and form complex mixtures of void and solid. One way around this consists of penalizing the perimeter of the unknown set. This approach was studied in collaboration with A. Chambolle, in dimension two, where we proved existence and C^{∞} regularity of the boundaries away from where they are pinned and forces are applied. In this talk, I will outline the proofs of these results.

CSRI POC: Bob Carr, 845-8562

Title: Benchmarking Overhead and Overlap in Cluster Computing

Speaker: Bill Lawry,
University of New Mexico

Date/Time: Thursday, July 25, 2002, 1:00-2:00 pm

Location: Building 980 Room 24

Brief Abstract: While modern cluster computing offers the benefits of parallel computation, the communication tasks can hinder computation. For example, if a node's network component interrupts the OS to handle incoming messages then communication will compete with the application for host CPU cycles. Similarly, if the network component requires application intervention in order for two nodes to complete message synchronization then overlap of computation with message payload transfer may be lost. Significant performance gains are possible in cluster computing by minimizing overhead while maximizing overlap.

I have co-developed a benchmark ported through MPI to detect communication overhead, concurrency, as well as blockages in cluster communication systems. COMB (the Communication Offload MPI-based Benchmark) uses two measurement methods and both use pair-wise communication. The first method uses a relatively artificial communication pattern in order to gage the communication overhead as experienced by an application during peak performance of the network. The other method uses a more realistic pair-

wise communication pattern which checks for overlap based on a minimum of non-blocking MPI calls followed by computation. While these methods do not model the complex communication patterns in actual applications, they do provide detailed and precise characterizations of system performance which has been useful to system analysts.

CSRI POC: Rolf Riesen, 845-7363

Title: A Multigrid Method for the Optimization of Systems Governed by Differential Equations

Speaker: Robert Michael Lewis
College of William & Mary
Department of Mathematics

Date/Time: Thursday, August 1, 2002, 10:00-11:00 am (PST)

Location: Bldg. 921, Rm. 137(Sandia-CA)
Bldg. 980, Rm. 95 (Sandia-NM)

Brief Abstract: We discuss a multigrid approach to the optimization of systems governed by differential equations. Using several model problems as illustrations, we present analytical and numerical results that suggest why, in many cases, a multigrid method would be effective for such optimization problems. Interestingly, multigrid methods may be effective for the solution of the optimization problem even when multigrid is not appropriate for the solution of the governing differential equation.

To try to explain how multigrid works for optimization, we will examine some of the effects of locality and non-locality in frequency and space in the optimization of systems governed by differential equations. As we describe, a central role is played by the qualitative nature of the Hessians of the objective and constraints for such problems, and the fact that the behavior of the system is governed by a differential equation.
This is joint work with Stephen Nash of George Mason University.

CSRI POC: Tamara Kolda, (925) 294-4769

Title: Motif-Based Mining of Protein Sequences

Speaker: Agatha Liu
Washington University

Date/Time: Tuesday, September 17, 2002, 1:30-2:30 pm

Location: Building 980 Room 95

Brief Abstract: We introduce CASTOR, an automatic, unsupervised system for protein motif discovery and classification. Given amino acid sequences for a group of proteins, CASTOR identifies statistically significant motifs and constructs a classification of the proteins based on the presence of these motifs by performing motif discovery and refinement in a top-down and recursive manner. The members of each class are likely to share a function, and the motifs associated with the class are likely to be related to the function.

We evaluate CASTOR's performance on the G protein-coupled receptor (GPCR) superfamily. The results show that the classification constructed by CASTOR is in better agreement with a manually curated classification than one constructed by another automatic, unsupervised system that clusters proteins based

on pairwise, global sequence similarity in a bottom-up manner. Furthermore, while manually constructed classifications tend to be strictly hierarchical, the non-hierarchical ones constructed by CASTOR suggest that complex, unusual functional relationships among classes may be more abundant than expected.

We also apply CASTOR to the mammalian olfactory receptor family for which very little functional information is available. Upon the application of CASTOR, we infer the potential functional roles associated with the generated motifs and classes by integrating various complex data from other sources, such as mutation experiments and ligand binding assays. Among other functional insights gained, we obtain results that support previous hypotheses on structural integrity and post-translational modification. We also propose and provide evidence for a combinatorial molecular mechanism that supports and potentially explains the complex ligand binding behavior. We additionally define sub-sequences that capture structural features of these receptors and study the motifs present in the sub-sequences.

CSRI POC: Daniel Mark Rintoul, 844-9592

Title: Towards Large-Eddy Simulation of Turbulent Flows in Complex Geometries

Speaker: Professor Mahesh Krishnan
University of Minnesota

Date/Time: Friday, December 6, 2002, 9:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Large-eddy simulation (LES) is a computational approach where one spatially filters the Navier-Stokes equations and directly computes the energy-containing scales of motion, while modeling the effects of the unresolved scales. LES has traditionally been restricted to fairly simple geometries; computation of engineering flows has relied largely on the Reynolds-averaged Navier-Stokes equations. We will discuss a computational approach to LES of turbulent flows on unstructured grids. Our objective is to reliably perform LES at high Reynolds numbers in engineering geometries. A novelty of our approach is that it is discretely energy conserving for grids composed of arbitrary computational elements. This makes it robust at high Reynolds numbers without the use of numerical dissipation. The formulation has been implemented for parallel platforms and has been successfully run on as many as thousand processors. Results will be presented for a range of flows from isotropic turbulence to flow in an engineering gas-turbine combustor.

CSRI POC: Stefan Domino, 284-4317
Sheldon Tieszen, Org, 09132, 844-6526

Title: Iterative methods for fluid-structure acoustics

Speaker: Jan Mandel
University of Colorado

Date/Time: Thursday, August 1, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: We will present two iterative methods for the time harmonic fluid-structure acoustic problem in fluid pressure/solid displacement formulation, discretized by finite elements. The first method is and extension of the FETI-H substructuring method. This method uses Lagrange multipliers at substructure interfaces and new duplicated degrees of freedom at the wet interface. The original degrees of freedom are then eliminated and the reduced system is solved by GMRES. Proper scaling of the coupled

system reveals that the fluid and the solid are coupled only weakly, and it is important for the convergence of the iterations. Substructure resonance is prevented by adding artificial complex terms on substructure interfaces, which do change the converged solution. The iterations are further stabilized by a coarse space built from plane waves or eigenvectors of the system matrix. This method is under development with Charbel Farhat and Radek Tezaur. The second method is multigrid with smoothing by GMRES, studied in the thesis of author's student Mirela Popa.

CSRI POC: David Day, 844-1868

Title: Comparison of Galerkin and Control Volume Finite Element Schemes for Advection-diffusion Problems

Speaker: Mario J. Martinez
Multiphase Transport Processes Department (9114)

Date/Time: Monday, August 5, 2002, 10:00-10:45am

Location: Building 980 Room 95

Brief Abstract: The intent of developers of the control volume finite element method (CVFEM) was to produce a locally conservative method while retaining the unstructured grid and generality of finite element methods (FEMs). Similarily, the streamline upwind control volume (SUCV) method is a control volume version of the streamline upwind Petrov-Galerkin (SUPG) method. These discretization methods are in current use or are candidates for new applications codes at SNL. We review the formulation of the CVFEM/SUCV, its conservation properties, and computer implementation. GFEM/SUPG and CVFEM/SUCV are compared via numerical error analysis in the context of advection-diffusion problems with respect to accuracy, conservation, degree of positivity, and convergence rate.

CSRI POC: John Shadid, 845-7876

Title: Improving the accuracy of density functional theory based calculations

Speaker: Ann E. Mattsson, SNL

Date/Time: Wednesday, December 18, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: When experimental data is unavailable or inconclusive, quantum mechanical calculations frequently are the best source of information on a physical system. Density Functional Theory (DFT) is a formally exact reformulation of the basic quantum mechanical Schrodinger Equation (SE) for the ground state of a system. However, to go from the computationally intractable SE wavefunction in 3N-variables to the 3-variables DFT density, a price has to be paid: the form of the exchange-correlation functional in DFT is unknown and must be approximated. The approximation used for this functional sets the fundamental limit on the accuracy of DFT-based calculations [1]. I will present two cases, the Pd(111)/α-alumina work of adhesion [2] and Al, Pt, Pd, and Mo vacancy formation energies [3,4], where the presently available functionals are not good enough to be predictive. I will discuss the origin of the errors and a procedure to correct for them *a posteriori*. Despite the success of this correction procedure, it is desirable to remove the errors at the source by improving the approximation for the exchange-correlation functional. How to do this is the topic of the second part of my talk. I will present the background and motivation for 'subsystem functionals' [5] and our recent findings in this area.

[1] A.E. Mattsson, Science **298**, 759 (25 October 2002).

[2] A.E. Mattsson and D.R. Jennison, Surf. Sci. Lett. **520**, L611 (2002).
[3] K. Carling et al., Phys. Rev. Lett. **85**, 3862 (2000).
[4] T.R. Mattsson and A.E. Mattsson, Phys. Rev. B **66**, 2141xx (Dec. 2002).
[5] R. Armiento and A.E. Mattsson, Phys. Rev. B **66**, 165117 (2002).

CSRI POC: John Aidun, 844-1209

Title: Compiler Optimization Seminar

Speaker: Michael J. McClosky
Perdue University

Date/Time: Wednesday, March 20, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: The talk will proceed according to the following outline

- Introduction
- Optimization Requirements
- Definitions
- Local (Basic Block) Optimizations
- Global Optimizations
- Loop Specific Optimizations
- Vector and Parallel Processing Optimizations
- Compiler Optimizations and High Speed Computing
- Compiler Optimizations and the Sandia Advanced Strategic Computer Initiative (ASCI)
- Summary
- Questions

CSRI POC: Erik DeBenedictis, 284-4017

Title: Tight-Binding for Real Materials

Speaker: Michael J. Mehl
Naval Research Laboratory

Date/Time: Thursday, February 28, 2002, 10:00-1:00 am

Location: Building 980 Room 95

Brief Abstract: Accurate first-principles quantum mechanical calculations for real materials are computationally limited to no more than 100-1000 atoms. At larger scales, atomistic potentials such as the Embedded Atom Method are used, but these methods may miss important physics driven by changes in the electronic structure, e.g., at cracks and defects. Parametrized tight-binding (TB) methods exist between these two extremes. Unlike atomistic potentials, the quantum mechanical behavior of the electrons is maintained, but the computational effort is much less than needed for comparably sized first-principles calculations. This talk describes the NRL Tight-Binding Method (NRL-TB), which maps the results of a limited set of first-principles calculations to a two-center non-orthogonal Slater-Koster TB Hamiltonian. The on-site Hamiltonian parameters are sensitive to the local environment and the hopping parameters are bond-length dependent. The method has been shown to successfully determine elastic constants, phonon frequencies, vacancy formation energies, and surface energies. In addition, TB molecular dynamics simulations are used to study thermal expansion and atomic diffusion.

We will discuss applications to spin-polarized systems, non-collinear magnetization, and multi-component systems, including MgB2.

CSRI POC: Peter Schultz, 845-7771

Title: Recent Developments in MATLAB

Speaker: Cleve Moler
The MathWorks (distributor of MATLAB)

Date/Time: Wednesday, January 16, 2002, 10:00-11:00 am (PST)

Location: Building 921 Room 137 (Sandia-CA)
Building 980 Room 24 (Sandia-NM)

Brief Abstract: MATLAB is a widely used software package with a GUI interface for development, testing, and visualization of mathematical algorithms and results. In this talk, we will discuss recent developments in MATLAB. Come find out what is new in MATLAB and what you new things you can do with it!

SNL Contact: Monica Martinez-Canales, (925) 294-3157

Title: I/O projects in the Unix I/O subsystem to deliver digital streaming video product

Speaker: Larry Morkre
Sun Microsystem

Date/Time: Monday, April 22, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Digital streaming video products have some of the toughest requirements on the I/O subsystem of any product group. They require not only high throughput but low patency. Any performance mishap will get reflected immediate in the video or in the audio quality. This is particularly true in the case of compressed video. This presentation will discuss some I/O projects that I participated in during my work at Pluto Technologies. One topic will discuss the Fibre Channel transporter software we developed to transport video data over FC using our own protocol. Another topic will be a list I/O driver that handled multiple I/Os with one system call. Further topics would include discussing using CMU's RAID frame software which allows user level development of kernel level drivers.

CSRI POC: Lee Ward, (240) 453-3893

Title: A Study in Support Vector Machines

Speaker: Todd S. Munson
Argonne National Laboratory

Date/Time: Monday, August 19, 2002, 10 a.m. (PST)

Location: Building. 921, Rm. 137 - Sandia-CA
Building 980 Room 95 – Sandia-NM (Videolinked)

Brief Abstract: We discuss interior-point and semismooth methods for solving quadratic programming problems with a small number of linear constraints where the quadratic term consists of a low-rank update to a positive semi-definite matrix. Several formulations of the support vector machine, a technique employed by the machine learning community for supervised learning, fit into this category. A related example is the Huber regression problem, which can also be posed as a quadratic program with the desired properties.

Support vector machines can be used, for example, when determining whether a tumor is malignant or benign. An interesting feature of the problems we consider is the volume of data, which can lead to quadratic programs with between 5 and 100 million variables and, if written explicitly, a dense Q matrix. The implementations of the two algorithms use linear algebra specialized for the support vector machine application. For the targeted massive problems, all data is stored out-of-core and we overlap computation and I/O to reduce overhead. Results are reported for several linear support vector machine formulations demonstrating that the algorithms developed are reliable and scalable.

CSRI POC: Tamara Kolda (925) 294-4769

Title: Computing Mountain Passes

Speaker: Todd Munson
Argonne National Laboratory

Date/Time: Thursday, November 7, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: We propose the elastic string algorithm for computing mountain passes in finite-dimensional problems. We analyze the convergence properties and numerical performance of this algorithm for benchmark problems in chemistry and discretizations for infinite-dimensional variational problems. We show that any limit point of the elastic string algorithm is a path that crosses a critical point at which the Hessian matrix is not positive definite.

CSRI POC: Patrick Knupp, 284-4565

Title: KAI's DVSM Run-Time Library

Speaker: Jeffrey Olivier

Date/Time: Monday, September 9, 2002, 9:45-10:30 am

Location: Building 980 Room 95

Brief Abstract: KAI's DVSM run time library, particularly the underlying memory consistency system which is based on Treadmarks. This will include a short background of Treadmarks as well as our contributions over this summer and our future goals and direction. These include enlarging the shared heap constraints, increasing the number of nodes, graceful shutdown in the event of failures, added support for message-based shared memory using one-sided get, and scalability analysis. Will give an overview of our results as well as the theoretical limits of the system.

CSRI POC: William McLendon, 845-0649

Title: Recent Developments in Derivative-free Algorithms for Non Smooth Unconstrained Minimization

Speaker: Dr. Ubaldo M. Garcia-Palomares
Universidad Simon Bolivar, Caracas, Venezuela

Date/Time: Wednesday, September 25, 2002, 10:00 a.m. (PST)

Location: Building 921 Room 137(Sandia-CA)
Building 980 Room 95 (Sandia-NM)

Brief Abstract: Efficient techniques for solving unconstrained minimization problems often require the computation of the first derivative. We are aware that first order information should be used whenever available. Nonetheless, real world applications deal with nondifferentiable functions or may preclude the use of derivatives, because the functional values arise either from a complex simulation package, or from inaccurate sample values. Not even a numerical approximation of the derivatives is a reliable approach. We will be dealing with nonsmooth functions f with directional derivatives defined everywhere and focus our attention on the problem of obtaining a point x* satisfying the NonSmooth Necessary Condition (NSNC) on a finite set D = {d1, ..., dm}, of m bounded directions. This condition will be defined and explained. It will be obvious that a point x* that satisfies the first order necessary condition also fulfills NSNC if f is convex. The opposite direction holds without convexity: If NSNC holds on D, if f is strictly differentiable at x*, and if the set D positively spans the space then x* satisfies the first order necessary condition. We review a recent algorithm that generates a sequence converging to an x* satisfying NSNC. Essentially the sequence satisfies a new decrease condition quite independent of derivative computations. Convergence is proved for either nonsmooth convex functions or strictly differentiable functions. We will emphasize some recent developments under the same ideas and state some still open questions. Finally, some numerical results on a sequential algorithm and hints of results with an asynchronous parallel implementation of the algorithm will be discussed. This talk strongly relies on the paper: García-Palomares and Rodríguez, New sequential and parallel derivative-free algorithms for unconstrained minimization, SIAM on Optimization 13-1 (2002) 76-97.

CSRI POC: Paul Boggs, (925) 294-4630

Title: Polyhedral techniques for approximation algorithms

Speaker: Ojas Parekh
Carnegie Mellon

Date/Time: Tuesday, January 22, 2002, 9:0010:00 am

Location: Building 980 Room 95

Brief Abstract: Approximation algorithms produce approximate solutions, along with a performance guarantee, to instances of NP-hard problems, providing a means of attack at problems which are widely believed to be intractable. Polyhedral analysis is an area of classical mathematics which has also enjoyed a growing and diverse audience due to the seemingly ubiquitous applicability of linear programming both as a practical and theoretical tool. My talk will focus on polyhedral approaches to the weighted edge dominating set problem in graphs, a fundamental network design problem in which we seek to cover all the edges of a graph with a minimum weight set of edges. I will present the first approximation algorithms for this problem, culminating with a 2-approximation algorithm, and show that the problem is in fact equivalent to a generalization of the famous weighted vertex cover problem, that of covering the edges of a graph with a minimum weight set of vertices. I will also present approximation algorithms for variants of the problem in which we seek an edge dominating set that must be connected or a tour.

CSRI POC: Robert D. Carr, 845-8562

Title: Toward the Multiscale Numerical Solution of Chemically Reacting Systems

Speaker: Dr. Linda R. Petzold
University of California, Santa Barbara

Date/Time: Tuesday, June 18, 2002, 10:00-11:00 am (PST)

Location: Building 921 Room 137 - Sandia-CA
Building 980 Room 24 - Sandia-NM (Videolinked)

Brief Abstract: Stochastic simulation enables the numerical simulation of the time evolution of a well-stirred chemically reacting system in a way that takes proper account of the randomness that is inherent in such a system. However, the computer times required to simulate over reasonable time periods are prohibitively long if the molecular populations of any of the reactant species are large. In cellular systems, for example, this is nearly always the case. At the same time, the deterministic reaction rate equations are well suited for modeling of systems with large molecular populations, but cannot accurately capture the evolution of species whose populations are small.

In this talk we will examine a range of models from stochastic simulation to the deterministic reaction rate equations. Since the models naturally segue to each other, there is some reason to hope that a multiscale numerical method that treats each species and reaction at the most appropriate scale could be developed. We will describe our progress on the first phase of this effort: the development of efficient and robust accelerated stochastic simulation algorithms inspired by the recently developed tau-leaping method of Gillespie. This is joint work with Dan Gillespie and Muruhan Rathinam.

CSRI POC: Tamara Kolda, (925) 294-4769

Title: Taylor series integration of differential-algebraic equations: Automatic differentiation as a tool for simulating rigid body mechanical systems

Speaker: Eric Phipps
Cornell University, Center for Applied Mathematics

Date/Time: Tuesday, February 26, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Rigid body mechanical systems serve as models of dynamical systems ranging from automobiles to the human body. According to Newton's Laws, the equations of motion for such systems can be formulated as a system of second order, ordinary differential equations, which can then be solved numerically by standard methods. However, for even relatively simple mechanical systems, the equations of motion can be algebraically quite complicated and difficult to derive. A simpler approach is to formulate the equations of motion instead as a system of differential-algebraic equations (DAEs), such as Euler-Lagrange DAEs. In this talk I will describe a technique for simulating mechanical systems based on Taylor series integration of Euler-Lagrange DAEs. The Taylor series integration method is capable of computing numerical solutions to differential equations with very high accuracy and large step sizes. Furthermore, by employing the technique of automatic differentiation, solutions to the governing DAEs and their derivatives can be computed solely from a specification of the Lagrangian and algebraic constraints. The combination of these techniques allows the efficient, accurate simulation of rigid body mechanical systems using a minimal amount of information to specify the model. These techniques have been implemented in MATLAB using a custom automatic differentiation library built in C++ that, via the use of

templates and object oriented methods, is capable of efficiently computing high degree tensor derivatives of matrix valued functions written in MATLAB or C++.

CSRI POC: Andy Salinger, 845-3523

Title: I. Parallelizing Associative Operations on a Computational Grid
II. Exploring Equivalence Classes of Semantics in Document Collections

Speaker: Bill Pottenger
Lehigh University

Date/Time: Monday, April 8, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: In the first part of this presentation I will discuss our research in developing a novel parallel- pipelined model of execution for computational grids. In [Kuntraruk and Pottenger 2001] we show that a speedup linear in the number of processors is achievable for applications involving associative operations (e.g., reductions) based on our parallel-pipelined model. The class of applications that leverages this model encompasses a wide range of operations that heretofore have not been amenable to linear speedup. Examples of such operations include parallel 'gather' in 'scatter/gather' algorithms, the parallel computation of dynamic last values, and the parallel execution of multiple-exit loops. In addition, in collaboration with our industry partner Data Synapse, Inc., we are developing an analytical model for massively parallel feature extraction in natural language processing that employs unused cycles on networks of PCs/workstations in a computational grid. Modeling of this nature is critical to the development of resource requirement estimation capabilities for applications executing under our parallel-pipelined model of execution on a computational grid. This is, I believe, an important area of overlap with Sandia's mission because much of the simulation and modeling work conducted at Sandia involves codes that rely heavily on reduction operations, and building middleware infrastructure that maximizes the utility of the various supercomputers at Sandia is an important research goal. The research we are conducting can be targeted directly at the ASCI computational grid currently under construction as part of the nation-wide, distributed nuclear weapons complex (NWC) involving CPlant and the ASCI red, white, and blue mountain supercomputers.
We are currently verifying and scaling our models on the Los Lobos cluster at the University of New Mexico.

In the second part of this presentation I will discuss our research into theories of information retrieval and textual data mining that form a basis for computation of object similarity employed in visualization tools such as VxInsight. The motivation for this work stems from the observation that, in general, object similarity functions lack a unifying theoretical framework. For example, although current research in Latent Semantic Indexing (LSI) shows improvements in performance for a wide variety of information retrieval systems, to our knowledge there is no theoretical basis for choosing the number of dimensions when forming a reduced representation term-document matrix from the Singular Value Decomposition (SVD). In recent work, however, we have demonstrated empirically and proven mathematically that a transitive path exists between every pair of co-occurring terms in the reduced representation term-term matrix. The implication of this result is that it may be possible to define a reflexive, symmetric and transitive term-term relation that defines equivalence classes of semantics [Kontostathis and Pottenger 2002]. A second example of the lack of a unifying theoretical framework is reflected by the absence of models that integrate both negative and positive similarities in an object similarity function. VxInsight, for instance, currently accepts input similarities that are non-negative [Davidson et al. 1998], and computes output similarities (for documents) that are non-negative [Borner et al. 2002]. Similarly, although the SVD employed in LSI results in both positive and negative values of similarity in the reduced representation term-term matrix, there is to our knowledge no sound theoretical basis for understanding the meaning of the negative values. Our recent research results have revealed, however, that such negative values may play a significant role in defining the aforementioned equivalence classes of semantics. The incorporation of such

negative object similarity values is intuitively an important piece of the puzzle that a unifying theoretical framework must address, and I will discuss our progress in developing a theory that unifies both positive and negative similarity values.

CSRI POC: Bruce Hendrickson, 845-7599

Title: Dynamic Programming Like It Has Never Been Before: Analyzing Complex Process Models Using an Idempotent-Algebraic ISS Small Gain Theorem

Speaker: Henry Potrykus,
University of Texas, Austin

Date/Time: Tuesday, October 22, 2002, 9:30-10:30 am

Location: Building 980 Room 95

Brief Abstract: Process models, especially those arising in the analysis of chemical engineering systems, are often complex (here taken to mean "involving many coupled equations"), nonlinear systems of differential equations. The stability analysis of said models, an important, basic question from an engineering point-of-view, traditionally takes two forms. First, linearization of the system's vector field around an operating point gives a local stability characterization. Otherwise, a Liapunov (energy) function for the system is sought, by which global nonlinear stability to a point may be concluded. As the problem now stands, Liapunov functions for complex, nonlinear systems are impossible to compute due to the systems' general asymmetry and frequent lack of simple (point) attractors.

I will present a method which partially fills this lacuna, and is complementary to the local stability analysis mentioned above: Here a general nonlinear input-to-state stability (ISS) small gain theory will be developed using the algebraic techniques first discovered through the study of the classical Bellman (shortest path) problem. The little bit of algebra used in the theory will be explained, the interplay between this algebra and the "real" algebra everyone is used to will be shown to be important, and the theory's relationship to some basic problems in graph theory will be presented.

Finally, we will not lose touch with what motivated the theory in the first place: the technique's applicability to complex chemical engineering systems (distillation columns, crystallization units, and continuous-flow reactors), will be described.

CSRI POC: Scott Mitchell, 845-7594

Title: High Performance Java for Scientific and Technical Computing

Speaker: Roldan Pozo
Leader, Mathematical Software Group
National Institute of Standards and Technology

Date/Time: Thursday, March 14, 2002, 9:30-10:30 am

Location: Building 980 Room 95

Brief Abstract: Among the many features of Java as program development platform, one of the most commonly cited shortcomings is its performance, particularly for computational-intensive codes. In this presentation we take a close look at how Java works and how it can be made to execute faster for scientific simulations and modeling. We will examine optimization strategies and bytecode transformations that

generate 100% Pure Java with 2-10x performance improvement, with speeds competitive with optimized C/C++ and Fortran. We will also look at ongoing research in compiler technology and distributed computing that allow Java applications to reach supercomputing status. Finally, we will discuss activities of the Java Numerics Group, an industry-wide consortium focusing on the development high performance tools, compilers, and libraries for high-end Java applications.

CSRI POC: Mike Heroux, (320) 845-7695

Title: Probabilistic Ordinal Optimization Approaches for Continuous-Variable Optimization Under Uncertainty

Speaker: Vicente Romero, Validation and Uncertainty Quantification Dept. 9133
Chun-Hung Chen, George Mason University

Date/Time: Wednesday, July 17, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Recent work in the Operations Research field has produced an optimal method ("OCBA") for apportioning Monte Carlo sampling amongst multiple uncertain or stochastic discrete systems in order to most efficiently resolve their statistical behavior or merit for the purpose of selecting the best (or several best) option(s). Even if the sampling must be truncated before a desired confidence of correct selection is achieved, the odds that the OCBA algorithm has truly identified the best alternative(s) are known at every stage of sampling.

OCBA can also be applied to continuous-variable (C-V) optimization problems involving uncertainty. Here improvement steps are taken based on ordinal ranking of candidate steps according to relative merit, rather than attempting to accurately resolve the absolute merit of each alternative. Confidence intervals (CI) on the correctness of the indicated ranking decrease at an exponential rate proportional to 1/e**N_T where N_T is the total number of Monte Carlo samples expended over all options locally being considered - whereas a vastly slower rate proportional to 1/N_j**0.5 for each of the J individual options governs convergence to absolute merit values. Adequate individual convergence is required for accurate gradients of objective function statistics. This can make gradient-based optimization relatively expensive in many cases. Alternatively, the ordinal ranking efficiency of OCBA can be augmented with considerations of local spatial correlation in the design space and then coupled with C-V ordinal optimizers such as Genetic Algorithms, Pattern Search techniques, DIRECT, Simplex methods, etc. This talk discusses the Probabilistic Ordinal approach to optimization under uncertainty, along with advanced sampling, selection, and optimization strategies for its implementation.

CSRI POC: Vicente Romero, 844-5890

Title: Using FEMLAB as a platform for studying time integration methods

Speaker: David Ropp (09214)

Date/Time: Wednesday, October 16, 2002, 10:00-10:30 am

Location: Building 980 Room 95

Brief Abstract: FEMLAB is a versatile computational tool based on MATLAB that solves PDEs using the finite element method. As such, it can also be used to study and compare different algorithms for solving PDEs. We have used it as a platform to study various time integration methods. This was not a straight-forward application of FEMLAB, though, and we had to write and rewrite certain routines and

misuse certain features. The benefit of this is that we have the computational and analytical capabilities of FEMLAB and MATLAB.

CSRI POC: David Ropp, 845-7431

Title: Multi-time Partial Differential Equations

Speaker: Jaijeet Roychowdhury, ECE Department and the DTC
University of Minnesota

Date/Time: Thursday, November 21, 2002, 9:00-10:00 am (MST)

Location: Building 980 Room 95

Brief Abstract: Many physical systems feature fast repetitive patterns whose characteristics change slowly. Examples include amplitude and frequency modulation in electrical or optical communications. Computing the dynamics of such systems efficiently has long been of considerable practical interest. A relatively recent approach towards doing so recasts the system's differential equations as PDEs in artificial time-scales. The resulting PDEs (termed Multitime PDEs or MPDEs) can often be solved far more efficiently than the original differential equations. In this talk, we will review basic concepts of MPDEs, recount their origins and development to date, and point out connections with other methods for fast-slow problems. We will outline the application of MPDE techniques to forced and autonomous systems, and show how they enable automatic macromodelling of linear time-varying systems. We will also discuss open problems and promising directions for further research.

CSRI POC: David Day, 844-1868

Title: Object-based Storage Devices: Yes it's a crazy idea, but is it crazy enough?

Speaker: Tom Ruwart
Chief Technology Officer at Ciprico, Inc.

Date/Time: Wednesday, June 12, 2002, 11:00-12:00 noon

Location: Building 980 Room 95

Brief Abstract: Abstract: Our ability to create data and use data storage devices have seen amazing growth rates over the past several years. The current software data storage technologies such as file systems and protocols are based on 30-year old technologies and are beginning to show signs of stress under the demands of current-day users. This presentation describes the effects of these demands and demonstrates how an Object-based Storage Device approach can alleviate the problems associated with current methods and technologies.

CSRI POC: Lee Ward, 844-9545

Title: Domain Decomposition Preconditioners: Theoretical Properties and Applications to the Compressible Euler Equations

Speaker: Marzio Sala, MOCS/SB/EPFL
Swiss Federal Institute of Technology of Lausanne, Switzerland

Date/Time: Friday, November 22, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: We consider domain decomposition preconditioners for the parallel iterative solution of sparse linear systems. Non-overlapping Schur complement and over-lapping Schwarz methods will be introduced. The attention is focused on two-level Schwarz methods. A class of algebraic coarse level operators, based on the concept of smoothed agglomeration, is presented. The resulting procedure, which does not require a coarser grid, is completely algebraic and well-suited for parallel computations on unstructured grids in two and three dimensions. First, for a model problem, convergence estimations are derived, and numerical results are reported to compare the proposed schemes with other two-level methods. Then, the proposed preconditioners are applied to the compressible Euler equations. A parallel adaptive pseudo-transient Newton-Krylov-Schwarz framework is outlined. Multidimensional upwind residual distribution schemes are used for the space discretisation, while an implicit time-marching scheme is employed for the time discretisation. The linear system arising from the Newton method applied to the resulting nonlinear system is solved by the means of Schwarz preconditioned Krylov iterations. A parallel grid adaptation procedure is used to increase the solution accuracy. Numerical results are presented for 3D problems of aeronautical interests on unstructured grids.

CSRI POC: John Shadid, 845-7876

Title: Computational and Theoretical Study of Liquids and Glasses Under Tension

Speaker: Vincent K. Shen
Princeton University

Date/Time: Tuesday, March 5, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Mechanical strength is a fundamental material property, and yet a rigorous microscopic theory for it is still lacking. In liquids, fracture is triggered by the formation of critically-sized bubbles, and it is responsible for a variety of technologically important phenomena, such as cavitation, explosive boiling, boiling heat transfer, and sonoluminescence. Using tools from molecular simulation and statistical mechanics, we have studied the initial stages of the process by which a pure, metastable liquid transforms into a stable vapor, known as homogeneous bubble nucleation. In studying the molecular-level processes that drive this particular phase transition, we have introduced a fundamental thermodynamic quantity called the degree of metastability that measures the degree of penetration into the liquid's region of metastability. Furthermore, we show that various properties of the critical bubble, such as its size, interfacial thickness, and work of formation, scale with this quantity independently of temperature.

While the potential energy hypersurface, or the energy landscape, has provided a wealth of information about the vitreous state of matter at the molecular level, the prospect for using the landscape formalism as a practical tool for predicting the macroscopic properties of glasses is also promising. As an illustrative example, we consider an important mechanical property in engineering applications employing amorphous solids: tensile strength. Computationally, the ultimate isotropic tensile strength of a glass can be determined by constructing the equation of state of its energy landscape. We have performed an extensive simulation study of the n-alkane glasses ranging from $n = 1$ to 48 to investigate the influence of molecular architecture on tensile strength. Remarkably, the ultimate isotropic tensile strength varies non-monotonically with chain length, exhibiting a maximum at $n = 3$. It is known that fracture in solids, in particular glasses, is initiated by voids or cracks, and it is therefore not surprising that the tendency to form such weak spots is directly related to tensile strength. As expected, the size and number of cavities at fracture correlate well with the tensile strength trends. The remarkable non-monotonic chain length dependence can be explained by a simple mean-field theory that highlights the effect of intermolecular packing effects on tensile strength.

CSRI POC: Andrew Salinger, 845-3523

Title: Electron and Hole Traps in the Bulk and at the Surfaces of Oxides

Speaker: Alexander Shluger
Department of Physics and Astronomy
University College London

Date/Time: Monday August 19, 2002, 9:15-10:15 am

Location: Building 980 Room 95

Brief Abstract: I will present the results of plane wave DFT and embedded cluster calculations and comparison of the electronic properties of defects in the bulk of MgO, SiO2, ZrO2 and HfO2. The electronic and atomic structures and the mechanisms of diffusion of anion vacancies and interstitial Hydrogen and Oxygen atoms and molecules in different charge states are considered. Defects in SiO2, ZrO2 and HfO2 are treated in conjunction with their electronic properties at the interface with Si. Electrical levels of these defects are calculated in the band gap and with respect to the Si conduction band. Their structures and properties are compared with the properties of similar defects in MgO and silicates. The nature of the electron and hole traps at surfaces of MgO is discussed. I will demonstrate that low-coordinated sites at surfaces of these materials can serve as traps for electrons, holes and excitons.

CSRI POC: Harold P. Hjalmarson, 844-8888

Title: Diffusion of Small Molecules and Atoms in Nanoporous Materials and Dense Solids

Speaker: Professor. David Sholl, Dept. of Chemical Engineering
Carnegie Mellon University

Date/Time: Monday, June 24, 2002, 9:30-10:30 am

Location: Building 980 Room 95

Brief Abstract: I will describe atomic-scale modeling of several materials that have applications as gas separation membranes. I will first discuss zeolite membranes, in which gas transport occurs by adsorption and diffusion of gas molecules inside the nanometer-scale pores of zeolites. By using a combination of atomically detailed Molecular Dynamics and continuum models, it is possible to directly predict the performance of zeolite membranes under practical experimental conditions. A related class of nanoporous materials with potential use as membranes are single walled carbon nanotubes. I will present results from atomistic simulations that strongly suggest that carbon nanotubes may have exceptional properties in terms of the gas transport rates they can create.

Finally, I will discuss the diffusion of atomic hydrogen through CuPd alloys. These alloys are thought to have favorable surface chemistry properties for hydrogen separations, but the existence of multiple solid phases complicates the choice of alloy composition. I will show how ab initio Density Functional Theory calculations have been used to understand the diffusion atomic hydrogen in the different solid phases of CuPd.

CSRI POC: Marcus Martin, 284-6355

Title: Improving The Memory Performance of Java Workloads

Speaker: Yefim Shuf
Princeton University

Date/Time: Thursday, June 6, 2002, 9:00 - 10:00 a.m.

Location: Building 980 Room 95

Brief Abstract: The rapidly widening performance gap between processor and memory speeds (i.e., the "memory wall" problem) and the growing importance of the Java platform make it crucial to improve the memory behavior of Java workloads. Various features of Java -- object orientation, dynamic linking, garbage collection, multithreading, and the absence of exposed pointers and pointer arithmetic -- create opportunities for improving memory system performance using runtime techniques within the Java Virtual Machine (JVM). The interaction of these features makes the realization of these opportunities non-trivial and challenging.

My research targets the "memory wall" problem in the context of Java. We have developed a better understanding of the memory behavior of Java workloads and how it compares with the memory behavior of Traditional scientific and technical workloads. In particular, we answer the following questions: 1) How can the data locality of Java applications be improved? 2) How can the overhead of garbage collection be reduced? 3) What new and traditional optimizations can be applied to Java programs? 4) What makes it difficult to apply some traditional optimizations to Java?

In this talk, I will discuss the progress we have made towards answering these questions. First, I will present some results from our comprehensive study of the memory performance of Java workloads. This study gives us important insights into the inherent high-level and low-level memory behavior of Java programs, identifies the sources of performance problems, and serves as the foundation for the rest of the talk. Next, I will introduce the notion of "prolific" types and propose a type-based framework that enables new techniques for memory management (a type-based garbage collection) and facilitates the application of known optimization techniques such as object co-allocation. Then I will discuss a novel locality-conscious algorithm for traversing reachable objects at garbage collection time, which reduces garbage collection pauses and enhances the locality of surviving objects. Time permitting, I will present our recent results on the impact of multiple page size support on Java application performance.

CSRI POC: Erik DeBenedictis, 284-4017

Title: MPI-3: Evolution, Revolution, or Status Quo

Speaker: Dr. Anthony Skjellum
MPI Software Technology, Inc and Mississippi State University

Date/Time: Wednesday, June 12, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: In this talk, opportunities, challenges, and issues with the MPI-1, and MPI-2 message passing interfaces are considered, and opportunities for improving MPI in future are considered against the background of evolving systems architectures, and application needs. The alternative, replacing MPI with a completely different explicit parallel programming model is also considered.

The design of message passing APIs, their interactions with applications, OS, and other middleware are considered. The design triple of (latency, bandwidth, overhead) are key driving issues in seeking improvements of such systems. The ability to offer interfaces that capture temporal locality to differing degress are considered. The need to capture newer requirements (e.g., fault handling) are considered as well.

Handling of legacy codes (always of paramount importance), is considered against the need to access a high degree of achievable performance as architectures change over the next five to ten years.

CSRI POC: Ron Brightwell, 09223, 844-2099

Title: Global Positioning Devices

Speaker: Mark Smith
Colorado State University

Date/Time: Friday, April 5, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Global Positioning Devices are becoming increasingly common for both recreational and professional services. As a recreational tool, they are valued by outdoor enthusiasts. As a professional tool, geologists, surveyors and numerous others depend on them for reliable, accurate information.

One of the most attractive features of GPS devices is their ability to communicate with external appliances, specifically, the personal computer. This is seen by the fact that most GPS devices on today's market have hardware that enables them to interface to an outside source.

As a programmer and owner of a GPS receiver, I have encountered numerous frustrations when trying to find useful personal computer software to communicate with my GPS. This is why I decided to write the South Fork GPS Library -- to encourage others to write GPS interfacing applications for their computers. The South Fork GPS Library is a set of C modules that handle communication, input, and data processing with GPS devices. It creates a "black box" that provides developers with a simple means to handle the information sent from a GPS device to a computer. The purpose of this software is to simplify GPS application development, thereby encouraging others to write their own software for processing GPS data. The library consists of three main components: communication, input parsing, and a "controller" component. The communication component is written to interface to a standard serial port and constantly send received information to the parsing component. The parsing component receives input from the communication component and then uses code built from flex and bison (the GNU equivalent of Lex and Yacc) to tokenize and parse the input. Lastly, the controller handles initialization of serial input and threads (one for communication and one for parsing) as well as releasing these resources back to the operating system once the program has terminated. Together, these components provide the end user with a handful of external variables that hold the most recent input sent from the GPS device. These variables include latitude, longitude, speed and bearing.

As stated previously, the purpose of this software is to encourage others to develop applications for communicating with GPS devices. Therefore, the software will be released under the GNU Public License so that developers may have free access to the source code and any subsequent modifications.

CSRI POC: Erik DeBenedictis, 284-4017

Title: Nonparametric model of random uncertainties in dynamical systems:
An overview

Speaker: Professor Christian Soize, Laboratory of Engineering Mechanics
University of Marne-la-Vallee, France

Date/Time: Friday, July 19, 2002, 10:00-12:00 noon

Location: Building 980 Room 95

Brief Abstract: We present a nonparametric model of random uncertainties for linear and nonlinear dynamical systems. This nonparametric model is a new approach allowing the data uncertainties and the modeling errors to be taken into account for dynamical systems. In part 1, we summarize the usual parametric probabilistic model of random uncertainties in structural dynamics and the new approach recently developed by the author, called the nonparametric probabilistic model of random uncertainties in dynamical system. This non-parametric model is constructed by using the maximum entropy principle and the role played by the constraints defining the available information is analyzed. Part 2 deals with the set of symmetric positive-definite real random matrices, called the positive- definite set of random matrices. This ensemble constitutes a new ensemble of the random matrix theory. In part 3, we present the nonparametric model of random uncertainties for linear and nonlinear dynamical systems. Part 4 is devoted to the comparison of the positive-definite ensemble of random matrices with the Gaussian Orthogonal (GOE) ensemble of random matrices. Part 5 deals with the validation of the nonparametric model of random uncertainties in vibration analysis using a comparison with the parametric approach. Part 6 is devoted to transient dynamics induced by shocks in structures with non-homogeneous random uncertainties; an experimental comparison is presented. In part 7, we present the frequency response function calculation of structures with non-homogeneous random uncertainties and with an experimental comparison. Finally, part 8 is devoted to the nonparametric model of random uncertainties for nonlinear dynamical systems.

CSRI POC: Steve Wojtkiewicz, 284-5482

Title: SAN-Based File Sharing

Speaker: Steve Soltis
Distributed file systems designer/architect, DataPlow Inc

Date/Time: Tuesday, February 12, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Storage Area Networks (SANs) are high-speed, channel-based networks that connect multiple computers directly to central pools of disk storage. SAN technologies, along with SAN-based distributed file systems, enable high-speed file sharing among multiple computers. In contrast to traditional LAN-based approaches, the benefits of SAN-based file sharing include improved system-wide performance and scalability, reduced time, effort, and costs managing large datasets, and support for storage consolidation. This presentation will introduce the use of SAN technologies to promote high-speed file sharing as well as discuss two SAN-based file systems from DataPlow, Inc.

CSRI POC: Lee Ward, (240) 453-3893

Title: Full physics full chemistry multi-scale modeling of materials

Speaker: Alejandro Strachan
Materials and Process Simulation Center (MSC)
California Institute of Technology

Date/Time: Wednesday, April 17, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: First principles predictions of materials properties requires bridging an enormous range of time and length scales: from electrons [Quantum Mechanics (QM)] to atoms [Force Fields (FF) and Molecular Dynamics (MD)], to mesoscopic (Phase Fields, Dislocations Dynamics) and macroscopic modeling (Finite Elements). Developing such tools will allow the prediction of material behavior at extreme conditions and the efficient design of materials with tailored properties and functionalities.

I will present recent work in multiscale modeling aimed at the development of first principles tools for the simulation of mechanical, chemical and electrical properties and processes in a wide range of materials. A key step in multiscale modeling is the development of first-principles-based force fields; FFs describe the atomic interactions in a computationally efficient way and allow the simulation, with molecular dynamics, of millions of atoms. I will describe new generation FFs that allows the accurate description of complex phenomena such as chemical reactions, charge transfer and polarization, and equations of state of various bulk systems in wide pressure ranges, for metals, ceramics and organics. I will focus on the application of these force fields in two key applications: i) plastic deformation and failure in metals and ii) chemical response to shock loading in a high-energy material (RDX).

We developed an *ab initio* Force Field for Tantalum and used it to characterize the key processes that control plasticity and failure, such as core structure, energy and mobility of dislocations, crack propagation and shock induced failure. Such information was used with micromechanical modeling to predict strain-stress curves of single-crystal Ta under uniaxial tension in a wide range of temperatures and strain rates; our first principles results are in good quantitative and qualitative agreement with experiments, capturing: i) the dependence of the yield point on temperature and strain rate, ii) the presence of a marked stage I of easy glide at low temperatures and high strain rates; and iii) the sharp onset of stage II hardening.

On the other hand, a recent breakthrough, the development of the First Principles based FF ReaxFF, allows the accurate description of complex chemical reactions as well as structural and mechanical properties of materials. I will report on the use of ReaxFF with MD to study the initial chemical events induced by shockwaves in the high-energy material RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine). We simulated shock propagation via high velocity impact simulations of RDX nano-slabs at various impact velocities and studied the induced chemical reactions. For sufficiently high impact velocities (>2 km/s) the RDX molecules decompose and react to form a variety of small molecules whose time evolution is reported. Such atomistic simulations with new-generation force fields are a powerful tool to characterize the initial stages of condensed-phase detonation in high-energy materials, providing relationships between mechanical load, the induced chemical reactions at atomic level and mesoscopic or macroscopic quantities such as local temperature and population of various species.

CSRI POC: John Aidun, 844-1209

Title: JMPL: Java Math Package Launcher Automating GUI Development and Maintenance for Science and Engineering

Speaker: Andrew Strelzoff
University of California, Santa Barbara

Date/Time: Tuesday, June 18, 2002, 1:00–2:00 pm (PST)

Location: Building 921 Room 137 – SNL CA
Building 980 Room 95 – SNL NM (Videolinked)

Brief Abstract: Developers of scientific applications typically lack the time and expertise necessary to produce and maintain high quality Graphic User Interfaces (GUI's). In this presentation we introduce JMPL: Java Math Package Launcher, an environment for developing and maintaining GUI front ends for scientific computing. JMPL assists the development process by automatically generating matching GUI skeletons from scientific applications. These skeletons may be further developed and customized in the JMPL editor before distribution to the community of users. The JMPL revision manager assists in the

process of updating the GUI as the underlying program changes. The result is an environment for continuing publication and improvement.

In this talk we will focus on two critical areas of the JMPL project. The first is the fundamental task of discovering interface requirements from a program and the translation of these requirements into an intermediate format, XML/UML. This intermediate format is then parsed by a browser, which dynamically constructs the GUI. The second area covered will be the development of an algorithm to estimate, which parts of an underlying application have changed. This is the input to the revision manager, which will allow it to assist in keeping the GUI up to date. Preliminary results will be presented, and the future development and direction of the project including extension to object oriented languages will be discussed.

CSRI POC: Kevin Long, 294-4910

Title: Zero Copy - Efficiency and Performance Engineering for System Software in Communication

Speaker: Professor Thomas M. Stricker
Laboratory for Computer Systems, ETH Zuerich, SWITZERLAND

Date/Time: Friday, April 12, 2002, 10:30-11:30 am

Location: MO285 Room 14

Brief Abstract: The steady increase in hardware performance enables the processing of richer multi-medial data-streams (better video and sound) on commodity PCs. Solutions that required special purpose hardware in the past are encountered on regular PCs at present and will be implemented on highly mobile and wearable computers in the near future. Still all such high performance applications will rely on communication-system software that makes optimal use of the hardware resources offered by the platform it is designed for. As an example I am presenting a novel zero copy TCP/IP protocol stack for cluster computing based on Gigabit Ethernet. So far a true zero copy protocol stack seemed impossible with the simple hardware found in commodity PCs. For our solution we enlisted protocol speculation, a probabilistic technique, that is well proven in the field of processor architecture. Our specific results in cluster computing demonstrates the necessity of software performance engineering in communication systems. When developing new software for communication systems we must work towards a good understanding and a precise characterization of the performance and the efficiency in addition to the usual design goals of new functionality, provable correctness and standard compliance.

CSRI POC: Rolf Riesen, 845-7363

Title: High Resolution Methods for Time Dependent Problems with Piecewise Smooth Solutions

Speaker: Eitan Tadmor
University of California Los-Angeles

Date/Time: Tuesday, June 11, 12, 2002, 9:00-10:00 am

Location: Building 980 Room 24

Brief Abstract: A trademark of nonlinear time-dependent convection problems is the spontaneous formation of nonsmooth macroscale features, like shock discontinuities and non-differentiable kinks, which

pose a challenge for high-resolution computations. We overview recent developments of modern computational methods for the approximate solution of such problems.

High-resolution central schemes will be discussed as a prototype example for high-resolution local methods. The family of central schemes offers high-resolution "black-box-solvers" to an impressive range of such nonlinear problems. The main ingredients here are detection of spurious extreme values, non-oscillatory reconstruction in the directions of smoothness, numerical dissipation and quadrature rules. Enhanced spectral viscosity will be discussed as an example for high-resolution global methods. The main ingredients here are detection of edges from spectral data, separation of scales, adaptive reconstruction, and spectral viscosity.

CSRI POC: Pavel Bochev, 844-1990

Title: Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time

Speaker: Shang-Hua Teng
Boston University/Akamai Technologies

Date/Time: Thursday, October 3, 2002, 10:00-11:00 am

Location: Building 980 Room 95

Brief Abstract: Theorists have been challenged by the existence of remarkable algorithms that are known by scientists and engineers to work well in practice, but whose theoretical analyses are negative or inconclusive. The root of the problem is that algorithms are usually analyzed in one of two ways: by worst-case or average-case analysis. The former can improperly suggest that an algorithm will perform poorly, while the latter can be overly optimistic because the random inputs it considers can bear little resemblance to those encountered in practice. We propose an analysis that we call smoothed analysis that can help explain the success of many algorithms that both worst-case and average-case cannot. In smoothed analysis, we measure the performance of an algorithm under slight random perturbations of arbitrary inputs. In particular, we consider Gaussian perturbations of inputs to algorithms that take real and complex inputs, and we measure the running time of algorithms in terms of the input size and the variance of the perturbations. We show that the simplex algorithm has polynomial smoothed complexity.

The simplex algorithm is the classic example of an algorithm that performs well in practice but takes exponential time in the worst case. In the 1980's the simplex algorithm was shown to converge in expected polynomial time on various distributions of random inputs, most notably by Borgwardt and Smale. However, the last 20 years of research in probability and numerical analysis have taught us that these random instances have very special properties that one should not expect to find in practice.

For every matrix A with entries of absolute value at most 1, every vector c, and every vector b whose entries are 1 or -1, we show that the simplex algorithm using the shadow-vertex pivot rule takes expected time polynomial in 1/sigma and the sizes of A and c to solve:

minimize c'x
s.t (A + \sigma G) x <= b

If A is well-scaled, then the solution to this program is an approximation to the original.

This is joint work with Daniel Spielman of MIT.

CSRI POC: Bruce Hendrickson, 845-7599

Title: Fixed-Polynomial Approximate Spectral Transformations for Preconditioning the Eigenvalue Problem

Speaker: Heidi K. Thornquist
Rice University

Date/Time: Thursday, November 14, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Arnoldi's method is often used to compute the eigenvalues and eigenvectors of large, sparse matrices. Various techniques can by employed to improve the convergence of this method, some resulting in considerable expense. We will present an efficient preconditioning scheme that accelerates the convergence of Arnoldi's method to the eigenvalues of smallest magnitude for Hermitian and non-Hermitian matrices. This scheme uses a fixed-polynomial operator to approximate the spectral transformation. When used with the Implicitly Restarted Arnoldi (IRA) method, this achieves the effect of a spectral transformation without factoring a matrix.

Numerical results indicate that this employing this preconditioning scheme with IRA is very effective. It is just as accurate as IRA in regular mode, more accurate than using preconditioned GMRES started anew for each linear solve, and considerably less expensive than either. Furthermore, using a fixed-polynomial operator with IRA compares favorably with two other current methods: Jacobi-Davidson and Locally Optimal Block Preconditioned Conjugate Gradient.

This work was supported in whole or in part by the Los Alamos National Laboratory Computer Science Institute (LACSI) through LANL contract number 03891-001-99-49, as part of the prime contract (W-7405-ENG-36) between the Department of Energy and the Regents of the University of California.

CSRI POC: Rich Lehoucq, 845-8929

Title: Variational ALE method

Speaker: Pururav Thoutireddy
California Institute of Technology

Date/Time: Tuesday, June 4, 2002, 9:00 – 10:00 am

Location: Building 980 Room 95

Brief Abstract: Variational Arbitrary Lagrangian Eulerian (VALE) method is Finite Element Method for PDEs extended to include horizontal variations. In this method in addition to nodal variables, nodal coordinates are also sought as solution to the problem by appealing to variational principles. Since this method is variational, it preserves symmetries and has good energy behavior. In addition, since both reference (nodal) and deformed coordinates are variables one can argue that Lagrangian and Eulerian formulations are special cases in this method.

Since mesh itself is the solution, nodes "flow" so as to give the optimal mesh, which is unique. The resulting mesh adaption scheme is devoid of error estimates and mesh-to-mesh transfer, hence interpolation errors. Further, this mesh adaption scheme is ideally suited for parallel computations, as the topology of the mesh remains the same, which obviates dynamic load balancing. In addition optimal property of the mesh adaption scheme makes it ideal for problems involving steep gradients as in the case of fracture, shock capture and shear band propagation.

This method also enables solution of a whole new class of problems involving change in reference configuration, as in the case of solution for optimal shape of precipitates in precipitation hardening materials.

In this seminar above mentioned salient features of this method will be demonstrated with numerical simulations in the areas of fracture mechanics, materials and shock capture.

CSRI POC: James S. Peery, 845-9336

Title: A Secure Server System

Speaker: Sonja Tideman

Date/Time: Friday, August 9, 2002, 9:45-10:30 am

Location: Building 980 Room 95

Brief Abstract: System security is never perfect; the most secure system can be cracked if a talented enough cracker decided to attack it. If we assume that all systems can be cracked, an effective security system should be capable of immediately detecting a security breach and isolating the compromised server to prevent further damage and protect sensitive information. The compromised server must be removed from the network and all client connections terminated. Depending on the nature of the server, terminating all client connections could be very costly for the client. A better approach would be to allow another server to take over for the compromised server. I will present a protocol for a network of servers to be able to monitor each other and to take over an existing client connection when required. This fail-over is completely transparent to client applications.

The server side of the protocol consists of a general purpose monitoring system and a logging system. The monitoring system monitors all servers within the network using multiple applications that can be plugged into the monitoring system. These monitoring applications may continuously monitor any number of system characteristics such as system load, health, or security, in order to detect and identify potential risks. The monitoring system does not place any restrictions on what the monitoring applications can do. The logging system records information about all current client connections. The client side operates as usual unless one server takes over for another server. In this case, the client verifies the security information of the new server, and, assuming the security information is valid, the client will accept messages from the new server and reject all messages from the original server.

CSRI POC: Rolf Riesen, 845-7363

Title: An Evaluation of the Integration of Reconfigurable Hardware with the Network Interface in Cluster Computer Systems

Speaker: Keith Underwood
Clemson University

Date/Time: Thursday, March 14, 2002, 2:00-3:00 pm

Location: Building 980 Room 95

Brief Abstract: Demand for high performance computing continues to remain strong despite the ever increasing performance of commodity desktop machines. One approach to high performance computing uses a cluster of these commodity systems (e.g. a Beowulf Cluster) to provide higher performance. While Beowulf Clusters provide scalable coarse-grain parallel computing, they fail to scale for fine-grain

computations. An- other approach, reconfigurable computing, uses programmable hardware to augment a commodity workstation with high-performance data-flow processing; hence, it exploits fine-grain parallelism but has limited scalability. In many regards, the two technologies are complementary and have much to gain from being integrated together. Unfortunately, a straight-forward integration of the two technologies yields a system that has an even narrower range of application than either technology independently.

The work presented focuses on incorporating reconfigurable computing in Beowulf Clusters effectively. To be effective, addition of reconfigurable computing to a cluster must benefit a wide range of applications. The fundamental issues are the commodity memory subsystem, the commodity interrupt mechanisms, and the commodity I/O bus (PCI) used in commodity systems. When a reconfigurable computing card must share access to these low performance commodity resources with a high-performance net- work card, the result can be drastically reduced performance. The proposed solution is an intelligent network interface (INIC) that incorporates a reconfigurable computing fabric with a high speed network interface to be used in a Beowulf Cluster. Placing reconfigurable computing in this data path can enhance cluster performance by: improving protocol processing performance, providing hardware implementations of collective operations, and providing high-performance data-flow processing of network traffic. The presentation will focus on the ability of an INIC to accelerate applications by shifting application kernels, along with the required protocol processing, onto the network interface.

CSRI POC: Ron Brightwell, 844-2099

Title: Error Estimates and Poisedness in Multivariate Polynomial Optimization and Derivative-free Optimization

Speaker: Professor Luis N. Vicente
University of Coimbra, Portugal
IBM T.J. Watson Research Center, New York

Date/Time: Monday, November 18, 2:00-3:00 pm (MT)

Location: Building 980 Room 95 (Sandia-NM)
Building 921 Room 137 (Sandia-CA)

Brief Abstract: The purpose of this talk is twofold. We show first how to derive error estimates between a function and its interpolating polynomial and between their corresponding derivatives. The derivation is based on a new concept of well-poisedness for the interpolation set, directly connecting the accuracy of the error estimates with the geometry (well-poisedness) of the points in the set. The prior existing error estimates make use of less intuitive geometrical conditions. Our approach provides also an appropriate algorithmic framework to ensure well-poisedness.

In the other part of the talk, we will briefly describe derivative-free optimization techniques based on geometrical concepts like well-poisedness and positive bases, pointing out what are the key ingredients for convergence and making a link to the first part of the talk.

CSRI POC: Tony Giunta, 844-4280

Title: High-Order Unstructured Grid Methods for Time-Domain Electromagnetics
Speaker: Tim Warburton
Department of Mathematics and Statistics
University of New Mexico

Date/Time: Thursday, May 9, 2002, 9:00-10:00 am (MST)

Location: Building 980 Room 95 (NM)
Building 921 Room 137 (CA video link)

Brief Abstract: We discuss the formulation, analysis, and computational performance of a high-order accurate computational method for the time-domain solution of Maxwell's equations. The formulation supports a fully unstructured discretization, using finite-element/finite-volume grids as well as heterogeneous materials.

The scheme employs a high-order nodal basis based on multivariate Lagrange polynomials, offering efficiency, accuracy and a straightforward implementation. The equations are satisfied in a quadrature-free discontinuous Galerkin/penalty fashion, resulting in a highly parallel formulation. The discretization is fully body conforming. Novel far-field truncation techniques have been developed, and latest results will be shown.

This work is done in collaboration with Jan S. Hesthaven and Cedric Chauviere, Brown University.

CSRI POC: David Day, 844-1868

Title: Automatic Algorithm Recognition and Replacement, A New Approach to Program Recognition

Speaker: Zhaofang Wen

Date/Time: Tuesday, February 19, 2002, 11:00 am – 12:00 noon

Location: Building 980 Room 24

Brief Abstract: Parallel computation will become the norm in the coming decades. Unfortunately, advances in parallel hardware far outpaced parallel applications of software. There are currently two approaches to applying parallelism to applications. One is to write completely new applications in new languages. But abandoning applications that work is unacceptable to most nonacademic users of high-performance computers. The other approach is to convert existing applications to a parallel form. This can be done manually or automatically. Even partial success in doing the job automatically has obvious economic advantages. Here we present a fundamentally new theoretical framework for finding poor algorithms in an application program and replacing them with ones that parallelize the code.

CSRI POC: Erik DeBenedictis, 284-4017

Title: Communication Offload MPI-Based (COMB) benchmark

Speaker: Riley Wilson
University of New Mexico

Date/Time: Thursday, September 19, 2002, 9:45-10:30 am

Location: Building 980 Room 95

Brief Abstract: COMB is an MPI benchmark for evaluating the effectiveness of communication offload in a parallel system. It is comprised of the Poll method and the Post-Work-Wait method. These methods have been used to evaluate many systems, including several at Sandia National Labs. Among many other aspects, systems were tested to observe the effect of allowing and disallowing the use of the cache for data.

CSRI POC: William Lawry, 284-4206

Title: Nonlinear Elimination and Function Modeling in Aerodynamic Analysis and Design

Speaker: David Young
The Boeing Company

Date/Time: Thursday, January 24, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: Just as Schur complement methods have been important for practical solution of large linear systems in industry, the nonlinear elimination method while a classical method has been crucial to success in aerodynamic design and analysis. We will outline the important applications of this method in wing design and optimization and discuss in what situations it is preferable to applying Newton's method to the entire system of necessary conditions for optimality. We will illustrate the method using computational examples from airfoil design. Another area we will discuss is the practical aspects of applying design optimization in engineering practice. As with all technological innovations, there is initial resistance that must be overcome. It is necessary to constantly deal with customer issues and to provide interfaces and work hard on reliability to gain customer acceptance.

CSRI POC: Bart van Bloemen Waanders, 284-6746

Title: A Combinatorial Optimization Algorithm for Large-scale Biological Data Clustering and Applications

Speaker: Ying Xu
Protein Informatics Group, ORNL

Date/Time: Tuesday, August 20, 2002, 9:00-10:00 am

Location: Building 980 Room 95

Brief Abstract: I will talk about our recent work on a new general framework for data clustering, based on the concept of minimum spanning trees. This clustering framework has been applied to various biological data analysis problems, including regulatory binding site identification and microarray gene expression data interpretation

CSRI POC: Steve Plimpton, 845-7873

Chapter 6. Fellowships

The Computer Science Research Institute supported two students during CY2002 through theDOE High Performance Computer Science (HPCS) Fellowship administered by the Krell Institute. The objective of the DOE High-Performance Computer Science Fellowship program is to encourage the training of computer scientists by providing financial support to talented students to enter a period of study and research in computer science with an emphasis on high-performance computing, accompanied by practical work experience with researchers at Los Alamos National Laboratory, Lawrence Livermore National Laboratory, and Sandia National Laboratories.

The fellowship program requires a program of study that will provide a solid background in high-performance computing. Examples of research specializations of interest to the program include:

- Parallel and novel architectures, including clusters of SMPs
- Three-dimensional scientific visualization
- High-speed network interconnects
- Parallelizing compiler technology
- Parallel systems software (OS kernel technology, file systems, etc.)
- Performance evaluation and modeling
- Scalable computer security
- Object-oriented scientific programming frameworks
- High-performance computing software component frameworks
- Parallel scalable algorithms

After the start of the fellowship, each participant will participate in a research assignment (practicum) at Los Alamos National Laboratory (LANL), Lawrence Livermore National Laboratory (LLNL), or Sandia National Laboratories (SNL) working with high-performance computing researchers. During the summer of 2002, Brian Barrett, an HPCS fellow studying at Indiana University completed his practicum at Sandia National Laboratories working with CSRI researcher Ron Brightwell.

Students must be United States citizens planning full-time, uninterrupted study toward a Ph.D. degree at a U.S. university. Students who are senior undergraduates or in their first or second year of graduate study in computer science are eligible to apply for the DOE High-Performance Computer Science Graduate Fellowship (DOE HPCSF).

For additional information regarding the Krell Institute and the HPCS fellowship, please see the web site http://www.krellinst.org.

Chapter 7. For More Information

For more information about the CSRI, please contact one of the following:

David E. Womble

Computer Science Research Institute, MS 1110
P.O.Box 5800
Sandia National Laboratories
Albuquerque, NM 87185-1110

dewombl@sandia.gov
(505) 845-7471 (voice)
(505) 845-7442 (fax)

Barbara J. DeLap

Computer Science Research Institute, MS 1110
P.O.Box 5800
Sandia National Laboratories
Albuquerque, NM 87185-1110

bjdelap@sandia.gov
(505) 845-7002 (voice)
(505) 284-4520 (fax)

Deanna R. Ceballos

Computer Science Research Institute, MS 1110
P.O.Box 5800
Sandia National Laboratories
Albuquerque, NM 87185-1110

drcebal@sandia.gov
(505) 845-7753 (voice)
(505) 284-4520 (fax)

Appendix A. Distribution:

External:

Dr. William Reed
DP-141/Forrestal Building
U.S. Department of Energy, Defense Programs
1000 Independence Avenue SW
Washington, DC 20585

Dr. Jose Munoz
DP-141/Forrestal Building
U.S. Department of Energy, Defense Programs
1000 Independence Avenue SW
Washington, DC 20585

Internal:

1	MS 0151	Thomas O. Hunter, 09000
1	MS 0321	William J. Camp, 09200
1	MS 0139	Michael O. Vahle, 09900
1	MS 0139	Robert K. Thomas, 09904
5	MS 1110	David E. Womble, 09214
1	MS 9018	Central Technical Files, 8945-1
2	MS 0899	Technical Library, 9616
1	MS 0612	Review & Approval Desk, 9612 For DOE/OSTI