SAND REPORT

SAND2002-1254
Unlimited Release
Printed May 2002

Computer Science Research Institute 2001 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

SAND2002-1254
Unlimited Release
Printed May 2002

Computer Science Research Institute 2001 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 at Sandia National Laboratories during the period January 1, 2001 to December 31, 2001.

Contents

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.

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.

2.1 Algorithms and Computational Mathematics Focus Area:

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.

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.

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.

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.

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.

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.

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).

2.2 Enabling Technologies Focus Area:

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.

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.

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.

2.3 System Software Focus Area:

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.

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.

2.3.2 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.

2.3.3 **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.

2.3.4 **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.

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

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.

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.

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.

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.

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.

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.

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.

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.

4. Summary of Activities in CY01

During 2001, the CSRI hosted 153 visitors from 75 different universities, laboratories or companies. These visits were hosted by 54 different Sandia staff who participate with the CSRI. In terms of the activities listed above, the CSRI sponsored 14 long-term research projects, 17 summer or long-term faculty visits, 32 summer students, 5 graduate fellows, and 5 workshops/conferences. Detailed descriptions of research projects, sabbaticals and seminars are given in the appendices.

5. For more information

For more information, please contact

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: Research Projects

Computer Science Research Institute
Sandia National Laboratories

Title: Automated Multi-Level Substructuring Technique

PI: Jeffrey K. Bennighof , University of Texas, Austin

Dates: June 1, 2001 – August 31, 2001

CSRI POC: Richard Lehoucq

Project Summary:

This project is concerned with applying our Automated Multi-Level Substructuring (AMLS) technique to the solution of very large eigenvalue problems and, in some cases, directly to applications for which partial eigensolutions are typically sought, such as frequency response analysis of structure finite element models having tens of millions of degrees of freedom or more.

In AMLS, a finite element model is divided into thousands of substructures in a tree topology using nested dissection. The finite element representation is transformed to one in terms of substructure eigenvectors whose eigenvalues are below a given "cutoff" value. The number of unknowns is reduced by orders of magnitude by truncating at the substructure level. Then either frequency response or a partial eigensolution of the original model is approximated using the substructure eigenvector subspace.

AMLS has been tested on dozens of production automobile body models that have millions of degrees of freedom. It delivers ample engineering accuracy on frequency response over broad frequency ranges, or on the partial eigensolution typically used for approximating the frequency response. This is accomplished in less time on a single workstation processor than the conventional block Lanczos eigensolution requires on a multiprocessor Cray T90. Moreover, AMLS is an inherently parallel approach.

Evidently, the AMLS approach can be of great benefit in the next generation of structural analysis problems, which are of interest at Sandia. The two accepted paradigms for solving very large eigenvalue problems are the "shift and invert" block Lanczos algorithm using a sparse direct solver for the iterations, implemented in commercial code at Boeing, and the Lanczos implementation used in Salinas, in which the domain decomposition iterative solver "FETI" (Finite Element Tearing and Interconnecting) is used in place of the direct solver.

Because of its reliance on a sparse direct solver, the first approach cannot be expected to perform well on problems having tens of millions of unknowns. In industry, the practical upper limit on problem size for this approach is generally accepted to be about 2.5 to 3 million unknowns. Solutions are obtained using an "out-of-core" algorithm, which requires the memory bandwidth and I/O capabilities of vector supercomputers to get even reasonably efficient usage of CPU resources.

Because it uses an iterative global solver, the second approach introduces difficulty with maintaining the orthogonality between vectors in the Krylov subspace that is assumed in the Lanczos algorithm. This approach has been used successfully to obtain dozens of eigenpairs, but is apparently not practical for obtaining thousands of eigenpairs because of the expense of solving the projected eigenvalue problem associated with the generated subspace.

Although experience with AMLS is currently limited to obtaining several thousand eigenpairs for models with up to 7.5 million unknowns, the method looks promising for achieving Sandia's goal of obtaining 10,000 eigenpairs for models having 100 million unknowns. The comparison is not perfect, but AMLS has obtained about 1,000 eigenpairs on a problem of order 2.9 million on a single processor in about the same amount of time that Salinas has required to obtain 20 eigenpairs on a problem of order 1 million on 128 processors.

I will spend one and one-half months at Sandia this summer collaborating with Dr. Rich Lehoucq and investigating the potential of AMLS for Sandia's applications. I will also work with Rich on numerical analysis which would provide a firmer theoretical foundation for AMLS.

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

PI: Professor Alok Choudhary, Northwestern University

Investigators: Dr. Wei Keng Liao, Research faculty
Professor G. Thiruvathukal
Two graduate students

Dates: November 1, 2001-December 31, 2002

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.

Tasks:

1. Identify the Overlapping Regions from MPI Derived Datatype
   Study of finding the overlapping access on the MPI-IO derived datatype level.
2. Contiguous Region in File Access
   Study the case of contiguous file access from all processors. Implement this task in the unit of bytes in order to take care any possibility of user defined derived datatypes.
3. Non-contiguous Region in File Access
   Extend the previous task to cover more complex file access. Devise a solution in byte level to cover any types of data mapping between file and memory buffer. 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 E. Dennis, Jr.,
Rice University

Dates: July 26, 2001 – June 30, 2004

CSRI POC: Juan Meza, Bill Hart

Project Summary:

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 Juan Meza, 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.

Title: Optimization Algorithms for Networks and Scheduling

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: Uncertainty Characterization and Management in Computational Mechanics

PI: Roger G. Ghanem, The Johns Hopkins University

Dates: July 1, 2001 – August 31, 2001

CSRI POC: John Red-Horse

Project Summary:

Given the growing significance of this uncertainty quantification (UQ) within Sandia, I expect that I will be devoting a significant portion of my stay at Sandia to broaden my perspective on UQ, and to better implement some of the techniques I have been developing into Sandia's analysis software. In the context of UQ, I would very much like, if you deem appropriate, to give a short course during my stay at Sandia on the subject of Uncertainty Characterization and management in Computational Mechanics. This should provide me with the chance for an informal interaction with the UQ group at Sandia, which should help me integrate the developments at Sandia into my research and teaching. I am also very much interested in the development of Sierra,and would like to interact with its architects to learn more about its features and implementation. I am particularly interested in the manner in which UQ has been integrated into Sierra.

Additionally, I am also engaged in research in the area of remote sensing with applications to vulnerability assessment against hurricanes, earthquakes, landslides, and other natural disasters. I look forwad to the possibility of interacting with the SAR group at Sandia and adapting from them novel techniques in the area of remote sensing that could significantly affect our research. Miss Malek is currently assisting me in this project, and she would be one of the students to accompany me to Sandia.

One other R&D area of interest to me that I would like to interact on with Sandia engineers relates to Information Technology. In particular, I have recently started an effort trying to integrate and manage distributed analysis software packages using CORBA. I am aware of one group within the Computational Physics Division at Sandia that was engaged in similar research, integrating Alegra and Coyote using CORBA. Developing standards for propagating uncertainty in such computational systems is essential for their usage in a distributed computational network, in particular when applied to the analysis of complex engineering problems.

Title: Terascale Simulation-Constrained Optimization

PI: Omar Ghattas, Carnegie Mellon University

Investigators: Larry Biegler and Anthony Kearsley, Carnegie Mellon University
Matthias Heinkenschloss, Rice University
John Tolle, University of North Carolina

Dates: June 1, 2000 – 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.

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.

Tasks:

The focus of our efforts will be to:

- Target steady state as well as time-dependent PDE constraints
- Seek SAND (Simultaneous Analysis and Design) methods that exploit PDE structure of constraints
- Develop methods that can scale to:
    - large numbers of state variables and state constraints
    - moderate to large numbers of decision variables and inequalities
    - large numbers of processors
- Integrate into modular simulation environment (SIERRA)
- Capitalize on existing and emerging parallel PDE solver libraries (AZTEC, PETSc)
- Demonstrate on well-chosen applications in the areas of:
    - optimal design and control
    - parameter estimation

Specific tasks that support our goals include:

- Investigate SIERRA as an environment to integrate optimization into.
- Evaluate PDE-based simulation codes under SIERRA as candidates for optimization.
- Determine suitable case studies.
- Develop and adapt SQP (Sequential Quadratic Programming methods) for steady PDE-constrained problems in the following order:
    - no inequality constraints
    - simple decision variable inequalities
    - decision and state inequalities
- Develop and adapt SQP for unsteady PDE-constrained problems in the following classes:
    - no inequality constraints
    - simple decision variable inequalities
    - decision and state inequalities
- Demonstrate SQP methods on challenging problems of interest to Sandia
- Organize workshops on PDE-constrained optimization to encourage cross-fertilization between optimization and simulation communities

Research Issues:

The project tasks engender the following research issues, which we will be addressing:

- Integration of optimization with PDE solver codes that do not form the true Jacobian. Many PDE codes do not form the exact Jacobian matrix of the PDE; often cheap approximations of the Jacobian are constructed at the expense of larger number of iterations. These approximations include using lower order spatial approximations than in the residual, avoiding differentiation of complex constitutive laws, and abandoning representation of the matrix (as in matrix-free Krylov solvers that approximate the matrix-vector product by directional differencing). However, optimization requires the exact Jacobian of the PDE and its adjoint in the optimality conditions (unlike the PDE solver, which requires only the residual). While optimization methods may be able to accommodate inexact Jacobian information in some stages of the optimization, the level of inaccuracy has to be controlled by the optimization and is tightened as the optimization iterates approach the solution. Coarse Jacobian approximations, which are appropriate for PDE solves, are usually insufficient for the optimization. One remedy is to employ the directional differencing idea whenever Jacobian-vector products are needed, but this doesn't work for adjoint methods, since the transpose of the Jacobian is needed. Accommodating approximate Jacobians is essential for successful integration of optimization with contemporary PDE solver codes.

- Incorporation of ad hoc PDE solver globalizations into the optimization framework. Traditionally, optimizers have used line search and trust region methods to assure global convergence. However,

many PDE solver codes use other domain-specific techniques to assure robustness, including multiple grid sequencing (coarser grids exhibit lower degrees of nonlinearity), parameter continuation (e.g. on Reynolds number), and pseudo time-stepping (solving a time-dependent problem to reach the steady state). It is desirable to integrate these ad hoc techniques with line search and trust region optimization globalizations.

- Efficient use of second order derivative information in the optimization. Integration of sufficiently good second order derivative information is crucial for the performance of the optimization methods considered in this proposal, but poses interesting challenges. For example, quasi-Newton methods often do not scale well with respect to the number of decision variables, M. Only if applied consistently with the PDE formulation and suitably initiated, they can be expected to show a convergence behavior that is independent of the discretization level. In some cases this has been proven analytically. Enforcing consistency and finding an appropriate initialization, however, can be difficult for complex problems. Mesh independence results are not available for limited memory quasi-Newton methods used in practice. Newton methods, on the other hand, promise a number of iterations independent of M. However, M linearized PDE solves (with differing righthand sides) are required to form the reduced Hessian at each iteration in a Newton rSQP method. This is particularly onerous when Krylov methods are used to solve the PDEs, as is typical for large scale problems on parallel machines. Iterative linearized PDE solves introduce inexactness in the Hessian. Another difficulty of Newton methods is obtaining second derivative information. One possibility for getting around these problems is using matrix-free Newton-Krylov methods, preconditioned with quasi-Newton approximations and approximate Jacobians. This combines the rapid convergence of Newton methods with the low per-iteration cost of approximate methods. This and other combinations of limited memory quasi-Newton methods and matrix-free inexact Newton methods need to be explored. Preconditioners for the large scale linear optimization subproblems need to be investigated.

- Handling very large numbers of inequality constraints. A particular difficulty of PDE optimization problems is that inequality constraints can be posed pointwise (e.g. temperature or stress might be bounded throughout the domain), which means that the number of inequalities can scale with the state variable dimension. Moreover, strict complementarity assumptions often imposed for finite dimensional, non-PDE based problems can typically not be expected to hold in this context. As the discretization is refined, strict complementarity will often be lost. For some state-constrained problems, strict complementarity is never satisfied. There is a need for optimization methods that scale well with large numbers of inequality constraints provided by PDE problems. Many active set methods, for example, have been known to scale combinatorially with the number of inequality constraints, which is disastrous for our problems. Interior-point methods, on the other hand, have been demonstrated to scale quite well for some problems. However design and convergence analyses of interior-point SQP methods is an active research area. In addition, for some PDE based optimization problems, active set and projection methods tailored to the problem have been successful.

- Aggregating inequalities to improve the performance of adjoint methods. Adjoint sensitivity methods require a linearized PDE solution for each active constraint, whereas direct methods require a solution for each decision variable. When there are few active constraints, adjoint methods are greatly preferred. It would be very desirable to preserve this advantage of adjoint methods in the presence of many active constraints. One way to do this is to aggregate inequalities, and in fact this is what interior point methods do. A promising alternative for aggregating constraints is the KS (Kreisselmeier and Steinhauser) function, which reduces large numbers of inequality constraints to just a few. KS functions can be derived that provide a convex over-approximation of the feasible region. Moreover, through the adjustment of the aggregation parameters, this region can be tightened to arbitrary accuracy. In addition to preserving the efficiency of adjoint methods in the presence of many inequalities, KS aggregation will be useful in addressing the combinatorics of inequalities.

- Permitting inexact Jacobian solves within optimization iterations. In a SAND method such as reduced SQP, the nonlinear PDE simulation does not have to be converged at each optimization iteration--only linearized PDEs are satisfied. Avoiding nonlinear solution in intermediate iterations is a big win. An even bigger win is to avoid exact solution of the *linearized* system. The reason is that for most state equation solvers, the number of (nonlinear) Newton iterations is usually independent of mesh size, whereas the number of (linear) Krylov iterations increases with mesh size for most popular preconditioners. It is highly desirable, therefore, to accommodate inexact solution of systems involving the Jacobian of the state equations. Additionally, in many cases exact Jacobians are too difficult or expensive to compute, and one prefers to approximate them. Control of inexact Jacobian solves and of inconsistent Jacobians within an SQP method will be crucial for the performance of the SQP optimization. There are a few approaches that deal with these issues (inexact Newton-SQP methods, trust region globalization allowing inexact linear system solves). The existing approaches require that the linear systems can be solved with an accuracy specified by the SQP method. The SQP method will tighten the accuracy requirement as the iterates approach the solution. However, much work remains.

- Addressing time-dependent PDEs within optimization. Time-dependent PDE optimization is a difficult problem. The optimality conditions include a PDE for the adjoint variables (or Lagrange multipliers) which is a final value problem, i.e. it is integrated backward in time from a final value of the adjoint variables. The difficulty is that the righthand side for this adjoint problem depends on the state variables, which are computed through an initial value problem by integrating forward in time. The most work-efficient way to do this is to store the entire time history of the state variables, and then invoke them in reverse chronological order when solving the adjoint problem. Unfortunately, storage of the time history of the states is impossible for large-scale problems--memories on parallel systems are designed for a small number of state variables vectors, since parallel PDE solvers usually require O(N) storage. The storage-efficient alternative (recompute the states forward in time as needed at each time step while integrating the adjoints backward in time) is terribly work-inefficient, requiring O(T) forward solves, where T is the number of time steps. Fast compression schemes are needed to overcome this basic problem of memory vs. work inefficiency. In addition, multiple shooting approaches, carefully adjusted to the PDE context, can be used to decouple the problem and offer additional ways to balance work and storage requirements.

  An additional difficulty occurs when the decision variables are time-dependent (e.g. when they represent boundary controls). A dense reduced Hessian arises, which is problematic because it couples decision variables at all time steps. Again, this is contrary to the principles of linear memory scaling in parallel PDE solvers, in which just one or several time steps are held in memory and operated on at any one time. One possible means of overcoming this problem is a block decomposition, with each block corresponding to a time step. A quasi-Newton update of each block is then possible.

Title: Research in Finite Element Simulations

PI: Max D. Gunzburger, Iowa State University

Dates: May 1, 2001 – April 30, 2002

CSRI POC: David Womble

Project Summary:

There are a variety of potential research projects and collaborations that I will 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 invoived, 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 unertainty, 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 prcatical 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 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.

Other potential activities

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

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 advisingthe 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 leecture 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: Resource Management for Heterogeneous Clusters

PI: Steve J. Chapin, Syracuse University

Investigators: Dimitris Katramatos, University of Virginia
Marty Humphrey, University of Virginia

Dates: March 23, 2000 – May 31, 2002

CSRI POC: Neil Pundit

Project Summary:

Clustered systems have become an attractive building block for supercomputing and high- performance distributed computing. To obtain maximum computing power, and to gain full leverage from existing computing systems, these clusters are often combined into larger distributed systems. These systems are called heterogeneous or federated clusters, and are a significant special case of metacomputing (or grid computing). Heterogeneous clusters do already exist in real systems, such as the Computational Plant at Sandia National Labs.

These systems represent a prime opportunity for high-performance computing, and we understand how to assemble the hardware to build large clusters. As is (unfortunately) usual, software maturity has lagged, particularly for cluster management and administration. For example, standard message-passing libraries such as MPI allow programmers to construct programs that can run on several platforms, but there is essentially no support for mapping that computation onto a particular cluster. Such a mapping might want to take into account differences in machine architecture/speed, the interconnection network topology, or the number of processors per machine.

Unfortunately, this often has adverse effects on application efficiency because the program decomposition may not map well onto the cluster architecture (or equivalently, the default mapping software may not do a good job of mapping that program onto the architecture). While the program decomposition may be expecting a mesh, the actual network may be a combination of crossbar switches connected in an alternative topology (e.g. ring network or modified hypercube). In such a situation, a naive placement might place processes that communicate frequently (those adjacent in the mesh) on topologically distant nodes.

To avoid these situations, we need to devise mechanisms to properly map computations onto system configurations. These mechanisms must include 4 basic components:

- A representation of the communication properties of the computation including the process decomposition, communication topology, frequency, size, and pattern of communication.
- A representation of the system, including static properties such as processor type, configuration, network topology, as well as dynamic properties such as load.
- Agents for collecting the state information and making it accessible to allocation agents.
- Allocation agents to map programs to available resources.

What we have here, in fact, is a special case of the grid computing problem, and we can apply the lessons learned building grid computing systems in the less chaotic environment of clustered systems. This means that our solutions will be simpler and more efficient than general-purpose grid computing mechanisms. While our solution will be independent of any particular grid computing system, we will build it so that it can be integrated with the two major grid computing systems (Globus and Legion).

As proof of concept, we will develop prototype mechanisms and demonstrate their efficacy by using them to schedule the MP-LINPACK benchmark suite and applications that use the MP salsa library. This supports our focus on making broad classes of programs use clusters more efficiently, rather than attempting optimization tuning for a single program.

Problem Description

Clusters can vary widely in configuration, depending on the number of nodes, processors per node, processor architecture, interconnection network, and network topology. The C-plant cluster reduces much of this variability, but there are still significant differences between portions of the cluster (cf. the redundant 2-D torus topology of Siberia vs. the modified recursive hypercube of the Alaska subcluster). The result is that what is logically one cluster is really at least four sub-clusters combined into one (with ongoing acquisitions, this number will only increase). When examining a machine or system, it is important to understand the regularity of the communications infrastructure and of the processors. Likewise, it is important to know whether the program has a regular decomposition. One of the most important factors in determining application performance is the mapping of the computation onto the architecture, i.e. deciding which of the job's tasks will run on which computational nodes. An efficient mapping will closely match the application's communications patterns to the underlying communications topology, while a poor mapping will not (cf. our earlier example of a mesh-based computation being mapped onto a hypercube).

In parallel machines, such as hypercubes or the Intel Paragon, the processors are homogeneous and the communications network is extremely regular; this vastly simplifies the process of mapping regular programs onto the architecture. On the other end of the spectrum, general distributed systems such as the Internet have no regularity in their structure. Clusters represent a middle ground they have semi-regular communication structures, and usually have only a few architectures.

In the same way that parallel machines exhibit regular structure, parallel applications usually have regular structure. One of the most common structures is five-point stencil, wherein the new value of a data point is a (possibly computationally complex) function of its current value and the current value of each of its four neighbors in a grid (hence the name 5-point stencil). Other structures require regular communication in a three-dimensional topology, or in some cases, all-to-all communication (each process must communicate with every other process).

There are two obvious issues here related to the efficiency of execution of the applications program. The first is the compatibility of the program decomposition with the system topology, and the second is how well that compatibility is exploited via the mapping of processes to processors in the system.

For example, consider a five-point stencil. If we are dealing with a mesh architecture, there is a direct mapping of stencil processes onto mesh processors. Thus, the five-point stencil is perfectly compatible with a mesh architecture. However, we must still properly perform the mapping to exploit that compatibility. A poor choice of mapping could result in poor application performance. Thus, the program and system structures dictate an upper bound on the performance we can achieve, but it still requires proper mapping support to achieve that performance. In the past, this mapping has been done by hand it is the intent of this proposed work to automate this process.

Tasks:

To provide proper mapping support, we propose the following solution:

- We will devise a standard representation for application descriptions, capturing factors such as computational needs and communication patterns, topology, and size.
- We will devise a standard representation for system structure, including architecture, communication topology and characteristics, and software configuration.
- We will develop cost models to assist in the mapping of applications to systems. The cost modeling functions will take both application and system descriptions and estimate the cost of performing particular operations (such as creating threads, migrating processes, etc.).
- Each SU will run a management daemon on the SS0 node, which will maintain up-to-date status information for use in scheduling decisions.

- An allocation manager will coordinate with the SU management daemons to maintain a snapshot of the system state. When a job is submitted to the system, the allocation manager can invoke the cost functions to determine a good mapping for the job tasks and to allocate resources. This does not have to be a single persistent daemon, although conceptually it is. We will replicate this so that there is no single point of failure in the system.

If time permits, we will develop schedulers that are separate from the allocation manager, allowing the allocation manager to focus on resource control while the schedulers focus on mapping tasks to resources. We will also take advantage of existing software, such as PBS or the Maui Scheduler, to handle some of the allocation management issues (in particular, Maui or PBS might provide some equivalent functionality to that of the management daemon). Future research in this area will encompass economic models for scheduling, and contract-based negotiation for the use of resources.

The software components we have briefly described are independent of any particular control system (such as Globus or Legion). We do not intend to integrate our initial prototype with either system, but our design will be intended for later integration.

We have two clusters available for development of our software. The first is the Orange Grove cluster at Syracuse University, which is a heterogeneous cluster primarily comprising 16 single-processor Alpha machines and 48 dual-processor Pentium machines, connected by Myrinet and Fast Ethernet. We also have the Centurion cluster at the University of Virginia available for our use. Like Orange Grove, Centurion has both Alpha and Pentium processors and is connected by Myrinet and Fast Ethernet. The two clusters are connected by the vBNS. After initial development and testing on our clusters, we can deploy the software on Sandia's clusters during the summer and fall of 2000.

Title: Benchmarking OS Bypass Implementations of MPI

PI: Arthur B. Maccabe, The University of New Mexico

Investigator: Bill Lawry, The University of New Mexico

Dates: August 6, 2001 – December 31, 2001

CSRI POC: Ron Brightwell

Project Summary:

This project will address the development of a portable benchmarking tool to compare implementations of MPI based on OS Bypass. This tool will focus on the degree to which the OS Bypass implementations improve the availability of the host processor for application use. In addition to providing a basis for comparing different approaches to OS Bypass, this tool will be important in guiding the OS Bypass implementation of Portals 3.0.

OS Bypass is commonly recognized as feature of many high performance communication protocols including: Portals 3.0 (a message passing protocol developed jointly between Sandia National Labs and the University of New Mexico), GM (Glenn's Messages, developed by Myricom), and VIA (the Virtual Interface Architecture). The essential goal of OS Bypass is to reduce the CPU overhead associated with the implementation of a communication protocol by offloading parts of the protocol implementation to a programmable NIC (Network Interface Card). Implementing part of the protocol on the NIC means that the NIC does not need to involve the host processor in the processing of every incoming packet. In concept, incoming data packets "bypass" the operating system. In reality, we have simply moved the appropriate policies of the operating system from the host processor to the NIC.

To see the importance of OS Bypass, consider a gigabit networking fabric based on 1500 byte packets. At peak bandwidth, the inter-arrival time for network packets is 12 μsec. When we consider that interrupt latencies on a PCI bus (the time it takes the processor to respond to an interrupt from a device on the PCI bus) are approximately 10 μsec, one problem becomes evident: the network will saturate the interrupt handling capability of the PCI bus. Moreover, when we consider that the interrupt latency of the Linux kernel (the time it takes the Linux kernel to respond to an interrupt once the interrupt has been presented at the CPU) is approximately 3 μsec, another problem becomes evident: a minimum of 25% of the host processor will be dedicated to handling incoming network packets. These problems will on become more pronounced as network speeds continue to increase. Two alternatives to OS Bypass have been proposed an implemented: jumbo frames and interrupt coalescing. Jumbo frames increase the frame size from 1500 bytes to 9000 bytes. This decreases the interrupt pressure by a factor of six, when end-to-end bandwidth is the critical factor affecting application performance. There are two issues that limit the benefit of jumbo frames as network speeds increase: the amount of data transmitted in a typical message, and the physical page size of the host system. To see the advantages of jumbo frames in a 10 gigabit network, the frame size would need to be 90,000 bytes. If we continue to grow the frame size, the actual frame size will be dominated by the expected message size and we will no longer see any benefit from jumbo frames. In addition to the expected message length, frame sizes are naturally limited by the page size of the host system because this is the largest unit that can be used in a single DMA transfer.

In interrupt coalescing, the NIC throttles the rate at which it interrupts for the host processor by delaying interrupts until a fixed number of packets have arrived or until a timer expires. This reduces the interrupt pressure, but still requires that the host processor process every incoming packet. Moreover, small incoming messages are now subject to a high variability in their latency depending on the state of the NIC when the packet arrives. This variability in latency can have a significant impact on application performance.

While a number of research groups have put forward the arguments presented in the previous section in support of OS Bypass techniques, none has developed benchmarks to demonstrate the effectiveness of an

OS Bypass implementation. Interestingly, the only network benchmarking tool that reports CPU utilization is *netperf*, a tool designed to benchmark implementations of the IP (Internet Protocol) suite, a protocol suite that does not lend itself to OS Bypass.

We will develop a portable benchmark tool that will allow us to contrast implementations of MPI (the Message Passing Interface) based on OS Bypass. We have chosen to benchmark MPI because of its obvious importance in high performance scientific computing. Because the goal of OS bypass is to retain high communication performance while minimizing the impact on the host processor, we will use sustained bandwidth weighted by host processor slowdown as the basis for comparison. Bandwidth will be measured using the standard ping-pong test.

Host processor slowdown will be measured in a way that is similar to the CPU overhead measurement implemented by *netperf*. In *netperf* this is a two step process: first, a simple counting application is run on an unloaded system for a fixed period of time; next, the same counting application is run for the same period of time, while the maximum sustained bandwidth is being measured. CPU overhead is reported as the ratio between the two counts.

While this methodology is appropriate for protocols in the IP suite, it is not appropriate in the context of MPI. Because the second measurement involves two processes (a counter process and a communication process), the measured overhead includes the cost of process context switches. In the context of MPI, it would be better to integrate the counter measurement into the communicating process so that there is a single process per compute node. Unfortunately, this will introduce a polling overhead that is not easy to account for. The first task of this research will be to identify an appropriate methodology to measure the two counts: the count attained during the bandwidth measurement and the count attained when no communication is in progress. Once this step is complete, the metric we will use to compare MPI implementations is bandwith * count during bandwidth measurement/count on an unloaded node

Parameters to the benchmark tool will include: the size of the message used in the ping-pong test, whether the receive message is pre-posted, the measurement time, the minimum number of measurements to be taken, and the desired confidence level of the results. Outputs will include means, standard deviations, minimums, and maximums for the two counts and the sustained bandwidth.

In addition to providing a good basis for comparing MPI implementations, this benchmark tool should help evaluate design alternatives in the OS Bypass implementation of Portals 3.0. The current generation of programmable NICs have slow processors (less that 100 MHz) and limited local memory (up to eight MB). The Portals 3.0 API was designed to enable offloading almost all of the communication protocol onto the NIC. Due to limited resources, it is clear that only a part of the communication protocol can actually be offloaded.

Tasks

- Develop benchmark tool. In this task, we will complete the design of the overhead measurement methodology and develop an implementation of the benchmark.
- Measure OS Bypass implementations of MPI. In this task, we will use the benchmark to compare the performance of MPI. As a minimum, we will compare a Portals based implementation with a GM based implementation, both on Myrinet.

Title: Development of algorithms for the determination of medial surfaces

PI: Sunil Saigal, Carnegie Mellon University

Dates: June 11, 2001 - September 30, 2001

CSRI POC: Scott Mitchell

Project Summary:

The project relates to the development of algorithms for the determination of medial surfaces of arbitrary three-dimensional geometries. Medial surfaces are required for efficient mesh generation for finite element analysis of large components. Medial surfaces may additionally be used for cleanup of CAD geometries and to determine the complexity of the mesh generation operation *a priori*. Since large components are to be analyzed under the ASCI program that may require considerable computational effort, parallel implementations of the algorithms developed will also be addressed under the program initiated with CMU. The developments will be made within the existing Sandia mesh generation toolkit called CUBIT. The basic medial surface algorithm to which enhancements will be made has recently been acquired by Sandia.

The purpose of Dr. Saigal's visit to the CSRI is three fold. First, as his collaboration with Sandia on mesh generation is just beginning, he will work directly with researchers to learn and document their development styles and practices so that he and his students may cleanly integrate their work with Sandia's. It is important that the research results, obtained by Dr. Saigal and his students, be fully integrated back to Sandia. This is an important aspect of establishing a longer-term research project, which both Dr. Saigal and the meshing group seek. He will also begin research on medial axis transformations, including the possibilities of extending current methods to parallel regimes. Second, he will learn first-hand the meshing problems from a Sandia perspective to help guide future research projects. Sandia's high speed computing environment makes it's meshing needs somewhat unique from industrial partners he has worked with in the past; he now seeks to more fully understand these different issues. Working with Sandia users and taking the CUBIT training class along side them will be an integral part of achieving this understanding. Finally, where possible, he would like to discuss material modeling, and MEMS techniques with Sandia researchers, and share results from his research with them.

Title: Large Scale Eigenvalue Methods and Balanced Model Reduction

PI: Dan C. Sorensen, Rice University

Dates: May 24, 2000 – December 31, 2001

CSRI POC: David Womble

Project Summary:

Over the past ten years we have been developing methods and software for large scale eigenvalue problems. Our approach, called the implicitly restarted Arnoldi method, underlies our ARPACK software. This package is widely used at Sandia and throughout the world. Hence it is important to continually improve and enhance the package with new features and capabilities.

We intend to improve on the existing deflation schemes, develop a more efficient implementation for symmetric and Hermitian problems, develop a Cayley transformation template that is designed specifically for stability and bifurcation analysis of dynamical systems, include templates for implementing spectral transformation through sparse direct factorization, and also through pre-conditioned iterative solvers.

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.

A fundamental model reduction problem involves systems of the form $x' = Ax + Bu$, $y = Cx$, where A, B, C are real n by n, n by m, and p by 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 shall investigate a model reduction method for such systems through low rank approximation of certain system Grammians. Our focus will be on the development of methods and software for (very) large systems.

We have developed a new computational approach to balanced model reduction that has a high potential for addressing some fundamental difficulties with existing dimension reduction techniques. These issues are central to the development of robust and widely applicable software. We intend to develop a computational methodology that will address the following problems.

- Provide rigorous bounds on the error in the response of the reduced system.
- Naturally preserve fundamental system properties such as stability.
- Be fully automatic once a desired error tolerance is specified.

Although these topics have been addressed, especially for low-dimensional systems, several issues remain open, in particular for very high-dimensional systems. We intend to investigate subspace projection methods, and to develop an (implicit restarting approach) that will (iteratively) produce best approximations of specified rank k to controllability and observability Grammians. These approximations are produced in factored form and can be used to construct a balanced model reduction of specified accuracy or dimension. Computational experience with model problems possessing attributes similar to those arising in applications, indicate very low rank approximations may be possible in practice.

Title: Processor in Memory Enhanced MPP System Architecture for Acceleration of Data Intensive Applications

PI: Dr. Thomas Sterling, California Institute of Technology

Investigator: Maciej Brodowicz, California Institute of Technology

Dates: August 1, 2001 – July 31, 2002

CSRI POC: Neil Pundit

Project Summary:

Massively Parallel Processing systems are being developed to provide tens of Teraflops and are being planned to deliver 100 Teraflops and more for DOE mission critical application domains. The advanced microprocessors employed in these structures are excellent at performing register-to-register numerical operations but are less effective at data-intensive memory-oriented operations due to the long communications latencies and deep memory hierarchy. As a result, efficiencies less than 10% are often experienced running real world application codes on these big machines. As planning for the next generation of MPPs proceeds, the possibility of incorporating a new class of semiconductor device, Processor-in-Memory (PIM), that merges CMOS logic with DRAM memory blocks on the same die, may be considered as a possible means of addressing this disparity and greatly improving future system response time and efficiency. This proposal is for a two year applied research project that will investigate the means and effect of integrating smart memory within the memory hierarchy of future DOE MPPs to greatly facilitate execution of major interdisciplinary codes. The project will extend the work of the Gilgamesh MIND architecture project being conducted at Caltech's Center for Advanced Computing Research. The goal of the proposed project is to reduce the response time of an MPP incorporating MIND PIM chips by a factor of 2 compared to an equivalent scaled system employing conventional "dumb" DRAM memory.

Processor-in-Memory (PIM) semiconductor technology makes possible the merger of CMOS logic and DRAM memory on the same silicon die to enable a new class of processing structures. Such architectures can be employed in scalable arrays on a stand-alone basis or as an embedded capability within large MPPs replacing some or all of the DRAM main memory with PIM based smart memory. At very small cost in terms of the chip real estate, the classical "von Neumann bottleneck" (the historic separation between processors and their main memory) can be significantly mitigated if not eliminated along with many of its deleterious effects through the use of PIM technology.

The Gilgamesh project is developing a new generation of PIM architecture (MIND or "Memory, Intelligence, and Networking Devices") that incorporates mechanisms critical to the manipulation of memory in support of data intensive operations including those related to irregular data structures. The MIND chip architecture, unlike most other PIMs, supports the direct handling of virtual pointers in the memory itself, providing in situ translation of virtual to physical addresses. MIND also supports direct inter-memory chip communication, avoiding the bottlenecks of using the shared system memory bus and the system processors for data memory movement control. In addition, MIND incorporates an innovative hardware multithreaded instruction issue mechanism simplifying resource management and hiding chip wide latencies. Scalable arrays of MIND chips are enabled by an active message transport layer and a message driven task instantiation protocol known as *parcels*. Distributed shared memory address translation mechanisms simplify efficient global data access. Together, these enhanced features make MIND a flexible and powerful memory component for diverse roles from embedded computing in flight and spaceborne missions to data oriented accelerators for mainstream server (e.g. data base, web search) and high-end technical computing structures such as MPPs.

Massively Parallel Processing (MPP) systems are being developed that will extend the peak performance regime beyond 100 Teraflops. These systems integrate high-speed COTS microprocessors with high-speed moderate latency networks in large configurations of thousands of nodes. Such mammoth systems are

efficient at performing floating-point register-to-register arithmetic operations but may be far less effective at global memory data manipulation, especially related to irregular data structures with low temporal locality access patterns. As a consequence, overall efficiencies of such systems can decline below ten percent for important real world application codes. To date, no attempt has been made to overcome this deficiency through a synthesis of MPP and PIM architectures, although the possibility has been investigated by both the HTMT and DIVA research projects. The purpose of such an integration would continue to exploit the cost and performance properties of COTS while taking advantage of the logic embedded in the PIM memory for the data oriented operations that the conventional system microprocessors do inadequately. This proposal solicits sponsorship of a two-year project to investigate the potential of employing PIM technology at the main memory layer of the memory hierarchy in lieu of conventional DRAM. The goal of the proposed project is to achieve a real reduction in overall response time of a fixed scale MPP augmented with MIND chips of a factor of two.

The strategy for accelerating computations on large scale MPPs is to replace some or all of the main memory DRAM with PIM chips based on an enhanced MIND architecture and perform time-consuming data manipulation operations locally within the memory. Performance benefits accrue through several factors including orders of magnitude increase in available memory bandwidth, highly parallel fine-grain processing, reduction in latency between memory row buffers and on-chip logic units, data related overhead operations performed by memory chips and not system processors, and reduced contention for shared system resources.

By migrating data intensive operations from the system processors to the smart MIND memory chips, the processors are either free to perform other work for which they are more suitable or, in the case of time critical data tasks, experience greatly reduced wall clock time required to perform such tasks compared to performing such tasks by the system processors alone. Examples of data oriented operations to be performed within the MIND array main memory are the following:

- Rapid data search and replace
- Zero-time locality actions, i.e. touch once operations
- Associative searches for relational databases
- Pointer-chasing and tree walking, for irregular data structures
- Parallel prefix operations
- Gather and scatter operations
- Sparse matrix operations
- Data structure and matrix transpose
- Global synchronization
- Proactive presaging of work in system processor caches through “percolation”
- Adaptive mesh refinement
- Memory management, demand paging, and some load balancing
- Histogramming

This project will investigate the space of opportunity afforded by the potential heterogeneous processing structures of conventional MPPs incorporating advanced PIM-based smart memories. The objectives of the project are:

1. Estimate potential performance benefit of PIM enhanced MPP for example mission-related applications,
2. Develop architecture enhancements to Gilgamesh MIND chip architecture to support heterogeneous operation in the context of MPP system structure, and test with FPGA-based prototype,
3. Derive software methodology for decoupling and coordinating actions by system processors and MIND logic related to compiler and operating system,
4. Implement, test, and evaluate PIM-based MPP concept through empirical studies with prototype.

Two application kernels will be selected by the sponsor as a basis for study and evaluation. It is likely that one will relate to genome mapping while the other will employ adaptive mesh refinement. Actual existing codes will be employed and modified to work with the new scheme and execution traces will be used as a

basis for estimating the level of benefit that is feasible. An overall end-to-end response time reduction of 50% (a factor of 2X gain in sustained performance) is the target goal for real world applications compared to conventional MPP structures without MIND enhancement. Traces of the selected applications will be analyzed, partitioning the work between system processors and MIND logic depending on best suitability. Estimates of anticipated performance gain will be derived to assess the relative merit of the proposed concept.

The base architecture of the Gilgamesh MIND will be extended to permit integration as part of an MPP system architecture. This requires the addition of a system memory bus interface and control protocol. The existing MIND prototype, under development using SRAM and FPGA components will be modified to incorporate an external bus interface. This will probably be a PCI bus interface to permit easy connection to existing hardware systems and will provide the framework for conducting software experiments using COTS hardware and software components for low cost and rapid progress. Two prototype cards will be manufactured in each of the two years for testing. They will be interfaced with existing PC/workstations in a cluster configuration for proof of concept. This work will leverage on-going efforts and research investment to support this effort through a cooperative arrangement with the Gilgamesh project.

While the necessary code modifications will be made by hand for the purposes of experiments to be performed and to minimize software development costs, a detailed analysis of the compiler and software functionality required to support this form of decoupled computation will be conducted. These include:

- Compiler allocation of division of labor between system processors and MIND logic,
- Coordination and synchronization code between system processors and MIND logic,
- Consistency mechanisms between caches and TLBs of the system processors and the MIND data storage, and
- Coordination the operating system directory table and the MIND logic address translation mechanisms,
- System initialization and booting that now must include MIND resources, and
- System exception handling.

A specific means of accomplishing each of the above will be developed to ensure feasibility of the approach. Some limited software experimentation will be performed, in particular as it relates to the prototyping experiments.

First year tasks:

1. Select benchmark applications in consultation with Sandia and Cellera. Acquire or create kernels and produce representative traces for studies.
2. Analyze benchmark kernel traces to devise partitioning of activities between system processors and MIND logic.
3. Derive an estimate of performance gain of MIND augmented MPP system architecture.
4. Map kernel codes onto MIND instruction set architecture and modify this ISA if necessary to enhance performance for Sandia/Cellera codes.
5. Develop MIND prototype system memory bus interface and implement with second-generation units.

Second year tasks:

1. Port code segments to MIND prototype to validate feasibility of approach and perform execution experiments of decoupled computation on single node and multi-node configurations. This will be done by hand with only basic software support.
2. Develop detailed specification for system software enhancements required for future real world operation. These include desired compiler optimizations.
3. Develop advanced memory management methodology for future integration of operating system (e.g. open source Linux) and consistency mechanisms.
4. Provide detailed assessment of all aspects of the proposed concept, its implementation, experimental results, and future work.

Title: Pattern Search for Nonlinear Optimization

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

Dates: March 6, 2001 – December 31, 2001

CSRI POC: Tammy Kolda

Project Summary:

This project will continue the development of pattern search algorithms for nonlinear optimization problems of the form:

minimize $f(x)$
subject to $x \in S \subseteq \Re^n$,

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. 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, has already been published.

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. 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.*

Title: Analysis of message passing environments on large clusters, and future directions

PI: Pete Wyckoff, Research Scientist, Ohio Supercomputer Center

Investigators: Dhabaleswar Panda, Professor, Ohio State University
P. Sadayappan, Professor, Ohio State University

Dates: December 1, 2000 – February 28, 2002

CSRI POC: Neil Pundit

Project Summary:

Message passing software is a critical component of massively parallel machines, and much effort has been expended to achieve fast, reliable communications. Little work has been put into understanding the available design space and identifying particularly beneficial aspects of existing message passing systems. Meanwhile, networking hardware continues to evolve in ways that force changes in message passing software. It is crucial to develop a basis for comparison of network software, learn from the successes and failures of previous designs, and understand how protocols must change to take advantage of improvements in hardware.

Overview: As cluster computing continues its rise in popularity, and availability of inexpensive high-performance hardware components improves, the focus for investment in large-scale parallel computing is rapidly shifting from hardware to software.

One of the main barriers preventing a cluster of individual computers from functioning as a true supercomputer is the lack of a robust network. Although the physical wires connecting the machines may competently move bytes from one end of the machine room to the other, it is the protocols and software that are responsible for ensuring the rapid performance and correct results of parallel programs.

Much time has been spent developing custom protocols for a wide range of network passing hardware, and one must imagine that a large fraction of that work would be similar given the identical goal of reliable high-speed communication. However, there is no source for an unbiased comparison of the attempted systems, and no common language by which to classify their aspects. Analyses of this sort are only seen in technical papers announcing new communication protocols and relegating historical ones to sections titled "Other work."

Due to continual improvements in semiconductor technology, network hardware is improving in aspects of switching speed, power consumption, and density. Software which will be used on newer generations of hardware must adapt to accommodate these changes to avoid succumbing to newly-emerging bottlenecks in communication.

In particular, Myrinet from Myricom is a frequently used message passing infrastructure, and it is shipped with essentially no software by which to use it. Many organizations have undergone the trials of writing their own message passing software to exploit the intrinsically high bandwidth, low latency network offered by Myrinet. And each variation of this piece of software has landed somewhere different on the design space for the problem.

We will study available designs and implementations for high-performance Myrinet software with respect to three components which are critical to large-scale technical cluster computing: reliability, performance, and scalability.

Reliability is the mechanism by which messages are guaranteed to be delivered without corruption, and how that is implemented. Some assume the Myrinet hardware is reliable, or place the onus on the application program. Those that implement some sort of retransmission algorithm must decide to put the

functionality on the NIC or in the host, and if timers will be used or a scheme of positive and negative acknowledgements from the receiver.

Performance is a complex metric which starts with single-communication latency and bandwidth measurements, but becomes more complex when considering the interaction of multiple simultaneous transfers and collective communications.

Scalability addresses the production nature of message passing software. Without it, systems are limited to only small sizes. Limitations to scalability may come from the caching of connection state on the NIC, insufficiently wide bit fields, limited capability to use multiple processors on an SMP node and multiple NICs, use of collective algorithms with linear or worse order, and many other sources**.**

The interactions of these three are complex. To generalize, reliability often reduces performance due to the need for more bookkeeping and network traffic for each message, and it reduces scalability by the requirement of more saved state per message. Conversely, high performance is often achieved by pre-allocating receive buffers and caching much of the message state at the expense of scalability, and avoiding buffer copies, checksums, and timer interrupts at the expense of reliability.

Detailed descriptions of how each system approaches the above aspects, and the trade-offs made in their design decisions will be provided. Other considerations important to implementers will be noted, although not studied as extensively. Systems initially planned for consideration are: GM, LFC, LCI, PM, AM-II, FM, VMMC2, BIP, BDM, Hamlyn, U-net, Trapeze.

To make concrete our understanding of the results in the previous section, we plan to make modifications to one particular Myrinet software implementation, GM, and to study the different trade-offs. For example, the retransmission timer in GM is implemented on the NIC, and packets to be resent are constructed entirely in the NIC with no aid from the host. First removing this functionality from the MCP, then reinserting it in the host, would allow us to determine exact costs for this operation, and the effect on scalability from freeing up NIC resources which track message state. Artificially injecting failures to cause more frequent triggering of the retransmit code path would allow us to determine the trade-off point for implementing retransmission in the slower host as a function of bit error rate. Timing MPI applications under changes in GM will point out which modifications are important for current hardware.

Looking further on in time, changes to network hardware are destined to occur as driven by the insatiable demand for ubiquitous computing by society at large. Standards for new network protocols and hardware for both networks and main system busses have been arising recently with much industry support, including ST, VIA, Quadrics, and InfiniBand (SIO and FutureIO). It is likely that forthcoming changes will necessitate design changes in message passing software to continue to reap maximum performance from the underlying network. We will consider the effects to message passing of general changes in network parameters: higher bandwidth, faster off-board processors, faster system busses, tighter processor/network integration, and so on. Available documentation for InfiniBand and Quadrics will be referenced in particular. Changes in hardware may be accompanied with changes in the APIs used to access the hardware, as is the case with InfiniBand and VIA, and not all changes may work to the benefit of those doing message passing communications. Attention to interactions along the full ``stack'' from application down to hardware is necessary.

Deliverables:

1. Understand and analyze the reliability mechanism in GM, by listing all the different fault scenarios which GM handles, and using that to build state-transition diagrams to show how GM handles the faults.
2. Derive a cost model to the state-transition diagram in terms of how much overhead (instruction count, interactions with host) GM requires to handle recovery for each fault.
3. Analyze the impact of different types of faults and their frequencies on the overall performance. We can develop a small simulation framework to evaluate the performance. This simulation framework

needs to be parameterized so that we can evaluate the impact of host speed, NIC speed, and network link speed, in addition to the failure characteristics.

4. Look 1-3 and investigate whether there are better solutions. If so, then we first incorporate them into the simulation model and see which of the solutions are beneficial.
5. Based on the above step, we modify GM and study its performance on actual hardware, including machines in the CIS nowlab and at OSC. The performance evaluation of our changes on GM can be compared with the simulation model, validating the study.
6. Deliver the simulation components and modified GM.

Appendix B: Sabbaticals

Computer Science Research Institute
Sandia National Laboratories

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. Raphael Haftka
University of Florida

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

Dr. Robert Preis
University of Paderborn, Germany

Dr. Virginia Torczon
College of William and Mary

Descriptions of projects by Drs. Berggren, Haftka, Pothen, and Preis follow in this section. The project of Dr. Torczon is joint with Dr. Michael Lewis and is in the section of CSRI funded projects.

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: October 1, 2001 – December 31, 2002

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 boundries, 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 sutdies 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: Global Optimization under Uncertainty

PI: Raphael Haftka
University of Florida

Dates: January, 2001 – May, 2001

Project Summary:

The design of an engineering system involves a set of design variables that defines the freedom of the designer to change the system and a set of `fixed' parameters that define aspects of the system that cannot be changed. In structural design, for example, design variables may control the geometry and material properties of structural members, while the fixed parameters may include other geometry and material properties that cannot be changed as well as the loads.

In order to evaluate the performance of a given design, we normally perform computer simulations of its response to its environment. Design optimization of a system typically requires dozens or hundreds of simulations even when we only seek a local optimum close to our initial design. To search for a global optimum may require many thousands of simulations.

The evaluation of the reliability of a single design often requires sampling the same design under slightly perturbed values of the design variables and fixed parameters. In structural design, loads, material properties, and geometry are varied based on statistical data about manufacturing tolerances and loading events (such as earthquakes). Thus, reliability calculation for a single system may be as expensive as deterministic optimization of that system. Optimization for improved reliability, and in particular, global optimization for reliability may require millions of system response evaluations. Consequently, global optimization for improved reliability has not been attempted except for the simplest of design problems.

Our research group at the University of Florida and Virginia Tech has been active in seeking relief for the computational burden of both global optimization and local optimization subject to uncertainty. We have explored the use of response surface techniques and of coarse grained parallel computation this purpose Tackling the problem of global optimization under uncertainty is a logical continuation of this earlier and ongoing work.

In discussions with Sandia colleagues, I found that there is much interest in both topics in the Optimization/Uncertainty Estimation Department (9211). The work of Bill Hart on genetic algorithms and global optimization, John Red-Horse (uncertainty estimation), and the DAKOTA group (Mike Eldred and Tony Giunta, software for optimization, uncertainty estimation and parallel computing) is particularly relevant to the topic I am proposing. In addition, I was inspired by the work of Vicente Romero of the Verification and Validation Department on using sampling techniques from reliability methods for global optimization.

In the past five years, there has been growing interest in global optimization methods that work on a population of designs and use it to estimate the statistical attributes of superior designs. These estimation of distribution algorithms can also be viewed as a logical improvement on genetic algorithms. Genetic algorithms also work with population of designs, but they do not try to derive statistical information from the population that could accelerate the search for the optimum.

Genetic algorithms are similar in their operation to directed breeding programs. Consider, for example, a program to improve the hardiness of cattle for semi-arid regions. If you simply pick superior animals for further breeding you will running a traditional genetic algorithm. If on the other hand, you map out the genome of the individual animals, and use statistics to identify genes and gene complexes associated with hardiness, you will be performing EDA. You may then be able to make more specific decisions on pairs of animals to breed based on their genetic makeup rather than just their hardiness.

Because EDA methods can benefit from expanded sampling, they seem to be natural targets for synergy between global optimization and sampling for uncertainty quantification. In the extreme case where all the uncertainty comes from design variables the match between the two types of sampling is perfect, and the cost of global optimization subject to reliability may be reduced by an order of magnitude. In addition, EDA methods may also benefit reliability calculations of a single design, because this often involves search for the most probable failure condition, which is a global optimization problem.

During my sabbatical at Sandia, I would like to devote it to explore this potential synergy between EDA algorithms and sampling for uncertainty quantification. Aside from the algorithmic challenge of combining the two, I plan to explore the use of parallelization for these algorithms. They can benefit from a large number of simulations performed in parallel, and can also benefit from parallel calculation of structural response and the derivatives of the response with respect to problem parameters.

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.

Appendix C: Seminar Abstracts

Computer Science Research Institute
Sandia National Laboratories

Title: Aerodynamic Optimization with Variable-Fidelity Models

Speaker: Dr. Natalia Alexandrov
Multidisciplinary Optimization Branch
NASA Langley Research Center

Date/Time: Tuesday, January 16, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: This work discusses the application to aerodynamic optimization of a first-order approximation management framework based on sequential quadratic programming. Approximation management is a methodology aimed at maximizing the use of low-fidelity models in iterative procedures with occasional but systematic recourse to higher-fidelity models for monitoring the progress of the algorithm. We discuss the scheme's convergence properties and present a computational demonstration on aerodynamic optimization of a 3D wing and a 2D two-element airfoil. For the 3D wing optimization problem, the variable-fidelity models are represented by Euler analysis performed on meshes of varying degrees of refinement. For the two-element airfoil, computation with variable-fidelity physics models (Navier-Stokes vs. Euler) is attempted.

CSRI POC: Anthony A. Giunta, 09211, 844-4280

Title: Aerodynamic Design Using the Euler and Navier-Stokes Equations

Speaker: Professor Juan Alonso
Stanford University

Date/Time: Tuesday, May 15, 2001, 10:00 am (PST)

Location: Bldg. 940, Auditorium (Sandia-CA)
Bldg. 980, Rm. 24 (Sandia-NM)

Abstract: The design of high-performance aircraft and spacecraft requires the use of high-fidelity models in a variety of the participating disciplines that usually include aerodynamics, structures, propulsion, mission performance and stability and control. These models are usually coupled to gradient-based optimization techniques in order to obtain performance improvements over the baseline design. Unfortunately, for high-dimensional design spaces (typically, N ~ 1,000 in our applications), the calculation of accurate sensitivities is an extremely demanding computational task that dominates the overall process of computational design. This talk will present an overview of a variety of advanced techniques that we are currently using for the computation of sensitivities in both single- and multiple-discipline design studies. Among them, the adjoint method for the Euler and Navier-Stokes equations, coupled aero-structural adjoints for design in the presence of elastic deformations, and the complex step will be discussed in comparison to the more traditional finite difference method. Several examples of the use of these techniques in aerospace design will be presented, drawing on our recent experience in transonic commercial transport design, low boom supersonic aircraft, and supersonic transport aero-structural design.

CSRI POC: Monica Martinez-Canales, 8950, (925) 294-3157

Title: Using Motion Planning to Study Protein Folding

Speaker: Nancy Amato
Texas A&M University

Date/Time: Tuesday, November 6, 2001, 10:00 am (PST)

Location: Bldg. 940, Auditorium (IMTL) (Sandia - CA)
Bldg. 980, Room 24 (Sandia - NM)

Abstract: Motion planning arises not only in robotics but in many other areas such as intelligent CAD (virtual prototyping), mixed reality systems training and computer-assisted operation), and even computational biology and chemistry protein folding and drug design). Surprisingly, a single class of planners, called probabilistic roadmap methods (PRMs), have proven effective on problems from all these domains. Strengths of PRMs, in addition to versatility, are simplicity and efficiency even in high-dimensional configuration spaces. Moreover, PRMs are (almost) embarrassingly parallel. In the first part of this talk, we introduce the PRM framework, briefly describe several PRM variants developed in our group, and discuss parallelizing PRMs. In the second part, we concentrate on our recent application of PRM-based motion planning techniques to protein folding. Our focus in this work is to study the protein folding mechanism assuming we know the native fold. Therefore, instead of performing fold prediction, we aim to study issues related to the folding process, such as the formation of secondary and tertiary structure, and the dependence on the initial conformation. Our results on several small to moderate sized proteins (60-150 amino acids) indicate that the PRM-based technique generates folding pathways that are in agreement with experimental data. Our technique naturally supports the study of folding pathways starting from any desired denatured starting conformation, and also appears to differentiate between proteins where secondary structure forms first and those where the tertiary structure is obtained more directly.

If time allows, we will describe initial promising results using PRMs for ligand/protein binding; this work utilizes haptic user input. More information regarding our work, including movies, can be found at http://www.cs.tamu.edu/faculty/amato.

CSRI POC: Juan Meza, 8950, (925) 294-2234

Title: A Comparison of Factorization-free Eigensolvers with Application to Cavity Resonators

Speaker: Peter Arbenz
Institute of Scientific Computing
Swiss Federal Institute of Technology (ETH)

Date/Time: Tuesday, September 18, 2001, 10:30 am

Location: Building 980 Room 95

Abstract: We investigate solvers for the symmetric eigenvalue problem A x = M x lambda with positive definite M that do not require the factorization of either A or M. Preconditioned Rayleigh quotient minimization and Jacobi-Davidson methods are compared. Our model problems come from large scale finite element discretizations of the new radio frequency cavity ring cyclotron installed at the Paul Scherrer Institute in Villigen, Switzerland.

CSRI POC: Rich Lehoucq, 9214, 845-8929

Title: Eulerian-Lagrangian Coupling via a Level Set Technique.

Speaker: Marco Arienti
CalTech

Date/Time: Thursday, August 30, 2001, 10:30 am

Location: Building 980 Room 95

Abstract: In this talk, a numerical method is presented for coupling an Eulerian compressible flow solver with a Lagrangian solver. Applications of interest are fast transient problems involving fluid-solid interactions. The algorithm allows stand-alone Eulerian and Lagrangian codes to perform fully coupled simulations with minimal additions and virtually no modifications.

Standard integration of the Euler equations over a Cartesian mesh and interface reconstruction via a level set technique are employed. To treat the irregular boundary cells, the Eulerian computational domain is augmented by a thin layer of Cartesian ghost cells. The Lagrangian mesh evolves concurrently, driven by the traction boundary conditions imposed by the Eulerian flow.

During the development of the algorithm, several numerical tests were designed to measure the rate of convergence and the accuracy of the coupling. Some of these problems for one and two dimensions will be presented. They include a test case consisting of an isotropic elastic solid and a compressible fluid in a fully coupled setting where the exact solution is available.

CSRI POC: Daniel E. Carroll, 09231, 845-8069

Title: Advances on a Scaled Least-Squares Neutron Transport Method

Speaker: Travis Austin
University of Colorado

Date/Time: Monday, April 16, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: The neutron transport equation models the flux of neutrons within some background material. The neutrons either stream through the material uninterrupted or collide with nuclei followed by absorption or scattering. In this talk, we will consider a minimization problem for a scaled least-squares neutron transport functional; the solution to this minimization problem is also a solution to the transport equation. Several aspects of the scaled least-squares neutron transport method will be studied. First, we look at the scaled interior functional and the boundary functional. For the interior functional, we explain the need for the scaling operator. Without this scaling, minimizing the functional in the optically thick regime using standard trilinear elements gives a numerical solution which poorly approximates the true solution. Second, we talk about a numerical solution method for minimizing the least-squares functional; the method we consider is multigrid. In the optically thick regime, multigrid converges very slowly for a P_N discretization of the angular flux, a standard trilinear finite element approximation of the moments, and pointwise Gauss-Seidel smoothing in the multigrid algorithm. Using a new finite element space for the first-order moments and a block Gauss-Seidel smoothing, we can guarantee fast convergence. Third, we look at a new scaling for the interior functional when the scaling operator is anisotropic. The scaling guarantees ellipticity in a scaled H^1-like norm.

CSRI POC: John Shadid, 09233, 845-7876

Title: High-Performance Algorithm Engineering for Gene-Order Phylogenies

Speaker: David A. Bader, University of New Mexico
(Joint work with Bernard M.E. Moret, Tandy Warnow, Stacia K. Wyman, and Mi Yan)

Date/Time: Thursday, March 8, 2001, 3:00 pm

Location: Building 980 Rm. 95

Abstract: Phylogenies derived from gene order data may prove crucial in answering some fundamental questions in biomolecular evolution. Yet very few techniques are available for phylogenetic reconstruction based upon gene order and content, and these are (for the most part) computationally expensive. High-performance algorithm engineering offers a battery of tools that can reduce, sometimes spectacularly, the running time of existing approaches. We discuss one such application, in which we started with the method known as ``breakpoint analysis'' (developed by Sankoff and his colleagues) and produced a software suite, GRAPPA, that demonstrated a million-fold speedup in running time (on a variety of real and simulated datasets), by combining low-level algorithmic improvements, cache-aware programming, careful performance tuning, and massive parallelism. We show how these techniques are directly applicable to a large variety of problems in computational biology. (Supported in part by DOE CSRI 14968, NSF Grants CAREER 00-93039, ITR 00-81404 and DEB 99-10123.)

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: A-Posteriori Error Estimates for Godunov Finite Volume Methods and Generalized Discontinuous Galerkin Methods on Unstructured Meshes

Speaker: Tim Barth, Physics Simulation and Modeling Office
NASA Ames Research Center

Date/Time: Wednesday, December 5, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: A-Posteriori Error estimates for Godunov finite volume methods are presented which exploit the two solution representations inherent in the method, viz. as piecewise constants u_0 and cellwise q-order reconstructed functions $R^0_q\ u_0$. Using standard operator duality arguments, analytical and numerical results for advection, diffusion, and advection-diffusion equations are presented. We then considered the discontinuous Galerkin FEM in abstract form, B(u,v)=F(v), and a generalized Petrov-Galerkin variant on finite dimensional subspaces written in abstract form as B(R^p_q u_p, v_p)=F(v_p). Once again, two solution representations are inherent in the Petrov-Galerkin formulation and a-posteriori error estimates are derived and numerically tested.

CSRI POC: Pavel Bochev, 844-1990

Title: Fault Tolerance in LAM/MPI

Speaker: Brian Barrett
Computer Science Department, Indiana University

Date/Time: Monday, July 2, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Large clusters comprised of commodity hardware are an attractive solution for parallel computing using MPI. However, there are currently no MPI implementations that allow user applications to reliably survive the death of one or more nodes during execution. The LAM implentation of MPI, an open-source implementation maintained at Indiana University, while supporting some fault tolerant capabilities, currently falls well short of providing true reliable execution in an unreliable environment.

This talk will discuss the current fault tolerant capabilities of LAM/MPI, as well as ongoing research into providing portable fault tolerance in MPI. In particular, I will discuss the algorithm presently used by LAM's run-time environment to recover from faults, planned improvements to this algorithm, how this will apply to the MPI programmer, and a proposed model for an MPI program to survive failures during execution.

CSRI POC: Richard R. Drake, 9231, 844-2537

Title: Optimization theory and software

Speaker: Roscoe Bartlett
Carnegie Mellon University

Date/Time: Monday, January 15, 2001, 3:30 pm

Location: Building 980 Rm. 95

Abstract: Optimization theory and software provide important tools for many different areas of Process Systems Engineering (PSE). Optimization problems arise at nearly every level of PSE from high level business planning, to scheduling and design, to parameter estimation, to real-time optimization, all the way down to process control. A particularly successful class of optimization algorithms for solving Nonlinear Programs (NLPs) (a standard formulation optimization problems often take) is Successive Quadratic Programming (SQP).

This presentation describes efforts to apply Object-Oriented (OO) technology to SQP to develop an experimental framework for SQP called rSQP++. Three types of challenges and variabilities to deal with when developing such a framework are discussed: algorithmic variability, implementation variability, and NLP specific specializations. It will be shown how new OO methods can address each of the above challenges.

CSRI POC: Bart van Bloemen Waanders, 09211, 284-6746

Title: Support Theory: A Framework for Analyzing Preconditioners

Speaker: Erik Boman,
SNL

Date/Time: Monday, June 11, 2001, 12:00 noon

Location: Building 980 Rm. 95

Abstract: Preconditioning is important to accelerate the convergence of iterative solvers, but the design and analysis of preconditioners is challenging and remains as much an art as a science. We have developed a new set of techniques for analyzing preconditioners for symmetric positive semidefinite linear systems. Such systems arise frequently in applications, for example in the finite element solution of self-adjoint PDEs. Our techniques provide a mathematical toolkit for bounding the so-called support number, thereby obtaining a bound on the number of iterations required for iterative methods like conjugate gradients.

We show that for a special class of matrices, our support theory reduces to the support-graph theory developed by previous authors. We then demonstrate the utility of our methodology by constructing new preconditioners for some finite element problems, and devising a new analysis of incomplete factorization preconditioners.

Joint work with Bruce Hendrickson.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Sensitivity Equations for Thermal Fluid Systems

Speaker: Jeff Borggaard
Virginia Tech

Date/Time: Tuesday, March 6, 2001, 10:00 am PST

Location: Bldg. 940, Auditorium (SNL-CA)
Bldg. 980, Room 24 (SNL-NM)

Abstract: We discuss several advantages to using continuous sensitivity equations (CSE). These equations are obtained directly from a PDE rather than the common strategy of differentiating the discrete approximation of the PDE. This leads to flexibility in developing special numerical techniques for the coupled PDE/CSE system. These techniques include adaptive mesh refinement strategies that incorporate sensitivity variables, solving the CSE as a post-processing step, developing algorithms that exploit the linearity of the CSE, and avoiding so-called mesh sensitivities which would arise when differentiating the discrete approximation. The latter can often be achieved in automatic differentiation software.

We also present a number of applications of sensitivity equations including uncertainty analysis, finding nearby solutions, ranking the significance of design variables, and computing gradients for optimization algorithms. To illustrate these applications, we study thermal fluid systems-coupled Navier-Stokes and energy equations - where we allow for temperature dependent fluid properties (viscosity, thermal conductivity, etc.).

CSRI POC: Monica Martinez-Canales, 8950, (925) 294-3157
Kevin Long, 8950 (925) 294-4910

Title: Market-Based Mobile-Agent Planning

Speaker: Jonathan L. Bredin
Dartmouth College

Date/Time: Tuesday, January 30, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: We promote markets as mechanisms to share computational resources. Markets by themselves do not solve all problems, however. While economists assume their agents are rational, engineers must synthesize rationality. As such, resource-usage and expenditure planning become key issues. In this talk we explore market-based mechanisms for resource control and algorithms to optimize an agent's performance under the mechanisms.

We present algorithms for resource control and planning and compare their performance with established methods for resource control. The use of our methods increases system throughput compared with traditional allocation methods. Additionally, our allocation supports task prioritization. We also show methods for reducing volatility in execution time though reservation systems. We conclude by sketching future work that applies our ideas to resource-usage forecasts, robust network routing, and anonymous trust networks.

CSRI POC: Grant S. Heffelfinger, 09235, 845-7801

Title: Issues involved in use of genomic and other biological information to model cellular function

Speaker: Roger Brent,
Director, The Molecular Sciences Institute
Berkeley, CA

Date/Time: Tuesday, February 20, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: At TMSI, we are working very hard to build quantitative models of cellular function. Scientifically this task is important in that success in it would enable us to make better predictions of the future behavior of biological systems given knowledge of their current state. Success would also enable these models to be used as design tools for future design and engineering of biological systems. In doing this work, we face and will continue to face significant experimental and computational challenges. I will describe some of these and our efforts to surmount them.

I will also review our thinking and progress on a closely-related issue, which is the design of computational representations that can help codify existing non-genomic biological knowledge, which is mainly qualitative. These structures will probably go beyond either relational databases and frame based or other object-like representations, to involve other means better suited to representing processes, events causality, and the passage of time. Construction of these representations would support quantitative modeling efforts, but also have significant impact on science and medicine. It is likely that eventual construction of such representations will be a large effort involving both public and private components.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: A Physicist's Perspective on Scientific Computing

Speaker: Philip Campbell, SNL

Date/Time: Wednesday, June 20, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: The evolution of numerical algorithms, languages, architectures, and computing philosophy are traced through the experiences of a computational physicist working with hydro and shock-physics codes, EOS and opacity codes, electromagnetic PIC codes, SPH and transport codes in various areas of research.

Much of the work on hydro codes was done on shared-memory vector machines. Later, electromagnetic PIC codes were adapted to the early NCUBE and Intel hypercubes and the CM-2. Up to this point most scientific work was done in Fortran. Pioneering work in C++ at Sandia is mentioned as well as some of the early problems in T3D development at Cray Research. Current work is mostly done on networked SMPs with Linux systems becoming more common.

Computer manufacturers are now encountering many benchmarks in C and C++ as well as Fortran 90. A few of the more interesting and challenging of these benchmarks at Cray and SGI will be described**.**

CSRI POC: Sue Goudy, 09223, 844-6083

Title: Getting a Clue: Knowledge acquisition through guessing in an uncertain game environment

Speaker: Steve Carrion
University of Michigan, Ann Arbor

Date/Time: Wednesday, December 12, 2001, 9:00 am

Location: Building 980, Room 95

Abstract: It is often the case that actions in a game environment need to be taken on less than complete knowledge of the world. Traditionally in these situations, the information at hand is used to formulate the next action in an attempt to draw closer to a goal. The goal here is to develop an implementation of the reversal of this method: the extraction of additional information using the actions available and prior knowledge. In other words the goal itself is information gain.

This is an implementation paper concerning the game environment 'Clue', focusing on the development of representations, methodologies, and strategies to aid in the creation of a system to effectively and efficiently gain information in this setting.

CSRI POC: Erik DeBenedictis, 9223, 284-4017

Title: Numerical Solutions for Minimal Surfaces

Speaker: Ulises Cervantes-Pimentel
University of Illinois at Urbana-Champaign

Date/Time: Monday, January 29, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: The simplest boundary value problem for minimal surfaces is that of finding a surface of least area spanned by a single closed wire, known as Plateau's problem. Besides the traditional finite difference approach, there are numerical techniques using finite elements on surfaces to compute stable minimal surfaces as well as max flow/min cut type algorithms to obtain area-minimizing surfaces. In all this cases, in order to obtain reasonable accuracy the numerical estimates require decompositions too fine for current workstations. Some of these algorithms will be discussed as well as the newest finite element procedures for approximating minimal, including unstable, surfaces

CSRI POC: Robert Leland, 09226, 845-8439

Title: The Interspace Prototype: An Environment for Semantic Indexing, Retrieval, Interoperability

Speaker: Conrad Chang
University of Illinois

Date/Time: Friday, August 24, 2001, 3:00 pm

Location: Building 980 Room 100

Abstract: The goal of the Interspace project is to build an information analysis environment for distributed object repositories. The infrastructure will support semantic retrieval and indexing on textual domains, although the general techniques should apply to other media types. The prototype also supports integrated interoperability of semantic services based on statistical algorithms for information management. The integration provided by the environment will enable users to issue queries and obtain more relevant results through the process of query refinement and the aid of automatic document clustering. We have built a distributed implementation using NOW for performance improvement.

CSRI POC: Ron Brightwell, 9223, 844-2099

Title: First-Order System Least Squares (FOSLS) for Elliptic Grid Generation (EGG)

Speaker: Andrea Codd
Department of Applied Mathematics
University of Colorado at Boulder

Date/Time: Tuesday, April 17, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: Elliptic Grid Generation (EGG), using the Winslow generator, defines a map between a simple computational region and a potentially complicated physical region. It can be used numerically to create meshes for discretizing equations directly on the physical domain or indirectly on the computational domain by way of the transforming map. EGG allows complete specification of the boundary, and it guarantees a one-to-one and onto transformation when the computational region is convex.

A new fully variational approach is developed for solving the Winslow equations that enables accurate discretization and fast solution methods. The EGG equations are converted to a first-order system that is then linearized via Newton's method. First-order system least squares (FOSLS) is used to formulate and discretize the Newton step, and the resulting matrix equation is solved using algebraic multigrid (AMG). The approach is coupled with nested iteration to provide an accurate initial guess for finer levels using coarse-level computation. Theoretical and numerical results confirm the usual full multigrid efficiency: accuracy comparable to the finest-level discretization is achieved at a cost proportional to the number of finest-level degrees of freedom.

CSRI POC: Knupp, Patrick, 09226, 284-4565

Title: Columnsort Lives! An Efficient Out-of-Core Sorting Program

Speaker: Dr. Thomas Cormen
Dartmouth College

Date/Time: Thursday, December 13, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Suppose that you have a big parallel computer and you wish to sort a lot of data. So much data that it does not fit in the memory of your big computer (we call such problems "out-of-core"), nor does it fit on a single disk. How could you sort it efficiently?

In this talk, we present the design and implementation of a sorting algorithm that works under the above conditions. Our algorithm is based on Leighton's columnsort algorithm. We show how to relax some of the steps of the original columnsort algorithm to permit a faster out-of-core implementation. Our algorithm requires only four passes over the data. Although there is a limit on the number of items that can be sorted--as a function of the memory used per processor—this upper limit need not be a severe restriction, and it increases superlinearly with the per-processor memory. We define several measures of sorting efficiency and demonstrate that our implementation's sorting efficiency is competitive with that of NOW-Sort, a sorting algorithm developed to sort large amounts of data quickly on a cluster of workstations.

We also discuss three improvements to the algorithm. One uses threads for maximum overlap of I/O, computation, and communication, and this improvement pays off nicely. One reduces the number of passes from four down to three, but this improvement does not always pay off. And one adds a pass but in so doing increases the maximum problem size as a function of per-processor memory.

Joint work with Geeta Chaudhry and Len Wisniewski

CSRI POC: Cindy Phillips, 9211, 845-7296

Title: Improvements to the Debugger

Speaker: Jim Cownie and Rich Collier
Etnus (TotalView debugger)

Date/Time: Friday, November 9, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Jim and Rich will present a talk on the ASCI improvements that have been made to the debugger most recently. If there is interest, Jim and Rich will also address the new look of TotalView 5.

The ASCI improvements include:

- Execution Performance Enhancements for Massive Parallelism
- Support for Compaq Alpha Fortran OpenMP/NUMA directive
- Support for Memory Utilization Statistics
- Support for Attach to Subset of Processes
- Support for Non-Interactive CLI Scripts
- Support for Process-Thread Set Grouping in Motif UI
- Support for Calling Functions in the CLI
- Support for Parallel Evaluation of Expressions
- Support for Enhanced Message Queue Debug Library Capability
- Support for Non-Uniform Address Space Distributed Debugging
- Support for Variable Transformations
- Support for Arrays of Structure Components
- Support Displaying C and C++ Pointer Variables as Arrays
- Support Setting/Modifying Data Display Format and Precision

CSRI POC: Ron Brightwell, 9223, 844-2099

Title: Iterative methods and approximating large systems by small systems: some theory and many open questions

Speaker: Jane Cullum
Computer and Computational Sciences Division
Los Alamos National Laboratory

Date/Time: Tuesday, December 4, 2001, 10 am (PST)

Location: Building 980 Room 24 (Sandia – NM)
Bldg. 921, Room 137 ((Sandia – CA)

Abstract: We focus on the development and use of matrix iterative methods for obtaining small models for use in simulations of very large linear systems. The proposed iterative algorithms are not application specific.

We illustrate the use of such ideas within the context of modeling and simulation of interconnects (wires, planes) in very large integrated circuits (VLSI). Fast clock times and the high density of the circuits result in electromagnetic interference between the interconnects which can distort and delay the signals in the interconnects. The Partial Element Equivalent Circuit (PEEC) models of the interconnects which we use yield very large systems of delay-differential-algebraic equations. The corresponding matrix formulations are nonsymmetric and not sparse.

We look at the question of approximating such systems by small systems whose behavior mimics the behavior of the original system. The work involves boundary integral equations, electrical circuits, iterative methods for large matrix problems, systems theory, nonlinear eigenvalue computations and analysis, and leads to many open questions. We demonstrate a model reduction technique which appears to work well and which is naturally parallel. Theoretical underpinnings for this method exist when the time delays are zero. However, there are many unanswered theoretical questions concerning the application of this method to physically realistic cases with time delays. This talk will focus on the iterative method ideas and will not assume any a priori knowledge of circuits.

CSRI POC: Victoria Howle, 08950, 925-294-2204

Title: Algorithm Development for Gas Dynamics

Speaker: W. Wenlong Dai, Senior Research Scientist
Laboratory for Computational Science and Engineering
University of Minnesota

Date/Time: Wednesday, April 11, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: In this talk, I will review some of the algorithm development that I have been involved with in Paul Woodward's research group at the University of Minnesota during the last few years, including explicit and implicit schemes. I will focus on magnetohydrodynamical (MHD) equations for explicit schemes, which include characteristic formulations, Riemann solvers and the divergence-free condition. For implicit schemes I will choose gas dynamics, MHD and (simplified) radiation hydrodynamics as examples, and will focus on temporal accuracy, our own iterative (linear and nonlinear) solvers and their convergence rate. If possible, I will also discuss highly accurate schemes for Euler equations.

CSRI POCs: John Shadid, 09233, 845-7876 and Allen Robinson, 09231, 844-6614

Title: Dynamic Data Driven Application Systems

Speaker: Dr. Frederica Darema
Senior Science and Technology Advisor, NSF/CISE

Date/Time: Thursday, February 15, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: At a February 17, 2000 congressional briefing, meteorologists were asked why they missed predicting the track and magnitude of a major storm in January 24-25, 2000, that blanketed major cities from South Carolina to New England. One of the reasons cited by the scientists is that computer models (simulations) were not geared to incorporate changing conditions (like prevalent winds) as the many hours long computer simulations proceeded. On May 7, 2000, the national park service started a controlled burn near Los Alamos National Laboratory. Within a day, the fire was labeled a wildfire. Once again, the existing methodologies were unable to simulate what the behavior of the fire based upon real-time changing conditions, and the emergency response agencies were thus unable to take appropriate and effective actions to limit the propagation of the fire.

A new paradigm for application simulations will be discussed that is geared to enhance the analysis and prediction capabilities of modeling and simulations, thus eliminating shortcomings as those in the examples above. In this new paradigm, the applications will be able at execution time to accept and respond to dynamically injected measurements, and reversely, such applications will have the ability to dynamically control the measurement process. The synergistic and symbiotic feedback control-loop between simulations and measurements can open new domains in the capabilities of simulations, with high potential pay-off, and create applications and measurement methodologies with new and enhanced capabilities. It has the potential to transform the way science and engineering are done, and induce a major impact in the way many functions in our society are conducted, such as manufacturing, commerce, transportation, hazard prediction/management, and medicine.

The talk will address the challenges and the applications' and systems' software technologies needed to enable such capabilities. These include: application composition environments, performance engineering methodologies and new compiler technologies and systems software enabling dynamic runtime support, applications' algorithms amenable to perturbations by the dynamic data inputs. Such challenges clearly also present the need for a synergistic multidisciplinary research in applications, systems' and algorithms' areas. Current relevant research projects and programs will be discussed, together with additional efforts needed.

CSRI POC: Bart van Bloemen Waanders, 09211, 284-674

Title: Simulation of Flow Through Porous Media

Speaker: Clint Dawson, Professor, Aerospace Engineering
University of Texas at Austin

Date/Time: Tuesday, February 6, 2001, 10:00 am

Location: Building 940 Auditorium (SNL-CA)
Building 980 Rm. 24 (SNL-NM)

Abstract: Porous media problems are typified by flow equations, which relate velocities and pressure, and transport equations, which account for the conservation of certain chemical species. These mathematical models include coefficients which can vary by several orders of magnitude, point sources and sinks, and chemical reactions with widely disparate time scales. Steep gradients can occur in both fluid pressures and transported quantities. Simulating these features accurately and efficiently, while honoring the underlying conservation principles, are desirable goals when applying numerical solution techniques to these equations. In this talk, we will review some of the common numerical methods in use today for these problems, and discuss current research and future directions.

CSRI POC: Monica Martinez-Canales, 08950, (925) 294-3157

Title: The NetAlive Parallel Language and the Wireless Internet

Speaker: Eric DeBenedictis
NetAlive, Inc

Date/Time: Thursday, July 19, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: NetAlive is a computer language focused on making multiple computers work together on general-purpose tasks on a distributed computer network.

Background

The NetAlive language was developed by NetAlive, Inc., is patented, and became the basis of commercial products, such as: (1) knowledge management applications, where various servers on the Internet cooperate to "mine" data, and (2) wireless applications, where handheld devices cooperate with servers to carry out data access, bar code scanning, and other functions. The wireless applications are accessible on the Internet at http://www.netalive.com and have been used commercially by truckers, field service workers, and lawyers.

This talk will focus on how some of NetAlive's component technologies may be applicable to parallel supercomputing of the type performed at Sandia.

Key Features

- A NetAlive program encompasses all the processors working together, rather than having separate programs for different processors. This simplifies code development.
- The programmer need not create message-passing statements, but instead specifies only where code modules run. The compiler generates the message-passing protocol. This eliminates a source of error and simplifies programming.
- The modules of a NetAlive program include both computational functions and "widgets" of a Graphical User Interface (GUI) – and NetAlive includes a GUI page designer. This makes invoking NetAlive programs as easy as accessing a Web page.
- While NetAlive was not designed for supercomputing, it has a data parallel mode.

CSRI POC: Bob Benner, 9224, 845-7291

Title: Multiscale Nonrigid Data Registration Using Adaptive Basis Functions

Speaker: Thomas Disenbacher
Vanderbilt University

Date/Time: Thursday, February 22, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Data registration is the attempt to align, or register, two sets of data so that their local characteristics may be compared. For instance, in medical imaging, a person may have a magnetic resonance imaging (MRI) scan taken at one time and a positron emission tomography (PET) scan taken at another, and the researcher may wish to compare areas within both scans. Or, two different individuals may be scanned with fMRI, and one wishes to compare areas within the two separate brains.

Data sets are usually collected on a finite domain, which is typically idealized as the unit cube $[0,1]^3$. The image to be registered is often called the template, denoted by T(x), and the image it is to be registered to is called the subject, denoted by S(x). In the case of one individual and two imaging modalities, the registration problem can be thought of (roughly) as finding the correct amount to rotate and translate one data set to bring it into alignment with the second. In this case, we are solving the equation $S(x) = T(Ax+B)$, where A and B are rotation and translation matrices.

For the case of different individuals--and, in practice, for any two data sets we wish to align--we need to consider a more general transformation that incorporates nonlinearities. That is, the registration problem becomes one of finding a deformation field $\nu(x)$ (again on the domain $[0,1]^3$) such that $S(x) = A*T(x + \nu(x))$, where A in this case is a scaling matrix. More typically, instead of trying to achieve equality, we instead try to find the ν that optimizes some cost function $C(S,T,\nu)$ involving the template, the subject, and the deformation field itself. In this way, we can build in global or local preferences about what ways the template may deform in.

When performing registrations, it is crucial to maintain certain structures of T intact, as well as to keep the deformation field ν smooth. For instance, if we are comparing two different individuals' brains, we would like the hippocampus or hypothalamus (as a structure) to remain reasonably intact as one brain is registered to the other. And we expect structures not to overlap as they deform; that is, we in general expect $x+ \nu$ to be invertible. Current approaches to registration often impose conditions like these through heuristic means, but building them into models has proven to be more difficult due mainly to computational constraints.

We model the deformation field as $\nu:= \sum_{i=1}^n c_i \varphi_i$, where the c_i are multipliers for basis functions φ_i whose shape and location are chosen adaptively. This gives us flexibility in incorporating automatic basis function selection methods and a priori information about T and S into the model. That is, we start out with $\nu_0(x) = 0$, and adjust ν_0 locally wherever the fit between S and T seems worst, iteratively finding subsequent ν_i's that improve on the cost function $C(S,T, \nu)$. The basis functions φ that we use have local support, which speeds up computations.

By selecting and placing basis functions appropriately, the number of basis functions required to produce a good deformation is reduced over uniform grid methods. Thus, this method should provide both improved computational speed and improved accuracy over such methods.

CSRI POC: David Day, 09214, 844-1868

Title: An Efficient Approach To Multi-Objective Optimization Under Uncertainties

Speaker: Urmila Diwekar
Carnegie Mellon University

Date/Time: Thursday, June 21, 2001, 1:30 pm

Location: Building 980 Rm. 95

Abstract: Robust decision making under uncertainty is of fundamental importance in numerous disciplines and application areas. For many practical issues, decision making involves multiple, often conflicting, goals and poses a challenging and complex optimization problem. This talk presents innovations in the techniques underlying multi-objective optimization, in the characterization of uncertainties, and in the ability to develop and apply these methods outside of traditional application domains. The work presents innovative and powerful approaches for solving problems of optimization under uncertainty applied to large scale real world problems.

The problem of decision making under uncertainty is posed as a stochastic programming problem, which fundamentally involves constrained optimization of one or more probabilistic output functions constructed from multiple simulations for input parameter sets obtained by sampling uncertain input parameter distributions. Although valuable, the computational burden of this approach can be extreme and depends on the sample size used for characterizing the parametric uncertainties. The computational tedium presents a critical impediment to a widespread use of the stochastic analysis and optimization approach to robust decision making. This talk presents innovative computational strategies for significantly advancing the state-of-the-art in the area of optimal decision making under uncertainty by (a) improving algorithms for sampling over uncertain variables, (b) enhancing efficiency of (continuous variable) nonlinear programming algorithms by approximating derivatives through a re-weighting scheme, and (c) developing new algorithm for discrete optimization under uncertainty by using better error bounds for sampling. The enhancement in sampling is further used to derive a new and efficient multi-objective optimization algorithm, Minimizing Number of Single Objective Optimization Problems (MINSOOP), to generate a true representation of the whole Pareto set. The power and usefulness of approach is demonstrated using case studies in molecular modeling, environmental control, manufacturing, and management.

CSRI POC: Timothy G. Trucano, 09211, 844-8812

Title: Algorithms and Systems for High-Throughput Structural Biology

Speaker: Dr. Bruce Randall Donald
Department of Computer Sciences
Department of Chemistry, Dartmouth

Date/Time: Thursday, March 15, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: In the post-genomic era, key problems in molecular biology center on the determination and exploitation of three-dimensional protein structure and function. For example, modern drug design techniques use protein structure to understand how a drug can bind to an enzyme and inhibit its function. Large-scale structural and functional genomics will require high-throughput experimental techniques, coupled with sophisticated computer algorithms for data analysis and experiment planning. This talk will introduce techniques my lab is developing in two key areas: (1) data-directed computational protocols for high-throughput protein structure determination with nuclear magnetic resonance spectroscopy, and (2) experiment planning and data interpretation algorithms for reducing mass degeneracy in structural mass spectrometry for protein complex binding mode identification. These techniques promise to lead to fast, automated tools to pursue the structural and functional understanding of biopolymer interactions in systems of significant biochemical and pharmacological interest. Applications to the challenge of structural proteomics will be discussed.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: A Shortest-Path Approach to the Grid-Graph Partitioning Problem

Speaker: Dr. William W. Donaldson
University of Wisconsin

Date/Time: Wednesday, January 17, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Previous researchers have demonstrated that dividing a grid graph into exhaustive and mutually exclusive bands (stripes) will produce very good (and, in some cases, asymptotically optimal) partitions for regular grid-graphs. These earlier methods (striping heuristics) differed in the domains of application and the stripe-height selection process.

In this talk, I will present two results from my dissertation. The first proves that previous methodologies do not have reasonable run-times. The second result demonstrates a transformation from the original grid-graph-partitioning problem to a shortest-path problem. This transformation exploits certain properties within the set of feasible solutions and allows for the discovery in polynomial-time of an optimal set of stripe heights.

Computational results will be presented that demonstrate improved solution quality for general domains.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: An Additive Schwarz Preconditioner for the Spectral Element Ocean Model Formulation of the Shallow Water Equations

Speaker: Craig C. Douglas
Department of Computer Science, University of Kentucky
Department of Computer Science, Yale University

Date/Time: Thursday, December 6, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: We discretize the shallow water equations with an Adams-Bashford scheme combined with the Crank-Nicholson scheme for the time derivatives and spectral elements for the discretization in space. The resulting coupled system of equations will be reduced to a Schur complement system with a special structure of the Schur complement. This system can be solved with a preconditioned conjugate gradients, where the matrix-vector product is only implicitly given. We derive an overlapping block preconditioner based on additive Schwarz methods for preconditioning the reduced system.

This is joint work with Gundolf Haase (Johannes Kepler University) and Mohamed Iskandarani (University of Miami).

CSRI POC: Jonathan Hu, 294-2931 and Ray Tuminaro, 294-2564, 9214

Title: Fourier Analysis of Discrete Solutions of the Convection-Diffusion Equation

Speaker: Howard Elman
University of Maryland

Date/Time: Thursday, August 23, 2001, 9:30 am California PST

Location: Bldg. 921/137 with video link to Bldg. 980/95

Abstract: It is well known that discrete solutions to the convection-diffusion equation contain nonphysical oscillations when boundary layers are present but not resolved by the discretization. However, except for one-dimensional problems, there is little analysis of this phenomenon. In this work, we present an analysis of the two-dimensional problem with constant flow aligned with the grid, based on a Fourier decomposition of the discrete solution. For Galerkin bilinear finite element discretizations, we derive closed form expressions for the Fourier coefficients, showing them to be weighted sums of certain functions that are oscillatory when the mesh Peclet number is large. These expressions are then used to characterize the oscillations of the discrete solution in terms of the mesh Peclet number and boundary conditions of the problem. When streamline upwinding is included in the discretization, we show precisely how the amount of upwinding included in the discrete operator affects solution oscillations and accuracy when boundary layers are present.

Joint work with Alison Ramage of The University of Strathclyde.

CSRI POC: Ray Tuminaro, 9214, (925) 294-2564

Title: The Linux 2.4 Kernel

Speaker: Peter Espen, Organization 9143
RESPEC

Date/Time: Thursday, May 3, 2001, 1:30 pm

Location: Building 980 Room 95

Abstract: On Friday Jan 05, 2001 Linux Torvalds made a brief official announcement to linux-kernel that the Linux 2.4.0 kernel was released. This new kernel has many significant changes in architecture, filesystems, resource management, etc. from the previous 2.2.xx version Linux kernels. Many of these changes were designed in order to bring increased robustness and improved SMP scalability to the Linux kernel.

I will summarize the features of the Linux 2.4 kernel and highlight the significant changes from previous kernel versions that are present in the Linux 2.4 kernel.

CSRI POC: James Laros, 9224, 845-8523

Title: Preservation of Invariant Rectangles under Discretization

Speaker: Donald Estep
Department of Mathematics, Colorado State University

Date/Time: Thursday, September 27, 2001, 1:10 pm

Location: Building 980 Room 95

Abstract: An important issue in the study of reaction-diffusion equations is determining whether or not solutions blow-up. This carries over to numerical solutions, in which case we have the additional concern of determining if discretization either inhibits blowup or causes it to occur artificially. The existence of invariant rectangles inside of which solutions remain for all time is an important factor for addressing these issues in many cases. After presenting some motivating examples, we discuss the preservation of invariant rectangles under discretization in two ways. We construct special numerical methods that preserve invariant rectangles exactly and show how to use adaptive error control to preserve invariant rectangles in an approximate sense.

CSRI POC: Mark A. Christon, 9231, 844-8279

Title: Quantum Monte Carlo Algorithms: Making QMC Useful

Speaker: Michael Todd Feldmann
CalTech

Date/Time: Wednesday, August 22, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Quantum Monte Carlo (QMC) is a very powerful method for electronic structure prediction yet it has failed to be useful for those without large computational resources at their disposal. Our current focus is to address those computational issues that QMC will need to overcome in the near future to make it a useful method for the entire scientific community. Specific issues for QMC include massive parallelization for tens of thousands of cpus, on-the-fly decorrelation of serially correlated data on a parallel machine, heterogeneous networks, linear scaling, robust/efficient correlation function optimization, and software construction issues (maintainability, extensibility, portability, etc).

CSRI POC: David E. Womble, 9214, 845-7471

Title: Space-Time Multiscale-Multiphysics Computational Techniques

Speaker: Jacob Fish
Rensselaer Polytechnic Institute

Date/Time: Tuesday, December 11, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: In this talk I will describe an adaptive multiscale-multiphysics based design framework aimed at predicting the behavior of structural systems with strong spatial-temporal scale mixing and significant interaction of physical processes. The term multiscale-multiphysics based design framework is coined to emphasize that the behavior of the structure is assessed from the first principles, which are operative at smaller scales than currently resolved in simulations. A number of important applications fall into this category including: 3D woven architectures in aircraft engines, advanced airframes, tires, micro-electronic devices, and porous engineering materials such as honeycombs and truss-like materials. In these structures the size of the microstructure is comparable to that of structural details or to the wavelength of a traveling signal often leading to strong dispersion effects. This is further complicated by the fact that various physical processes, such as deformation, heat conduction, oxidation, stress corrosion, fatigue and fracture are operating at different spatial and temporal scales. The technical challenge is to use modern computing to develop new design concepts where material and structure are viewed as a single system.

CSRI POC: Mark Christon, 09231, 844-8279

Title: Extending the Direct Simulation Monte Carlo method to new Flow Regimes

Speaker: Michael A. Gallis
Microscale Science and Technology Department, SNL

Date/Time: Thursday, March 1, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: The Direct Simulation Monte Carlo method of Bird (DSMC) was introduced in 1969 as a numerical method to simulate hypersonic rarefied flow. Since then DSMC has been extended to cover a large number of problems ranging from plasma flow to subsonic neutral particle transport. Although the method can in principle deal with almost any flow regime, its use is limited because of the very high computational load at high pressures. In this presentation, examples of typical DSMC applications will be given, as well as results from an ongoing effort at SNL to extend the applicability of the method, by reducing the computational load at high pressures and low speeds. The aim of this work is to evaluate the suitability of DSMC, and that of particle simulation codes in general, in the flow regime typically found in MEMS flows.

CSRI POC: Sudip S. Dosanjh, 09233, 845-7018

Title: Solving Computational Genomics Problems on EARTH --
The Evolution of Fine-Grain Multithreading and Its Application

Speaker: Professor Guang Gao
University of Delaware

Date/Time: Monday, June 4, 2001, 11:00 am

Location: Building 980 Rm. 95

Abstract: In this talk, we begin by a brief review on the evolution of fine-grain multithreaded models and architectures --- in particular the EARTH (Efficient Architecture For Running Threads) model developed by the sparker and his associates. We outline the program execution model and architecture issues with fine-grain multithreading to the challenges of both regular and irregular applications. We have implemented the EARTH architecture model on a number of experimental high-performance multiprocessor platforms and we will present some recent experiment results on the effectiveness of fine-grain multithreading on irregular applications. We report our experience of mapping some computational genomics problems on EARTH: in particular the parallel solution of the Smith-Waterman algorithm and the whole genome alignment problem.

CSRI POC: Mark Daniel Rintoul, 09209, 844-9592

Title: Simulations of Peptide Thermodynamics by Replica Exchange

Speaker: Angel Garcia
Group Leader, Theoretical Biology and Biophysics Group, LANL

Date/Time: Wednesday, April 18, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: We study the thermodynamics of 16-21 amino acids long peptides that favorably adopt helical or beta hairpin structures at low T. The systems are simulated in explicit aqueous solution, under periodic boundary conditions, by a highly parallel replica-exchange approach that combines molecular dynamics trajectories with a temperature exchange Monte Carlo process. Replicas are simulated over a Trange of 275-500 K. The combined trajectories in T and configurational space allow a replica to overcome a free energy barrier present at one temperature by increasing T, changing configurations, and cooling in a self-regulated manner, thus allowing sampling of broad regions of configurational space in short (nanoseconds) time scales. The free energy landscape of the beta hairpin forming peptide system over a wide range of temperatures shows that the system preferentially adopts a beta hairpin structure. However, the peptide also samples other stable ensembles where the peptide adopts helices and helix-turn-helix states, among others. The helical states become increasingly stable at low temperatures, but are slightly less stable than the beta turn ensemble. The alpha helical structure forming peptides show a cooperative helix-coil transition. We will examine the role of sequence and force fields in the thermodynamics and kinetics of helix formation.

CSRI POC: Mark Stevens, 09235, 844-1937

Title: AMPL and Stochastic Programming

Speaker: David Gay
Lucent Technologies

Date/Time: Thursday, July 19, 2001, 10:00 am (PST)

Location: Building 980 Room 24 (Sandia-NM)
Building 921 Room137 (Sandia-CA)

Abstract: AMPL is an algebraic modeling language and system for expressing and manipulating linear and nonlinear optimization problems in discrete or continuous variables. AMPL has so far only dealt with deterministic problems, but as interest in stochastic optimization grows, we are working to add facilities that will make working with some kinds of stochastic optimization problems easier. In particular, "random parameters" will permit expressing some multi-stage problems while letting suitable stochastic solvers do their own sampling. This talk gives an overview of AMPL and our current work to extend it.

CSRI POC: Paul Boggs, 8950, (925)294-4630

Title: Short Course on Probabilistic Modeling of Uncertainty

Speaker: Professor Roger Ghanem
Johns Hopkins University

Date/Time: July 17, 19, 24, 26, 2001 (4 Sessions), 1:30 pm

Location: Building 980 Room 95

Abstract: Probability theory provides a rigorous mathematical framework for modeling certain types of uncertainty in physical phenomena. The representation of physical phenomena involving such uncertainties can be worked out in an elegant and mathematically rigorous manner as a problem over a product space involving two different measures. To qualify and quantify the solution to these problems requires the ability to develop approximations over these product spaces. This course will cover the mathematical foundation of approximation theory over product measure spaces with applications to the characterization of the uncertainty in predictions from mechanics-based models of physical phenomena. The topics covered will include: Topological spaces, Hilbert spaces, measure spaces, and product spaces; projections and best approximations; Polynomial chaos approximations in measure spaces; applications to a problems in solid and fluid mechanics and dynamics. Moreover, the concept of stochastic error estimation will be introduced as a parallel to deterministic error estimation. This will permit the design of resource allocation decisions in a manner as to control the overall error budget.

CSRI POC: John Red-Horse, 9211, 845-9190

Title: Dynamic Meshes, Dynamic Interfaces, and Hemodynamics

Speaker: Omar Ghattas
Carnegie Mellon University

Date/Time: Wednesday, June 27, 2001, 2:00 pm

Location: Building 980 Rm. 95

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 graduate student Ivan Malcevic, CMU colleagues Guy Blelloch, Gary Miller, and Noel Walkington, and University of Pittsburgh Medical Center collaborator Jim Antaki.

CSRI POC: Bart van Bloemen Waanders, 09211, 284-6746

Title: Parallel Computers of the 00's: Good news and Bad

Speaker: David Greenberg
Center for Computing Sciences

Date/Time: Monday, February19, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: ASCI red and the Cray T3E's, for better or for worse, marked the culmination of the massively parallel architecture. They combined top of the line commodity processors with interconnects which were (and still are) far ahead of the commodity world. Their programming environments were relatively bare bones but top programmers could get excellent performance. Now we are entering the era of commodity clusters. The IBM-SP, Linux clusters of various stripes, and the new Compaq Sierra Class machines all are scaling to 1000s of processors using interconnects which are best described as "commodity plus". The colony switch, Myrinet, and Quadrics all aim to provide better bandwidth and latency than ethernet but will they be good enough? The performance of SnRad, a radiation solver developed at Sandia, and Metis, a graph partitioner from U. Minnesota, will be used as an indicator of what we should expect from new machines and what we need to push for in their evolution. Performance results will be presented for ports to the Cray T3E and to the Sierra Class machines. We will discuss the use of advanced network features provided through the T3E's E-registers and the Quadrics elan co-processor. Advantages and disadvantages of using MPI and of using UPC and shmem-like calls will be discussed. We can debate whether the use of MPI with efficient sends and receives should drive our machine architecture desires or, if not, what features we will need and why. Inevitably the generality of the machines vendors sell will be bounded above by the requirements we give them. The good news is that better technology exists and is relatively affordable. The bad news is that vendors won't give it to us unless we insist on it.

CSRI POC: Steve Plimpton, 09209, 845-7873

Title: Linux High Availability

Speaker: Joseph Greenseid
Wesleyan University

Date/Time: Tuesday, December 11, 2001, 11:00 am

Location: Building 980, Room 95

Abstract: High Availability is a necessity for at least one, and usually more, systems in any business organization. There are many commercial solutions available for many commercial operating systems. However, with the rise of Linux, Linux based solutions have started to present themselves in this area. My presentation will be about High Availability for Linux.

The presentation will be broken up into three main parts. The first part will be an explanation of what High Availability is. The second part will cover some of the solutions that are currently available for Linux. The last part will be an example of how I tried to implement a Linux High Availability solution, how I failed at it, why I failed at it, and what the failure showed me about High Availability for Linux.

CSRI POC: Erik DeBenedictis, 9223, 284-4017

Title: Implicit-Explicit Splitting Methods for Stiff Ordinary Differential Equations

Speaker: Eric B. Haas
University of California, Los Angeles

Date/Time: Tuesday, August 28, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Stiff systems of ordinary differential equations are characterized as problems with solution components that vary on widely different time scales. Discretization methods for the solution of stiff systems are implicit in nature due to stability requirements. Unfortunately all such methods suffer a serious practical disadvantage in that the solution of an implicit system of nonlinear equations must be solved at each step of the integration process. This constitutes a formidable computational task.

It seems unjustified to solve large stiff initial value problems with fully implicit methods if the rapidly varying solution components originate from only a few distinct terms within the differential system. Therefore we describe a new approach to the problem of numerically integrating stiff systems of ordinary differential equations. In this approach an implicit step is combined with an explicit step to treat the different parts of the differential equation. If this "splitting" is done in an appropriate manner, the resulting method will inherit the stability properties of a stiff solver with the computational efficiency of an explicit solver. These compound algorithms can be designed to have a high order of accuracy.

Results on accuracy and stability are presented for linear multistep and Runge-Kutta type methods. Various extrapolation techniques are also explored. The desirable properties of these splitting methods are verified by asymptotic analysis and numerical examples of ODEs and PDEs via the Method of Lines.

CSRI POC: John Shadid, 9233, 845-7876

Title: PDE based optimization for inverse problems which evolve from Maxwell's equations

Speaker: Eldad Haber

Date/Time: Wednesday, July 25**,** 2001, 9:30 am

Location: Building 921/Room 137

Abstract: Inverse problems involving recovery of distributed parameter functions arise in many applications. Many practical instances require electromagnetic data inversion, where the forward problem comprises Maxwell's equations for moderate frequencies.

Realistic instances of such problems in 3D can be very computationally intensive and require solving systems of PDE's combined with very large scale optimization problems. Therefore, care must be taken in the selection and development of appropriate algorithms.

In this talk I will describe work we have been doing in the context of a project involving geophysical mining and oil exploration applications with the objective of making such computations practically feasible.

- For Maxwell's equations we use a Helmholtz decomposition reformulation to obtain a system which is similar to the Stokes problem. Following a finite volume (or a mixed finite element) discretization suitable for discontinuous coefficients we apply multigrid or ILU block preconditioning to obtain rapid convergence of a Krylov space method.

- For the inverse problem a nonlinear constrained optimization formulation is obtained and the necessary conditions for the inverse optimization problem yield a large, nonlinear, tightly coupled set of PDEs. We devise multigrid methods coupled with preconditioned Krylov solvers for the rapid solution of such systems.

CSRI POC: Co-Hosts: Ray Tuminaro (9214) and Paul Boggs (8950)

Title: Structural Optimization: What has Moore's Law done for us?

Speaker: Professor Raphael T. Haftka
Department of Aerospace Engineering
Mechanics, and Engineering Science, University of Florida

Date/Time: Monday, March 26, 2001, 10:30 am

Location: Building 880 Rm. B-9

Abstract: Using Moore's Law, the speed of computation has increased by about a factor of one million over the past 30 years. However, what is considered to be an adequate structural analysis by the aircraft or automotive companies still takes several hours of computer time. Consequently, many structural analysts continue to claim that structural optimization of meaningful problems is not feasible.

The presentation will survey what can be done in structural optimization in a three dimensional complexity space. One axis defines the complexity of the model--from simple beam models to three dimensional finite element models. A second axis defines the complexity of the structural analysis--from static linear analysis to transient analysis of structures with history dependent material properties. A third axis defines the complexity of the optimization--from deterministic local optimization to global optimization under uncertainty. I will show that while steady progress is being made on moving the problems we can handle away from the origin of the space, our appetite for complexity along all three axes increases apace with computer speed.

CSRI POC: Anthony A. Giunta, 09211, 844-4280

Title: Leveraging Population Information for Robust Design Using Genetic Algorithms

Speaker: Prof. Raphael T. Haftka
Department of Aerospace Engineering, Mechanics and Engineering Science, University of Florida

Date/Time: Thursday, May 3, 2001, 10:30 am

Location: Bldg 823, Rm. 2279

Abstract: Population data can be used to fit response surfaces to objective function and constraints in an optimization problem solved by genetic algorithm. These response surfaces can then used to estimate sensitivity to errors in the design variables and to help obtain a robust design. An example of the optimization of a composite laminate shows that this approach allows the inclusion of the effects of uncertainty with only a small increase in computational cost. For the example, the genetic algorithm required only about 60% more analyses than deterministic optimization. The response surface approximation can be easily updated without keeping track of the entire history of the population. In addition, a simple fading procedure is proposed that can effectively emphasize more recent information in the construction of the approximation.

CSRI POC: Anthony A. Giunta, 09211, 844-4280

Title: Parallel Computations Using Petra

Speaker: Mike Heroux, SNL

Date/Time: Wednesday, February 28, 2001, 12:00 noon

Location: Building 980 Rm. 95

Abstract: The Trilinos/Petra project provides parallel solvers and related tools for application designers and solver library developers. The Trilinos Solver Framework provides a uniform application programmer interface (API) to a wide variety of solver packages. The Petra library provides a common API for construction and use of distributed linear algebra objects and a rich set of basic functions for solver developers. In this talk we focus on the use of Petra for some important parallel computations. In particular we introduce the Petra Import and Export classes and show how they can be used to simplify the construction of matrices and vectors in parallel in a variety of settings, how to compute the transpose of a matrix in parallel and how to redistribute parallel linear algebra objects for better load balancing and algorithm performance.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Multiscale Phenomena: Green's Functions, the Dirichlet-to-Neumann Formulation, Subgrid Scale Models, Bubbles, and the Origins of Stabilized Methods

Speaker: Tom Hughes
Stanford University

Date/Time: Friday, March 2, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: This is an introductory lecture on multiscale concepts for the numerical solution of partial differential equations. An approach is presented for deriving variational methods capable of representing multiscale phenomena. The ideas are first illustrated on the exterior problem for the Helmholtz equation. This leads to the well-known Dirichlet-to-Neumann formulation. Next, a class of subgrid-scale models is developed and the relationships to "bubble function" methods and stabilized methods are established. The identification for stabilized methods leads to an analytical formula for "tau," the intrinsic time scale, whose origins have been a mystery heretofore.

CSRI POC: John N. Shadid, 09221, 845-7876
James Stewart, 09143, 844-8630

Title: Xyce(TM) Parallel Electronic Simulator - A New Sandia Capability

Speakers: Scott Hutchinson, Eric Keiter, Rob Hoekstra, SNL

Date/Time: Wednesday, April 18, 2001, 12:00 noon

Location: Building 980 Rm. 95

Abstract: The Xyce Parallel Electronic Simulator is being written, in collaboration with 01700, to support the simulation needs of the Sandia National Laboratories electrical designers. As such, the development has focused on providing the capability to solve extremely large circuit problems by supporting large-scale parallel computing platforms (up to thousands of processors). In addition, we are providing improved performance for numerical kernels using state-of-the-art algorithms, support for modeling circuit phenomena at a variety of abstraction levels and using object-oriented and modern coding-practices that ensure the code will be maintainable and extensible far into the future.

Although some high-level aspects of circuit simulation mimic those in implicit PDE solvers, there are also some dramatic differences. These stem largely from the extreme heterogeneous nature of the problems and the fact that a system of DAEs, as opposed to PDEs, is being solved. The talks will give an overview of the project and will attempt to highlight these differences with a view towards research needed for efficient solution of these problems.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Estimating One-Parameter Airport Arrival Capacity Distributions Using Stochastic Modeling

Speaker: Tasha Inniss
Trinity College

Date/Time: Tuesday, May 22, 2001, 10:00 am (PST)

Location: Bldg. 940, Auditorium (Sandia - CA)
Bldg. 980, Rm. 24 (Sandia - NM)

Abstract: During instances of capacity-demand imbalances, efficient planning and decision-making in air traffic flow management is contingent upon the "goodness" of the models that estimate airport capacity over time. Airport capacities are subject to substantial uncertainty as they depend on stochastic weather conditions. In this talk, I will present statistical models that we developed using a "seasonal clustering" technique to estimate airport arrival capacity distributions, which are required inputs into a class of stochastic ground holding models that determine the optimal amount of ground delay to assign to incoming flights. A method for adjusting the amount of assigned ground delay in a dynamically changing Ground Delay Program (GDP) will be presented. Based on an algorithm that we developed to compare total weighted delay (after adjusting appropriately), it will be demonstrated that the statistical models yield less total weighted delay than do current operational procedures.

CSRI POC: Monica Martinez-Canales, 08950, (925) 294-3157

Title: Preconditioning KKT Systems

Speaker: Ilse Ipsen
Department of Mathematics
North Carolina State University

Date/Time: Tuesday, July 17, 2001, 10:00 am (PST)

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

Abstract: We discuss the solution of systems of linear equations Ax=b, where the coefficient matrix A is large, sparse, real symmetric indefinite with KKT structure. These matrices arise from interior point methods in a wide variety of applications, including convective heat flow, putting, flight optimization between two cities, linear tangent steering, and space shuttle optimization. The blocks of the KKT matrix consist of structured bands.

We consider solution by preconditioned Krylov space methods. The Krylov space methods include GMRES and CG, while the preconditioner is constructed from a congruence transformation made up of incomplete LDL^T factorizations and judiciously chosen permutations that are designed to drastically reduce the bandwidth.

CSRI POC: Monica Martinez-Canales, 8950, (925) 294-3157

Title: From First Assembly towards a New Cyberpharmaceutical Computing Paradigm

Speaker: Sorin Istrail
Celera Genomics

Date/Time: Wednesday, May 16, 2001, 9:30 am

Location: Building 980 Rm. 95

Abstract: The new science of whole genomics will fundamentally change how pharmaceutical companies pursue the vital challenge of new and better drugs. Target discovery, lead compound identification, pharmacology, toxicology and clinical trials are likely to merge with the science of bioinformatics into a powerful system for developing new pharmaceutical agents. It will be possible to simulate the action of new molecules or therapeutic programs against diverse metabolic pathways prior to preclinical testing. Thus, a paradigm of cyberpharmaceutical testing will be available to the industry, speeding the selection of promising new agents, eliminating products that are likely to exhibit toxicity, and reducing the formidable costs and risks associated with the current paradigm of drug development.

We will report on the Celera's design of a whole genome shotgun assembler and its application to the sequencing of the Drosophila and Human genomes. We will also present some of the major emerging computational challenges of the above paradigm in the exciting new areas of *proteomics*, *structural genomics*, *expression profiling*, *SNPs* and *pharmacogenomics*.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: Exploring Nano-Scale Surface Properties of Functionalized Organic Thin Films

Speaker: Shaoyi Jiang, Assistant Professor
Dept. of Chemical Engineering, University of Washington

Date/Time: Friday, April 13, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Recent advances in molecular simulations and microscopic experiments have progressed to the point where there are enormous potential opportunities for application to numerous important problems involving materials, chemical, and biological sciences and technology. Our work focuses on exploring nano-scale surface properties of functionalized organic thin films for various applications using an integrated molecular simulation (molecular dynamics and Monte Carlo simulations, ab initio quantum mechanics, accurate force fields, and cell-multipole method) and experimental (scanning probe microscopy, surface plasmon resonance biosensor, and X-ray photoelectron spectroscopy) technique. In this talk, I will give an overview of research activities in my research group, including adsorption and fluid flow in micropores, friction properties of thiols/Au(111) and alkyl/Si(111), phase behavior of mixed thiols/Au(111), guest-host interactions in dendrimers, and molecular recognition for sensing and detection.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: Synthetic Implementation of Massively Parallel (MP) System Performance Tools

Speaker: Chu Jong
University of New Mexico

Date/Time: Tuesday, October 23, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Writing parallel and distributed scientific applications for High-Performance Computing (HPC) systems that make optimum use of computing resources is a very challenging task. Tools that assist performance monitoring and performance tuning can help programmers identify where, when, and how their application mismanage the system resources. Many parallel performance tools are available for computing systems with a low degree of parallelism. When using these tools on MP systems, however, the effectiveness of the tool diminishes quickly as the number of processor increases. Among system performance tools, very few can handle up to a hundred processors and none have demonstrated the ability of accommodating more than a thousand processors.

We have studied the scalability issues of MP system performance tools and proposed a fundamental solution, replacing the two-level data collection structure with a hierarchical data collection structure. We constructed three models, analytical model, simulation model, and implementation model, to generate user response time of MP system performance tools. From the results of our models, we proved that a hierarchical data collection structure increased the scalability of MP system performance tools in orders of magnitude.

This talk will focus on the implementation; I will address the implementation methods, the virtualize mechanisms that we used to construct virtual nodes from a single one, and the implementation results.

CSRI POC: Ron Brightwell, 9223, 844-2099

Title: Multiphysics Modeling Using the Finite Element Method and FEMLAB

Speaker: David Kan and Wayland Oong
Comsol, Inc.

Date/Time: Thursday, September 20, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: In this seminar, we will present a study of mathematical modeling in different fields of engineering and science. Specifically, we will model electromagnetic wave propagation, transport phenomena in chemical reactors, and stresses in solid structures under mechanical loads. All models will be realized through the use of the finite element package FEMLAB. We will consider the concept of multiphysics and its definition in the FEMLAB environment. The models will be built interactively from scratch so that the audience can give their input to alter the modeling process. Applications that will be modeled are:

- Wave propagation in microwave and opto-electronic systems
- Transport phenomena coupled to chemical reactions in monolithic reactors
- Displacements and stresses in mechanics
- Fuel cells

What is FEMLAB?
FEMLAB comprises tailored modeling solutions for Structural Mechanics, Chemical Engineering, and Electromagnetics. It facilitates modeling of all physical phenomena based on Partial Differential Equations (PDEs) such as heat transfer and fluid flow. The model equations, or physics, can either be applied by the user in a ready-to-use form, or specified freely to suit any type of physical phenomenon (linear, non-linear or time dependent). Several problems can be combined and coupled in a single model multiphysics, the benefit being a very straightforward modeling process with a minimum of non-realistic assumptions. Highlights of the new FEMLAB 2.2 are:

- New improved meshing algorithm in 3D
- New solvers for large 3D models
- IGES import and export
- New element library
- Extended multiphysics by coupling of 0D, 1D, 2D and 3D problems in a single model
- New postprocessing features in 1D, 2D and 3D.

CSRI POC: David L. Ropp, 9214, 845-7431

Title: An Infeasible Point Method for Binary Programs

Speaker: Anthony J. Kearsley
Carnegie Mellon University and
National Institute of Standards and Technology

Date/Time: Tuesday, June 19, 2001 2:00 pm (PST)

Location: Bldg. 940, Auditorium (Sandia - CA)
Bldg. 980, Rm. 24 (Sandia - NM)

Abstract: In this talk we introduce an approach for solving a class of interesting and challenging 0-1 programs. While by today's computational standards these programs are small in terms of numbers of variables and constraints, the developing and implementing of efficient numerical methods for locating optimal solutions still eludes researchers. The suggested method is based on a reformulation of these problems through a smooth approximation of the objective functions and an expansion of the number of variables that couple through the constraints. This formulation can be solved, in turn, using constrained optimization algorithms for nonlinear programming problems. Preliminary numerical results of employing a Sequential Quadratic Programming algorithm as a nonlinear solver are presented.

CSRI POC: Paul Boggs, 08950, (925) 294-4630

Title: N-Additive Runge-Kutta Methods for the Compressible Navier-Stokes Equations

Speaker: Christopher A. Kennedy, SNL
Combustion Research Facility, Livermore, California

Joint work with Mark H. Carpenter
NASA Langley Research Center
Aeronautics and Aeroacoustic Methods Branch

Date/Time: Monday, February 5, 2001, 9:30 am

Location: Building 980 Rm. 95

Abstract: Additive Runge-Kutta (ARK) methods are considered for application to the spatially discretized compressible Navier-Stokes equations. First, accuracy and stability are considered for the general case when N different Runge-Kutta methods are grouped into a single composite method. Then, implicit-explicit, N = 2, additive Runge-Kutta (ARK_2) methods from third- to fifth-order are presented that allow for integration of stiff terms by an L-stable, stiffly accurate explicit, singly diagonally implicit Runge-Kutta (ESDIRK) method while the nonstiff terms are integrated with a traditional explicit Runge-Kutta method (ERK). Coupling error terms are of equal order to those of the elemental methods. Derived ARK_2 methods have vanishing stability functions for very large values of the stiff scaled eigenvalue, $z^{[I]} \rightarrow -\infty$. All constructed methods retain high stability efficiency in the absence of stiffness, $z^{[I]} \rightarrow 0$. Extrapolation-type stage value predictors are provided based on dense output formulae. Methods have been optimized to minimize the leading order ARK_2 error terms, minimize the size of the Butcher coefficients, and maximize the conservation properties. Numerical tests of the new schemes on a chemical reaction inducing propagating shock wave and a two-equation example of a singularly perturbed initial-value problem confirm the predicted stability and accuracy of each of the methods.

CSRI POC: John N. Shadid, 09233, 845-7876

Title: APPSPACK: A Software Package for Asynchronous and Fault-Tolerant Parallel Pattern Search for Derivative-Free Optimization

Speaker: Tamara G. Kolda
Computational Sciences and Mathematics Research (08950)
Sandia National Labs, California

Date/Time: Tuesday, April 3, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: APPSPACK implements an asynchronous and fault tolerant parallel pattern search method for derivative-free optimization. This talk will explain what types of problems APPSPACK is most useful for. In particular, we will focus on engineering design and parameter fitting problems characterized by a small number of variables, expensive objective function evaluations, and no gradient information. We will give a basic overview of the algorithm, details of the software implementation, and numerical results. We will also query the audience to find out what types of problems they are interested in solving and what features would be most useful in future releases.

For more information on APPSPACK, visit http://csmr.ca.sandia.gov/projects/apps.html

CSRI POC: Tony Giunta, 09211, 844-4280

Title: A Software Reengineering Environment for Scientific Computing

Speaker: Suraj Kothari
Iowa State University

Date/Time: Wednesday, May 16, 2001, 11:00 am

Location: Building 980 Rm. 95

Abstract: We are developing a Software Reengineering Environment (SRE) equipped with specific capabilities for scientific computing including domain-specific automatic parallelization, coupling of models, restructuring of programs, and program comprehension. In this talk I will describe the SRE and how it is applicable in practice. A key idea behind our SRE is the use of domain-specific knowledge to address software reengineering problems, which are otherwise intractable. It enables a combination of the power of automation and high-level reasoning by human experts to yield solutions to such difficult reengineering problems as automatic parallelization. The domain-specific knowledge is about a class of problems. Thus, the reengineering process can be applied to not just one problem, but a class of problems. Moreover the framework is extensible to continue to incorporate new classes of problems. The SRE includes an interactive and integrated environment for program comprehension. It provides information filters to extract useful information from the program analysis data, which is often very large and complex.

The automatic parallelization capability is currently applicable to explicit finite difference programs. The input is a Fortran 77 program and the output is a SPMD program that uses MPI for message passing. We will present some results of parallelization experiments with well-known numerical weather prediction models. The talk will be accompanied by an online demonstration of the program comprehension and automatic parallelization capabilities of the current version of SRE.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Albuquerque High Performance Computing Center (AHPCC)

Speaker: Patricia Kovatch
University of New Mexico

Date/Time: Friday, May 18, 200l, 9:00 am

Location: Building 980 Rm. 95

Abstract: The Albuquerque High Performance Computing Center (AHPCC) has designed several production Linux superclusters for scientists and researchers to run a variety of parallel applications. The goal of these clusters is to provide easy-to-use high performance computing systems at reasonable prices. Superclusters are large scale clusters built from commodity parts with high performance interconnects, large storage and remote management tools. Details on the design, implementation and management of these systems will be discussed using examples from AHPCC's largest supercluster - LosLobos. This cluster has 512 733 MHz processors with 256 dual processor nodes.
Setting up the user application environment and tuning the systems for optimal performance will also be reviewed.

CSRI POC: Ron Brightwell, 09223, 844-2099

Title: Memetic Algorithms: Syntactic Model and Taxonomy

Speaker: Natalio Krasnogor
University of the West of England

Date/Time: Wednesday, July 27, 2001 3:00 pm

Location: Building 980 Room 95

Abstract: Genetic algorithms (GAs) combined with local search were named "Memetic Algorithms" (MAs). These methods are inspired by models of adaptation in natural systems that combine evolutionary adaptation of populations of individuals with individual learning within a lifetime. Additionally, Mas are inspired by Richard Dawkins concept of a meme, which represents a unit of cultural evolution that can exhibit local refinement.

In this talk we will review some works on the application of MAs to well known combinatorial optimization problems and the architecture of these algorithms will be studied. A syntactic model will be defined and a classification scheme based on a computable index D will be given for MAs. The existence of both a model and a taxonomy for MAs is of theoretical and practical relevance, i.e., they allow for more sensible and "fair" comparisons of approaches and experiment designs while providing with a suitable tool for developing novel MAs for new problems. In essence we introduce here a research program on Memetic Algorithms. (Time permitting) A particular insight on the new kind of MAs that can be extrapolated from our model will be given, specifically we will describe the "Simple Inheritance Multimeme Algorithm" for TSP and Protein Structure Prediction.

CSRI POC: Bob Carr, 9211, 845-8562

Title: A Posteriori error estimate for front-tracking for hyperbolic systems of conservation laws

Speaker: Marc LaForest
Department of Mathematics, Colorado State University

Date/Time: Thursday, September 27, 2001, 2:20 pm

Location: Building 980 Room 95

Abstract: In this talk, we present an a posteriori error estimate in the L^1 norm for front-tracking approximate solutions to hyperbolic systems of conservation laws. This result extends the L^1 stability theory of Bressan, Liu, and Yang and depends in an essential way on their L^1-equivalent functional. Our method shows that with the help of their functional it is possible to distinguish between the numerical error and the amount of entropy produced by the underlying physical system. The front-tracking approximations to which this estimate applies are piecewise constant approximations closely related to those of Glimm's scheme. Under the assumption of small total variation, the nonlinearities generated during wave interactions can be controlled in these approximations. The numerical error is measured by a quantity similar to the residual but computed from an exact Riemann solver.

CSRI POC: Mark A. Christon, 9231, 844-8279

Title: A Short Review of Semiconductor Device Modeling Codes

Speaker: Glenn Laguna, SNL

Date/Time: Thursday, April 19, 2001, 1:00 pm

Location: Building 980 Rm. 24

Abstract: Semiconductor device modeling codes attempt to solve the Poisson and continuity equations for charge carriers in devices as simple as a diode or as complex as a memory cell. Sandia's microelectronics center has used these kinds of codes, both for the prediction of operating characteristics of devices fabricated here, and analysis of the immunity of such devices to radiation effects such as single event upset. In this talk, I'll look at several commercial and academic codes that I have used in the past two years, with emphasis on what a user of the codes needs and which, if any, of the codes might be good candidates for parallelization.

CSRI POC: Neil Pundit, 09223, 844-2332 and Doug Doerfler, 09224, 844-9528

Title: Scalability Matters: Why we need to make bioinformatics programs scalable, and results from work on NCBI BLAST on various architectures

Speaker: Joseph I. Landman
Silicon Graphics

Date/Time: Tuesday, April 3, 2001, 11:00 am

Location: Building 980 Rm. 95

Abstract: The volume of genomics and proteomics data is growing faster than the ability of the canonical applications to process it. In order to analyze this exponentially growing volume of information, scalable algorithms need to be developed and be deployed on scalable computational platforms. The need for scalable algorithms arises from the nature of the exponentially growing volume of data, as higher throughput data collection techniques come online, the information growth rate also increases. This increasing growth rate stresses traditional algorithms. Algorithmic or programmatic inefficiencies and rate-limited processes ultimately will determine how much analysis can occur. This work focuses upon efforts initiated to overcome the non-scalable algorithmic limitations in BLAST, and adapt it to scalable computational platforms. Performance results will be presented, as well as a discussion on the need for scalability in bioinformatics, genomics, and proteomics programs.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: Modeling brittle fracture growth as an energy minimization problem

Speaker: Christopher J. Larsen
Math Department, Worcester Polytechnic Institute

Date/Time: Monday, September 10, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: There are many difficulties in modeling growth of brittle fracture of elastic materials. The main model for fracture growth, based on Griffith's criterion, addresses only the rate of crack growth, but not the direction, branching, or "brutal" crack formation. I will describe a model based on total energy minimization that does not require a crack path to be specified a priori, nor the presence of an initial crack. I will also address the problem of existence for this model.

CSRI POC: Bob Carr, 9211, 845-8562

Title: Pattern Search Methods for Nonlinear Programming

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

Date/Time: Thursday, March 8, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: We will discuss pattern search methods for general nonlinear programming. Pattern search methods are a class of derivative-free, direct search algorithms for nonlinear optimization. They are attractive for certain problems due to their relative simplicity and their applicability to problems where derivatives are unavailable or untrustworthy. They also possess attractive scalability properties for parallel computation.

The first part of the talk will briefly review pattern search methods for unconstrained optimization. We will also note the differences between unconstrained and constrained optimization that are of particular importance to the development of derivative-free methods for constrained optimization. We then discuss our recent work (with Virginia Torczon) on pattern search methods for general nonlinear programming.

Our work on extending pattern search methods to constrained optimization has proceeded along two lines. For bound and linearly constrained optimization, we have developed feasible iterate pattern search algorithms. Key to the algorithms is the way in which the local search patterns conform to the geometry of the boundary of the feasible region.

For general nonlinearly constrained problems we have developed a pattern search method based on an augmented Lagrangian method originally due to Conn, Gould, and Toint. The algorithm proceeds by successive bound constrained minimization of an augmented Lagrangian using a pattern search method. The stopping criterion proposed by Conn, Gould, and Toint for the solution of this subproblem requires explicit knowledge of derivatives. Such information is presumed absent in pattern search methods; however, we show how we can replace this with a stopping criterion based on the pattern size in a way that preserves the convergence properties of the original algorithm.

CSRI POC: Tony Giunta, 09211, 844-4280

Title: Time Dependent Calculations for Incompressible Flows on Adaptive Unstructured Meshes

Speaker: Paul Lin
Department of Mechanical and Aerospace Engineering
Princeton University

Date/Time: Tuesday, March 27, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Time dependent phenomena are important in a variety of incompressible flows, including free surface flows, unsteady wake flows, and time dependent geometries such as propellers. One difficulty in calculating time dependent flows is that the location of important flow features such as vortices or waves varies spatially with time. A grid providing an accurate resolution of the flow field at one time step may prove to provide insufficient resolution or be inefficient several time steps later. A solution to this problem is to use a grid which is also time dependent and can distribute points spatially depending on the current flow field. The specifics of how the enrichment and coarsening process depend on the local flow field gradients as well as the construction of techniques for maintaining a high quality unstructured mesh are crucial to the success of a mesh adaptation method.

This seminar will present an implicit multigrid driven algorithm for two-dimensional incompressible laminar viscous flows which has been coupled with a Delaunay based mesh adaptation method to better capture the important features of the flow. The method is validated by comparison with experimental results of low Reynolds number flow over a shedding circular cylinder.

CSRI POC: Andy Salinger, 09233, 845-3523

Title: I'd like to order one rack of LAM, well-done

Speaker: Dr. Andrew Lumsdaine and Jeff Squyres
Computer Science Department, Indiana University

Date/Time: Monday, July 2, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: LAM/MPI is a high quality open-source implementation of the MPI standard that is developed and maintained at Indiana University. LAM offers a complete implementation of MPI-1 and many features from MPI-2 including dynamic processes, one-sided communication, C++ bindings, and parallel IO.

This talk will present a brief overview of features available in LAM, followed by an in-depth discussion about how LAM/MPI is structured and implemented. MPI is actually the highest layer in LAM; it is built on top of Trollius, a user-level, daemon-based run-time environment. Trollius will be explored, and examples will be given to show how it provides many of the services used in MPI and the rest of LAM's run-time environment. Finally, LAM's abstraction for the communication layer in MPI -- the Request Progression Interface (RPI) -- will be discussed.

CSRI POC: Richard R. Drake, 9231, 844-2537

Title: Techniques in Support Vector Classification: Diverse Mathematical Ideas Applied to Practical Problems

Speaker: Shawn Martin, SNL

Date/Time: Friday, April 20, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: In this talk I will present the results in my Ph.D. dissertation, "Techniques in Support Vector Classification." Since classification is not a primary concern at Sandia National Labs, I will focus on the mathematical ideas applied in my dissertation rather than the relevance of my work to Support Vector Classification.

I will first present an application of the Veronese map from algebraic geometry to a problem in feature selection for Materials Design. Next I will present an application of the Gram-Schmidt procedure from linear algebra to the problem of kernel selection for Support Vector Machines. Here I must provide some background on Support Vector Classification, but I will also show how this result can be applied to a wide range of problems outside of classification. Finally, I will present applications of computational geometry and optimization to the development of a fast algorithm for training Support Vector Machines.

If time permits, I will also present some ideas (in various stages of completion) for future work. Included here is a very interesting idea for fitting curves/surfaces to data.

CSRI POC: William Hart, 09211, 844-217

Title: Domain Decomposition Preconditioners for Hermite Collocation Problems

Speaker: Gabriel Mateescu
National Research Council
Ottawa, Ontario, Canada

Date/Time: Monday, June 4, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: The computational kernel of scientific codes is often a linear system solver. A truly scalable solver has an O(N) number of operations, where N is the number of unknowns, and has a linear speedup of the parallel implementation. Krylov subspace methods may lead to scalable solvers when provided with robust and efficient preconditioners.

One-level domain decomposition preconditioners have good parallelism but lack robustness, while pure multigrid methods have good operation count but may require problem-specific choices for the smoother, restriction-prolongation operator, or grid coarsening. Multilevel domain decomposition is emerging as a mainstream strategy for robust parallel preconditioners. General purpose parallel linear algebra packages such as PETSc provide support for multilevel domain decomposition, but key components such as the coarsening strategy need to be provided by the user.

We describe two parallel domain decomposition preconditioners for linear systems of equations arising from the piecewise Hermite bicubic collocation discretization of second-order elliptic PDEs defined on a rectangular domain and with general boundary conditions. The preconditioner uses a three-level discretization scheme. In the framework of substructuring, we partition the domain into subdomains, edges, and vertices. The three-level discretization scheme employs a fine grid with step h on the first level; the second level is the edge-grid which has step h along each edge and step H on the perpendicular direction; the third level is the coarse grid with step H.

The construction of the preconditioner is performed in three main steps. In the first step, the unknowns and equations are ordered in a non-finite-element order, which induces an arrowhead structure of the operator matrix. In the second step, the edge and vertex subproblems are approximated by using the edge and coarse grid discretizations. In the third step, a global preconditioner is defined in terms of the original operator and operators arising from the special discretization. The special discretization provides the following benefits: it decouples the vertex subproblem from the subdomains and edges, decreases the coupling between the edge subproblems and the subdomains, and makes the edge subproblems mutually independent.

The resulting global preconditioner is robust and efficient. We describe the parallel implementation of the method and evaluate its performance using numerical experiments.

CSRI POC: Scott Hutchinson, 09233, 845-7996

Title: Development and Performance of Parallel Unstructured Mesh Multigrid Solution Techniques

Speaker: Dimitri Mavriplis
ICASE, and NASA Langley Research Center

Date/Time: Thursday, February 22, 2001, 1:00 pm

Location: Building 980 Rm. 95

Abstract: The development of an unstructured agglomeration multigrid algorithm will be discussed. This algorithm has been developed initially as a solver for steady and unsteady Reynolds averaged Navier-Stokes problems, but recently has also been applied to Radiation-Diffusion problems on unstructured meshes. Agglomeration multigrid constructs coarse grid levels using a graph algorithm which identifies groups of cells or control volumes to be merged together or agglomerated to form fewer but larger coarse level control volumes. Coarse level equations are obtained by Galerkin projection, using piecewise constant restriction and prolongation operators. The algorithm is closely related to algebraic multigrid, but the coarsening phase depends only on the grid characteristics rather than the stencil coefficients, thus resulting in static coarse levels for non-linear problems. The algorithm can be formulated as a non-linear solver (FAS multigrid), a linear solver (Correction Scheme), or as a preconditioner to a Krylov method, and efficiency comparisons using these three variants will be discussed. For anisotropic problems, directional (weighted-graph) coarsening as well as directional (line-based) smoothers have been developed, and convergence rates which are insensitive to the degree of grid stretching are demonstrated. The algorithm is parallelized through domain decomposition, using MPI and/or OpenMP for communication between the partitions assigned to individual processors. Scalability benchmarks on all three ASCI machines using up to 2048 processors are given for a large scale aerodynamic simulation. Comparisons of MPI versus OpenMP and combinations of MPI/OpenMP in a dual level hybrid parallel mode are also given.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: The DOGMA Parallel Computing Environment for Computational Genomics using Dynamic and Dedicated Clusters of Workstations

Speaker: David McLaughlin
Brigham Young University

Date/Time: Monday, November 26, 2001, 3:15 pm

Location: Building 980 Room 95

Abstract: The Distributed Object Group Metacomputing Architecture (DOGMA) provides an environment for users to deploy their pre-existing executables for parallel, distributed execution as well deploy newly written parallel programs. The target machines of the DOGMA systems are idle campus workstations as well as dedicated clusters and supercomputers. DOGMA is a task-based system, restarting failed tasks on other nodes to ensure all tasks get executed. The resource management portions are separated into a program called the BYUResourceManager which is designed to use a plugin scheduler, like the Maui Scheduler.

Because computing the phylogenies in computational biology is an NP-complete problem, only very small data sets can be computed exactly. And so approximation algorithms are used, some of which are not very scalable. More efficient algorithms are being explored. Solution surfaces is an attempt to visualize where to concentrate searching efforts for problems with large solution spaces. DNA sequence alignment, also an NP-complete problem, is essential to producing more optimal phylogenies. Historically, the phylogeny reconstruction problem has been performed in two stages: 1) aligning the sequences and then 2) searching for more optimal trees according to an optimality criterion. It has been shown, however that sequence alignment affects the search and should be performed at the same time. Some algorithms do this approach, and improvements to these are being explored."

CSRI POC: Erik DeBenedictis, 9223, 284-4017

Title: Finding Strongly Connected Components in Parallel in Particle Transport Sweeps

Speaker: William McLendon
Texas A & M University

Date/Time: Tuesday, July 10, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Discrete ordinates methods are commonly used to simulate radiation transport for fire or weapons modeling. The computation proceeds by sweeping the flux across a grid. A particular cell cannot be computed until all the cells immediately upwind of it are finished. If the directed dependence graph for the grid cells contains a cycle then sweeping methods will deadlock. This can happen in unstructured grids and time stepped problems where the grid is allowed to deform. To prevent this deadlock from occurring, we must detect and eliminate these cycles before the sweep is performed.

The number of cycles can be exponential in the number of vertices but the number of strongly connected components (SCCs) is at most linear in the number of vertices since a vertex is in at most one SCC. Therefore we are interested in finding all the SCCs of a directed graph.

A SCC of a directed graph G=(V,E) is a maximal set of vertices U Í V such that for every pair of vertices u and v in U, we have directed paths from u to v and from v to u.

Tarjan's classic serial algorithm for detection of SCCs runs linearly with respect to the number of edges and uses depth first search. However, depth first search is known to be difficult to parallelize – the special case of lexicographical depth first search is P-Complete, which means it is unlikely that a scalable parallel algorithm exists.

We have implemented a modified version of the Divide and Conquer Strong Components (DCSC) algorithm of Fleischer et al. It is a recursive divide and conquer approach for finding SCCs that does not rely on depth-first search. ModifiedDCSC was developed in C and MPI on the ASCI Red supercomputer maintained at Sandia National Laboratories and is currently in the process of being integrated into a multiphysics code.

CSRI POC: Glenn Laguna, 9224, 844-2308

Title: Impact of Sodium on Intracellular Calcium and its Implications in Neurotransmitter Release

Speaker: Shawn A. Means
University of New Mexico

Date/Time: Tuesday, March 13, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Crucial to the function of the nervous system in animals, neurotransmitter release propagates the signals between neurons, is instrumental in receiving information from sensory organs, and maintains control of muscles. A general description of the process involved in release of the chemicals and/or proteins known as neurotransmitters is well established. Calcium ions flood the terminus of a neuron and initiate machinery in the cell to release these neurotransmitters, which then diffuse across a miniscule gap of intercellular space and interact with the receiving neuron, sensor cell or muscle and hence transmit information between cells.

Whereas calcium is involved in intercellular communication, the sodium ion is the primary instrument for signal propagation down a neuron. After receiving an intercellular signal via neurotransmitter release, a neuron transmits this signal to the next cell by a sequence of sodium influxes triggered in domino-effect down its body and axon, where the calcium mediated release occurs again. However, sodium's involvement in the actual release of neurotransmitter beyond its role in signal propagation down the neuron is unknown. Evidence exists suggesting that sodium has indeed a modulatory control over neurotransmitter release in addition to calcium.

An exploration of a possible role in the impact of sodium on intracellular calcium is performed, utilizing a continuum, diffusion-reaction representation of a neuron's terminus (the site of neurotransmitter release) with two ionic species (Sodium & Calcium) and a protein species (Calmodulin) reacting with calcium. The boundary conditions are of special interest, where both point sources and sinks dramatically influence the interior diffusions and reactions. These nonlinear boundary conditions include depictions of ion channels (voltage-gated sodium and calcium), ionic pumps (calcium ATP-ase, sodium/potassium ATP-ase) and the mechanism of interest, the sodium/calcium Exchanger. This exchanger's direction of flux reverses under certain conditions which potentially affects neurotransmitter release by increasing intracellular calcium via indirect dependence on intracellular sodium.

CSRI POC: John Aidun, 09235, 844-1209

Title: Galerkin/Finite Element Analysis of Two-Layer Slot Coating

Speaker: Lawrence C. Musson
Dept. of Chemical Engineering & Material Science
University of Minnesota

Date/Time: Wednesday, April 11, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Deposition of multiple miscible liquid layers onto a substrate is a frequent goal in industrial coating processes. Coating and drying the layers together can greatly improve productivity and quality; it can also enable thinner layers to be coated than can be deposited alone at comparable speeds. In slot coating, liquid is pumped through a feed slot in an applicator die onto a substrate translating in its own plane past the die. Two layers of liquid can be deposited onto the substrate simultaneously by employing an applicator die with a pair of feed slots. The liquid bounded by the confining surfaces, i.e. the substrate and die, and upstream and downstream menisci, is called the coating bead. Coating is successful only when the bead is stable and can be achieved only within a range of operating parameters.

In this research, predictions of perfectly steady two-dimensional operating states were made by solving the Navier-Stokes system with the Galerkin finite element method and were accompanied by (1) arc-length parameter continuation of the ratio of upper to lower-layer liquid viscosity; and (2) evaluation of the consequences of ever-present three-dimensional perturbations, by linear stability analysis, and of ongoing small forced disturbances, by frequency response analysis. It was found that, for the flow rates tested, only over a limited range of viscosity ratio of about $0.27 < m < 2.2$ was the bead stable; outside of this range, unstable modes appeared at the upstream meniscus as it located near the bounds of the upstream slot gap. However, it was found that the range of viscosity ratio over which the bead was stable could be increased by manipulation of other operating parameters.

CSRI POC: Rod Schmidt, 09233 (Alb: 845-3057, offsite: 801-733-8568)

Title: Multigrid Algorithms for Discretized Optimization Problems

Speaker: Stephen Nash
George Mason University

Date/Time: Tuesday, October 23, 2001, 10:00 am (PST)

Location: Bldg. 940, Auditorium (IMTL) (Sandia - CA)
Bldg. 980, Room 24 (Sandia - NM)

Abstract: Many large optimization problems represent a family of models of varying size, corresponding to different discretizations. An example is the optimization of systems governed by differential equations. In such problems, one has a set of design variables along with a set of state variables, the two sets of variables being related through a set of differential equation constraints. The overall computational cost of optimization is determined by the level of discretization used to numerically solve the governing differential equations. If a fine discretization is used, one expects a greater degree of physical and mathematical fidelity to the problem under consideration, but the large number of state variables can make the cost of optimization prohibitive. We present a multigrid algorithm that uses solutions to optimization problems based on coarser discretizations, which are less expensive to compute, in a systematic manner to obtain the solution of the optimization problem based on a finer discretization. Of interest is the fact that the approach is applicable in situations where multigrid applied only to the solution of the differential equation might not be effective. We give evidence (both theoretical and numerical) that a multigrid approach can often be successful in the more general setting of optimization, and that the optimization setting offers a number of practical advantages for solving problems of this type.

CSRI POC: Paul Boggs, 8950, 294-4630

Title: Digital Heterodyne Filtering

Speaker: Karl Nelson
University of California, Davis

Date/Time: Thursday, March 29, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: Presents the development and application of adaptive heterodyne filter for use in Adaptive Line Enhancers. Adaptive heterodyne filter provide a means of translating a high order IIR filter with adaptation to enhance or attenuate sinusoids. Further, demonstrates properties of similar adaptive line enhancer filters and the existence of unique solutions for 2nd order IIR constrained filters.

CSRI POC: Scott A. Hutchinson, 09233, 845-7996

Title: Parallel Unstructured AMR and Myricom Networking for Beowulf Clusters

Speaker: Charles Norton and Thomas Cwik
Jet Propulsion Laboratory

Date/Time: Wednesday, August 29, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: We report on current activities in the development of the PYRAMID parallel unstructured AMR library. This Fortran-90 based library supports adaptive refinement, mesh migration, load balancing, partitioning, and automatic mesh quality control all in parallel. Current and future design issues are described including performance metrics associated with our transition to the 2.0 Gbit/s Myricom 2000 network on our 800 Mhz dual-processor Pentium-III Beowulf cluster.

CSRI POC: Bruce Hendrickson, 9226, 845-7599

Title: Dimensional Analysis Applied to Computer Performance Analysis

Speaker: Robert Numrich
CRAY

Date/Time: Tuesday, December 4, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: This talk will apply dimensional analysis to computer performance analysis. Dimensional analysis has been used in physics and engineering to express fundamental differential equations in dimensionless form so that whole families of systems can be solved at once rather than one at a time. It has also been used to discern the form of unknown equations based solely on the requirement of dimensional consistency. These same methods can be applied to computer performance modeling for both hardware and software. This talk will consider ways of applying the fundamental theorem of dimensional analysis, the Pi Theorem, to develop a simple model for memory contention for a shared memory multi-processor and for a simple model for remote communication.

CSRI POC: Erik DeBenedictis, 9223, 284-4017

Title: Structure-Based Design of Inhibitors of Acetylcholinesterase, Farnesyltransferase and Adenovirus Proteinase

Speaker: Yuan-Ping Pang, Ph.D.
Mayo Clinic Cancer Center
Department of Molecular Pharmacology and Experimental Therapeutics

Date/Time: Tuesday, March 13, 2001, 1:00 pm

Location: Building 980 Rm. 95

Abstract: Docking programs are used to search chemical databases for compounds that bind to specific 3-dimensional structures, e.g. the active site of an enzyme. Properties of EUDOC, a new docking program, will be described and examples will be given on how it has been successfully used to obtain new therapeutic agents.

Tacrine is a drug use to treat Alzheimer's disease. EUDOC found a second tacrine inding site in acetylcholinesterase. Tethering two tacrines with a chemical chain resulted in a dimeric analog 1,000-fold more potent and 10,0000-fold more selective than monomeric tacrine. The theoretical basis for this enhanced potency will be discussed.

Farnesyltransferase catalyses the covalent addition of the farnesyl group to a cysteine residue in proteins such as Ras. Using an in silico screening approach, the dimerization strategy, and combinatorial chemistry, a library of discrete dimers is being screened for potential farnesyltransferase inhibitors that will act anticancer agents. A new type of proteinase inhibitor has been discovered using the three-dimensional structure of the adenovirus proteinase and EUDOC. The drug binds close to the active site nucleophile and then irreversibly attacks it. The specificity of the drug is in the reversible binding. Only if it binds within 5 Å of the nucleophile does irreversible inhibition occur. The lead compound is currently being tested by the NIH as an anti-viral agent.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: Design and Performance Evaluation of Distributed Shared Memory Multiprocessors

Speaker: Hitoshi Oi

Date/Time: Monday, January 15, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: With a distributed shared memory (DSM) multiprocessor, programmers are provided with a globally shared memory address space on top of physically distributed memory modules. This view of the shared memory space is accomplished by sending request and data messages over the interconnection network that connects processing elements. Thus, the cost of inter-processor communication has a significant effect on the DSM multiprocessors.

In this talk, the following aspects of DSM multiprocessors that affect the inter processor communication are discussed: (1) interconnection network, (2) cache coherence protocol, (3) on-chip integration. The effects of these design choices on the performance are studied by analytical model and simulations.

CSRI POC: Scott Hutchinson, 09233, 845-7996

Title: High-performance Computing on Commodity-based Clusters

Speaker: Scott Pakin
University of Illinois at Urbana-Champaign

Date/Time: Friday, July 20, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Clusters of personal computers and high-speed networks are rapidly becoming a dominant platform for high-performance computing. However, the allure of good price:performance ratios often overshadows the difficulty of attaining good absolute performance from commodity components. In this talk, I will discuss my work in the area of high-performance computing on commodity-based clusters. First, I will describe Fast Messages, a low-level messaging layer designed specifically to be used as a building block for high-speed implementations of higher-level communication layers, such as MPI. Second, I will show how Fast Messages was used to support Dynamic Coscheduling -- coordinated thread scheduling across a cluster without the shortcomings of gang scheduling. And finally, I will present "nonblocking barriers," a new technique designed to complement coordinated thread scheduling by tolerating unresponsiveness that is internal to the application or that cannot be eliminated merely by coscheduling its threads.

CSRI POC: Ron Brightwell, 9223, 844-2099

Title: UIcluster Application

Speaker: Kevin T. Pedretti
University of Iowa

Date/Time: Tuesday, May 8, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: Mr. Pedretti's talk will present the UIcluster application developed as part of his thesis research. This program partitions EST (Expressed Sequence Tags) and other genetic sequences into "clusters" based on sequence similarity. Ideally, each cluster will contain sequences that all represent the same gene. In practice, this is not possible. The goals of UIcluster are to form clusters as accurately as possible and be computationally efficient. The importance of this result bears on several aspects, but the principle of these is creating non-redundant indices of genes. Such indices are an essential tool for assessing sequencing novelty rates and guiding geneticists in their research efforts.

If a naive approach such as an NxN comparison (N is the number of sequences input to the application) is taken, the problem is only feasible for very small data sets. For larger data sets containing tens or hundreds of thousands of sequences, a more scalable approach must be taken. UIcluster has been developed over the course of four years to solve this problem. Each version of the program has built upon the prior by incorporating new functionality and more sophisticated optimizations. The latest version of the application has been parallelized using the MPI (message passing interface) standard. Both the computation and memory requirements of the program can be distributed among multiple (possibly distributed) UNIX processes. Mr. Pedretti's talk will present approach and implementation of UIcluster, several real-world usage examples, and performance and accuracy measurements of the program.

CSRI POC: Doug Doerfler, 09224, 844-9528

Title: Combinatorial Algorithms for Adaptive Computation

Speaker: Ali Pinar
Department of Computer Science
University of Illinois at Urbana-Champaign

Date/Time: Thursday, April 26, 2001, 1:00 pm

Location: Building 980 Rm. 95

Abstract: Many scientific calculations require redistribution of their data and work during computation, either because the computation changes, or different components of the algorithm favor different partitions for better performance. Although effective algorithms and tools have been developed to find a new distribution of data, several problems remain open. First, data redistribution raises communication problems such as actually moving data to implement a redistribution subject to memory constraints and determining the new communication pattern after the new subdomains are assigned. Second, most partitioners rely on a priori information on the runtime of tasks, which is not always possible as in the case of overlapped Schwartz domain decomposition, ILU preconditioners, or complete factorizations. Finally, for a given data distribution some tasks can be assigned to one of many processors without altering the communication costs, and a clever assignment of these flexibly assignable tasks will improve performance. In this talk, I will briefly describe these problems and our solutions.

CSRI POC: Rich Lehoucq, 09214, 845-8929

Title: Memory-friendly Algorithms in Scientific Computing

Speaker: Alex Pothen
Computer Science, Old Dominion University, Norfolk VA
and ICASE, NASA Langley Research Center, Hampton VA

Date/Time: Monday, May 14, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: The web economy has lowered the costs of Gigabytes of memory and Terabytes of disk storage, enabling us to solve bigger problems than before. However, the gap between processor speeds and memory access speeds continues to grow every year, so that exploiting the larger memory and storage sizes requires the design of algorithms that can efficiently use these resources. We consider two problems: the design of sparse direct solvers efficient for external memory, and the design of kernels of iterative solvers efficient for caches. Using theory, simulations, and measurements of performance metrics via hardware counters, we show how to design memory-friendly algorithms in scientific computing.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: High Performance Computing

Speaker: Dr. Mahesh Rajan
Hewlett Packard Co., Palo Alto, CA

Date/Time: Monday, October 29, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Dr. Mahesh Rajan will speak on his work in High Performance Computing carried out at the Jet Propulsion Labs over the past 4 years. Dr. Rajan has spent many years in the high performance computing field and has practical experience on several platforms of interest to ASCI researchers. He will be available for interviews on Monday afternoon.

CSRI POC: John Noe, 9338, 844-5592

Title: A Multimodel Approach for Computational Simulation of Complex Systems

Speaker: Dr. John A. Reed
Mechanical, Industrial & Manufacturing Engineering Department
The University of Toledo

Date/Time: Wednesday, July 25, 2001, 11:00 am

Location: Building 980 Room 95

Abstract: In the development of large-scale, complex systems, the design process can be improved and development costs reduced by employing computational simulation tools to rapidly and inexpensively evaluate alternative designs. To be effective, computational simulation requires the use of higher-fidelity physics-based models early in the design cycle where they can be used to make trade-off decisions before a commitment to a final design is made. However, high-fidelity simulation of a complex system requires computing resources and simulation times which are beyond practical limits for use in industrial design processes. Furthermore, a high-fidelity model of the entire system may prove difficult to analyze and manipulate in a design environment. Accordingly, it is necessary to develop simulation techniques which permit the analyst or designer to combine models of varying fidelity based upon the particular physical processes being studied. The combined system models, or multimodel, enables engineers to analyze the system, focusing in detail on areas of interest, while at the same time optimizing use of computational resources and execution speed.

In this seminar, I will discuss recent efforts to develop several multimodels for aircraft engine simulation, concentrating on techniques used to incorporate higher-fidelity compressor and fan models with a low-fidelity system model. I will also describe the architecture of a general, object-oriented simulation environment being developed to support multimodeling.

CSRI POC: Gary Hennigan, 09233, 845-7558

Title: Hydrated ions and ion channel selectivity

Speaker: Susan Rempe
Los Alamos National Laboratory

Date/Time: Thursday, April 19, 2001, 1:00 pm

Location: Building 980 Rm. 95

Abstract: Biological membranes contain water-filled channels that selectively pass ions, thus controlling fundamental processes such as electrical signaling in the nervous system. Molecular-scale features of how channels select and transport ions remain largely unknown. Our studies focus on the interactions of various ions with water and small molecules that model the channel walls. Here I present surprising results for the water structure around several alkali metal ions and relate that to ion channel selectivity.

CSRI POC: Grant Heffelfinger, 09209, 845-7801

Title: Dynamic Interfaces and Production Software

Speaker: David Rogers, SNL

Date/Time: Tuesday, January 16, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Much research has been done into 'alternate' interface paradigms to help users view lots of data more intuitively. Solutions have ranged from pure 3D/VR worlds to interfaces with 'infinite zoom' such as Pad++. The discussion will also focus on the technology developed by dynaVu Inc. to solve this problem in a web browsing product. Then the discussion will be how these principles of dynamic interfaces were applied to a large simulation.

The production of an animated film requires practical solutions to complex problems. Several major tools will be addressed that I designed and lead teams to deliver. Challenges of integrating traditional animation including hand-drawn images and paintings with the digital artwork will also be reviewed.

CSRI POC: Philip Heermann, 09227, 844-1527

Title: Use of Probabilistic Ordinal Optimization for Continuous-Variable Optimization Under Uncertainty

Speaker: Vicente Romero, 9133
Chun-Hung Chen, George Mason University
Douglas Ayon, 9133 Summer Intern

Date/Time: Wednesday, June 27, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: A very general approach to solving optimization problems involving uncertainty is through the use of Ordinal Optimization (OO), where improvement at each step in the optimization problem is based only on a relative ranking of local alternatives, rather than on a crisp quantification of each alternative. This approach allows optimization under non-probabilistic and semi-quantitative descriptions of uncertainty. In fact, it can be seen as a formalism of the ordinal selection process that decision makers employ when deliberating over competing options whose outcomes are not estimable or projectable to first or higher orders -usually the case in the real world. A class of non-gradient-based optimizers that includes Coordinate Pattern Search (CPS) and Simplex methods can be used to perform Ordinal Optimization Under Uncertainty (OOUU). In cases where complete probabilistic information on the uncertain alternatives is available, other optimization approaches can also be employed, but probabilistic ordinal optimization provides a "Gold Standard" reference against which the accuracy and efficiency of other OUU methods can be compared (analogous to the role that Monte Carlo plays in uncertainty propagation). This talk will discuss a recent application of CPS-based probabilistic OOUU and planned efficiency enhancements to the method.

CSRI POC: Vicente Romero, 9133, 844-5890

Title: Numerical Modeling of Circulation in Lakes with Variable Topography

Speaker: David Ropp
University of Arizona

Date/Time: Thursday, February 15, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: This talk will describe the development of a model for long-time horizontal circulation in a shallow lake. Such a model can be useful in studying transport of nutrients or contaminants, or in navigation.

The development of the model has three parts. First, we derive the equations that capture the pertinent physics, and show that these equations are well-posed. Next, we construct a numerical scheme, using finite elements, to solve these equations, and show that this scheme is stable. Finally, we test the model on data from Lake Erie and compare with experimental results.

CSRI POC: Rich Lehoucq, 09214, 845-8929

Title: Performance Simulation

Speaker: Professor P. Sadayappan
Ohio State University

Date/Time: Wednesday, August 29, 2001, 2:00 pm

Location: Building 980 Room 95

Abstract: Abstract: Prof Saday(appan) will present initial results on impact of topology, routing, and mapping on performance of applications. He is a PI on a CSRI funded project with us. With numbers derived from CTH, this study reveals some interesting results.

CSRI POC: Neil Pundit, 09223, 845-7601, Glenn Laguna, 9224, 844-2308

Title: Simulating Polymer Blends with a Phenomenologically Derived Constitutive Model

Speaker: Richard Schiek
Technical University of Eindhoven, Eindhoven, the Netherlands

Date/Time: Thursday, March 29, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Pure polymers are often mixed or blended together to create materials with new properties. However, if the polymers are immiscible then the resulting blend will not be a continuous material. Rather, the blend will contain discrete domains, or droplets, of one polymer distributed within the other. The morphology of these droplets is determined by the physical flow or mixing history of the blend. To understand and predict the properties of a polymer blend, a model is developed based on the micro-structural phenomena encountered when two immiscible polymers share the same flow domain. This constitutive model incorporates the basic physical behavior of a drop under flow: drop deformation, breakup and coalescence. Predictions of this constitutive model are studied under planer, step shear and journal bearing flows.

CSRI POC: Laura Frink, 09209, 844-1910

Title: Fast 2D Electrical Impedance Tomography reconstruction for bubbly flow

Speaker: Professor William Schultz
University of Michigan
Department of Mechanical Engineering and Applied Mechanics

Date/Time: Tuesday, June 26, 2001 9:00 am

Location: Building 980 Rm. 95

Abstract: A singularity image method is applied to the electrical impedance tomography of gas-liquid flows in two-dimensional circular domain. Algorithms that use analytic complex functions, dipoles, and the Milne-Thomson circle theorem are described. Numerical experiments are provided to demonstrate the robustness of this technique. Numerical results show excellent reconstruction properties

Prof. Schultz is working with Sandia on reconstruction techniques for electrical impedance tomography, under a program funded by the NSF. Prof. Schultz is one of Paul Tortora's (9112) advisors.

CSRI POC: Tim O'Hern, 9112, 844-9061

Title: Optimization-Based Reference-Matrix Rezone Strategies for Arbitrary Lagrangian-Eulerian Methods

Speaker: Mikhail Shashkov
Los Alamos National Laboratory

Date/Time: Monday, January 15, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: The philosophy of the Arbitrary Lagrangian-Eulerian methodology (ALE), [1], for solution of multidimensional problems of fluid flow, is to move the computational grid (using the motion of the fluid as a guide) to improve the accuracy and efficiency of the simulation. One of the main elements is a rezone phase in which a "rezoned"grid is created that is adapted to the fluid motion. This paper develops a general rezone strategy that ensures the continuing geometric quality of the computational grid, while keeping "rezoned"grid at each time step as close as possible to the Lagrangian grid. Although the method can, in principle, be applied to more general grid types and in 3D, we restrict ourselves here to structured quadrilateral and unstructured triangular grids with fixed connectivity.

The unstructured grid is defined by positions of the nodes and connectivity. Between the node and each of its neighbors is an "edge-vector". Connectivity determines which edge-vectors form the columns of matrices used in the formulation of the objective function used in the optimization. The rezoning procedure consists of two phases consisting of a sequence of local optimizations followed by a single global optimization. Local optimization is performed in order to construct "reference"matrices to be used in the global optimization. For each given node we form a local patch from Lagrangian grid and construct local analog of condition number functional [2]. Minimization of this functional with respect to position of central node defines its 'virtual' location (the node is not actually moved). By connecting this virtually-moved node with its neighbors we define "reference" edge-vectors and matrices which represent the best locally achievable geometric grid quality.

The "Rezoned" grid is the grid which minimizes the global objective function. The global objective function measures the distance, in least-squares sense, between the matrices of the "rezoned"grid and the "reference" matrices. This objective function includes a "barrier" which penalizes grids with cells close to inverted. Direct optimization minimizes the global objective function. Numerical examples are provided to demonstrate the robustness and the effectiveness of our new method on model examples as well as for ALE calculations of hydrodynamics problems.

1. L. Margolin, Introduction to "Arbitrary Lagrangian-Eulerian Computing Method for All Flow Speeds", J.Comp.Phys., 135 (1997), pp.198-202.
2. P.Knupp, Matrix Norms and the Condition Number -A General Framework to Improve Mesh Quality via Node-Movement, Proc. of the 8th Meshing Roundtable, S. Lake Tahoe CA, 1999, pp.13-22.

CSRI POC: Dan Carroll, 09231, 845-8069

Title: vxWorks

Speaker: Richard Shinn
WindRiver Systems

Date/Time: Monday, July 23, 2001, 1:30 pm

Location: Building 980 Room 95

Abstract: Richard Shinn will talk briefly about the initial stages of the vxWorks boot sequence. The scope of his talk will cover device initialization, initial C environment setup, rom decompression and loading of the operating system. Mention will be made of the various alternative approaches which may be used to boot vxWorks.

CSRI POC: Jim Otto, 9223, 844-2185

Title: Modelling Self adaptation of mutation rates within Genetic Algorithms

Speaker: Dr. Jim Smith, Faculty of Computer Science and Mathematics
University of the West of England

Date/Time: Thursday, July 5, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: The use of "Self Adaptation" as a means of learning suitable mutation rates to use with Genetic Algorithms has been shown to show significant performance benefits I will briefly review the principles of Self-adaptation, and outline some research results.

I will then go onto to outline some approaches that have been taken to modelling GA behaviour, before introducing a dynamical systems model of a simple self-adaptive system. This model is used to predict the mean fitness of an evolving population as a function of time. The accuracy of these predictions are then tested by running a series of experiments using Genetic Algorithms with different population sizes. It will be shown that although there is a threshold below which the ``real'' populations do not closely follow the predictions, the model is still able to give us useful information about the behaviour of the ``real'' GAs, since the populations tend to get ``stuck'' at points close to certain eigenvectors of the infinite population model. Arguments are given which allow the prediction of which eigenvectors will be important.

The dynamics of the population evolving on a non-stationary environment are then considered, and some conclusions drawn about the nature of environmental change to which the algorithm will be able to respond.

CSRI POC: Bob Carr, 9211, 845-8562 and Bill Hart, 9211, 845-2217

Title: Optimization in Medical Diagnosis

Speaker: Ariela Sofer
Department of Systems Engineering & Operations Research
George Mason University

Date/Time: Tuesday, June 5, 2001, 10:00 am (PST)

Location: Building 940 Auditorium (Sandia – CA)
Building 980 Rm. 24 (Sandia – NM)

Abstract: In the past few years optimization has emerged as a promising tool in solving problems in medical diagnosis and treatment. This talk discusses the application of optimization to two problems in medical diagnosis: image reconstruction in positron emission tomography (PET), and biopsy protocols for diagnosis of prostate cancer.

(PET) is a medical imaging technique used for studying blood flow and metabolism of an organ. PET is used in studying brain function, and in diagnosing Alzheimers's disease and certain forms of cancer. Maximum likelihood reconstructions of PET images often have improved resolution over nonstatistical reconstructions, yet they are generally considered to be too expensive. This talk investigates the solution of 3-D maximum likelihood reconstruction problems via a primal-dual interior point method. The algorithm uses an early-terminated conjugate gradient to obtain an approximate Newton direction. We discuss a variety of computational issues such as data structures, preconditioning techniques, scaling of variables and the line search, and present computational results on problems involving over a million variables. This research is joint with Calvin Johnson, NIH.

Prostate needle biopsy is used for the detection of prostate cancer. The protocol of needle biopsy most commonly used in the clinical environment is an empirical technique which defines six symmetric locations on the prostate surface for needle insertion. Recent studies, however, have shown that a significant number of patients are not diagnosed at their first biopsy. We develop an alternative optimized biopsy protocol based on a 3-D distribution map of prostate cancer. This research is joint with Jianchao Zeng and Brett Opell, Georgetown University Medical Center.

CSRI POC: Monica Martinez-Canales, 08950, (925) 294-3157

Title: Preconditioning Strategies for Ill-conditioned Large Linear Systems.

Speaker: Masha Sosonkina
Assistant Professor, Department of Computer Science
University of Minnesota, Duluth

Date/Time: Wednesday, July 11, 200l, 9:00 am

Location: Building 980 Room 95

Abstract: In this talk, we describe how two preconditioning techniques, incomplete LU factorization with threshold and a multilevel recursive solver, have been adapted for linear systems arising in tire equilibrium computation. The heterogeneous material properties, nonlinear constraints, and a three dimensional finite element formulation make these systems difficult to solve by iterative methods. A combination of such techniques as shifting, filtering of small entries, pivoting in preconditioning, and a special way of defining levels for the multilevel recursive solver are shown to make these preconditioning strategies suitable for problems in tire design.

For difficult large-scale linear systems, a parallel version of a multilevel recursive preconditioning is used. We present an incremental approach for developing a distributed-memory preconditioner based on the Algebraic Recursive Multilevel Solver (ARMS) framework. The parallel overhead is localized only in the treatment of the local interface variables, i.e., the variables coupled with local and non-local equations. Thus the overhead amount can be easily adjusted depending on the problem difficulty and on the architectural characteristics. A few numerical experiments will illustrate the design choices.

CSRI POC: Mike Heroux, 9214, (320) 845-7695

Title: Rayleigh-Benard-Marangoni convection with a deformable interface

Speaker: Professor Alastair Spence
University of Bath, UK

Date/Time: Wednesday, June 6, 2001, 1:00 pm

Location: Building 980 Room 95

Abstract: In this talk some basic concepts of bifurcation theory will be introduced using ideas from singularity theory. Four bifurcations likely to arise in one parameter problems, namely fold points, Hopf points and symmetry-breaking bifurcations will be reviewed and their stability under discretisation discussed. Transcritical bifurcations will also be considered. Convergence and superconvergence results will be presented and illustrated with numerical examples.

CSRI POC: Andy Salinger, 9233, 845-3523

Title: Recent Advancements in the Efficiency of Global Atmospheric Models

Speaker: Bill Spotz
National Center for Atmospheric Research
Boulder, CO

Date/Time: Monday, June 4, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Numerical weather prediction and climate models are examples of scientific computing applications which affect both everyday life and national policy. Some computational fluid dynamicists are surprised to learn that spectral models are prevalent in the atmospheric science community. I will provide an introduction to global atmospheric modeling explaining why this is the case, how other methods compete against the spectral method, how distributed memory parallelism is implemented, and what drawbacks are present in the spectral model. I will then proceed to explain the theory and implementation of my own research, which is an alternative spectral representation stabilized with a projection. I will focus on improved floating-point operation counts and cache reutilization as well as load balancing issues. I will finish with speculation about the future of atmospheric modeling in a massively parallel world.

CSRI POC: John N. Shadid, 09233, 845-7876

Title: Assembling Mammalian DNA

Speaker: Ken Stanley
500 Software

Date/Time: Monday, January 15, 2001, 2:30 pm

Location: Building 980 Rm. 95

Abstract: Mammalian DNAs consist of approximately 3 * 10^9 bases (A, C, T or G). Efforts to decode mammalian DNA center around technology that allows selection of almost random contiguous pieces ranging in size from 2,000 to 300,000 out of the full DNA. Roughly 500 bases can be read from both ends of the smaller pieces. In addition, smaller subpieces ranging in size from 2,000 to 10,000 can be randomly selected out of the larger pieces, and again roughly 500 bases can be read from each end.

Were ours a perfectly random world, and one in which cost were not a factor, assembling mammalian DNA would be a solved problem. However, the DNA sequence is far from random and includes long stretches of exact and near exact copies. The selection of pieces out of the whole is not random, nor are the errors. The cost, presently estimated at $100 million is not insignificant.

I will explain the solution to the perfectly random problem and a couple approaches to the real problem.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Domain-based partitioning for parallel SAMR applications

Speaker: Johan Steensland
Uppsala University, Sweden

Date/Time: Monday, August 13, 2001, 2:00 pm

Location: Building 980 Room 95

Abstract: Domain-based partitioning techniques for dynamic structured grid hierarchies, occurring in structured adaptive mesh refinement (SAMR) methods, will be presented. Such methods for obtaining the numerical solution to partial differential equations yield highly advantageous ratios for cost/accuracy as compared to methods based upon static uniform approximations. Distributed implementations of these techniques have the potential for enabling realistic simulations of complex systems. These implementations however, present significant challenges in dynamic data-distribution and load balancing. That is, when SAMR is executed on a parallel computer, the work load will change dynamically. Thus, there is need for _dynamic_ load balancing in order to efficiently utilize the resources of the parallel computer. Inverse space-filling curve partitioning (ISP) is appealing for load balancing in parallel SAMR, because of its speed. In this work, ISP is considered as _part_ of a partitioning approach, which combines structured and unstructured techniques.

CSRI POC: Karen Devine, 9226, 845-7585

Title: Validation Warrants are Inverse Proofs

Speaker: Professor D. E. (Steve) Stevenson, Computer Science Department
Clemson University

Date/Time: Thursday**,** April 5, 2001, 10:00 am

Location: Building 980 Rm. 95

Abstract: I will present work in progress which proposes a foundations of validation based on a particular (and I believe) new view of probability. Bypassing the Bayes/Classical controversy, I start with the ideas of Jeffreys, Good, Polya, and E. T. Jaynes. Jaynes approached probability and statistics in science by what he termed "probability theory as logic." By this he meant that only probability models built from first principles should be admitted. I go one step further by observing that probability laws can be derived by inverting the rules of inference in logic in a sense to be made specific in the talk. We present the current formulation and give some examples of how such a theory can be used to validate a model.

CSRI POC: Timothy G. Trucano, 09211, 844-8812

Title: The Numerical Analysis of Bifurcations

Speaker: Professor Simon Tavener
Colorado State University

Date/Time: Wednesday, June 6, 2001, 2:00 pm

Location: Building 980 Room 95

Abstract: The convective motions that arise when a fluid layer is heated from below were most famously studied by Benard and Rayleigh in the early part of this century and have attracted numerous theoretical, experimental and numerical investigations in the intervening years. In an open container, fluid convection is not driven by buoyancy forces alone, but also by surface forces that arise from the temperature-dependent nature of the surface tension. Indeed, thermocapillary forces dominate buoyancy effects in sufficiently shallow layers. Computations of Marangoni-Benard convection are usually performed in two- or three-dimensional domains with rigid boundaries, and a contact angle of 90 degrees between the fluid and its container is generally assumed. Qualitatively different behavior can occur if the free surface is allowed to deform. Bifurcations arising in a two-dimensional domain with a deformable free surface can be computed using the finite-element method, by combining an orthogonal mapping of the physical domain with extended system techniques for locating singularities. I will show that even small deformations of the free surface result in surprisingly large changes in the critical Marangoni numbers at which convection occurs. Motivated by the Liquid Encapsulated Vertical Bridgeman (LEVB) crystal growth technique, I will then consider the two-fluid problem, concentrating on instabilities that arise when heating from above.

CSRI POC: Rich Lehoucq, 9214, 845-8929

Title: Coarse Bifurcation Studies of Bubbly Flows and Heterogeneous Systems

Speaker: Konstantinos Theodoropoulos
Department of Chemical Engineering
Princeton University

Date/Time: Tuesday, March 13, 2001, 10:30 am

Location: Building 980 Rm. 95

Abstract: Realistic models of physico-chemical processes are often microscopic (e.g. Monte Carlo, Molecular Dynamics, Lattice Boltzmann) or in the form of macroscopic PDEs. The objective of this work is to expand the capabilities of microscopic and of PDE-based timestepping codes by enabling them to perform tasks such as stability analysis, continuation and bifurcation computations.

The proposed methodology exploits a time-stepper-based computational approach introduced by Shroff & Keller [1] called the Recursive Projection Method (RPM). Consecutive calls to an existing time-stepping routine, that identifies the right-hand side of the problem, constitute the backbone of the method. RPM treats the timestepper, as a black box and adaptively identifies the low-dimensional invariant subspace P along which the convergence is the slowest. Integrations towards the steady state (Picard iterations) are substituted by Newton iterations in P using a small approximate numerical Jacobian and a Picard iteration in its orthogonal complement Q.

The approach is illustrated through the bifurcation/stability analysis of two systems:(i) The flow of individual bubbles in a periodic box simulated with a microscopic Lattice-Boltzmann (LB) timestepper. Such calculations are, currently, not accessible to direct microscopic codes. Coarse, i.e. averaged, mean-field steady states of the system correspond to stationary profiles of population moments in microscopic simulations.

The parameter-dependent behavior of a system is traditionally analyzed by first obtaining a coarse model (e.g. a mean field PDE) and then performing bifurcation/stability analysis of the PDE. Our methodology (exploiting RPM) allows the implementation of bifurcation studies using directly microscopic evolution rules without ever constructing the corresponding mean-field PDE [2]. Such an approach is essential when a macroscopic description of the system is unavailable in closed form. Coupling the LB simulator with RPM, has enabled us to perform systematic parametric studies. The coarse eigenspectrum of the system at different parameter values was calculated and a Hopf bifurcation was identified. Furthermore, the scheme was able to converge even on unstable steady states beyond the Hopf point.

(ii) A reaction-diffusion system, describing the oxidation of CO [3] on a periodically varying medium. This system, described by a set of time-dependent PDEs, exhibits travelling pulses with a small-scale modulation due to the variations in the medium. To perform bifurcation analysis on the coarse features of the system (i.e. without the small-scale variations) a homogenized equation should be derived taking in account the asymptotic behavior in the small-scale limit (such equations are, in principle, difficult to obtain). Bifurcation calculations should be then performed using this equation. Our approach allows for the direct coarse bifurcation analysis of the system without explicitly deriving the homogenized equations. Coarse Hopf bifurcations are identified and coarse spectra are computed.

CSRI POC: Andy Salinger, 09233, 845-3523

References Cited

[1] G. M. Shroff and H. B. Keller, SIAM J. Numer. Anal. 30, 1099 (1993).
[2] C. Theodoropoulos, Y.H. Qian and I.G. Kevrekidis, Proc.Nat. Acad. Sci. 97, 9840 (2000).
[3] M. Baer, N. Gottschalk, M. Eiswirth, and G. Ertl, J. Chem. Phys. 100, 1202 (1994).

Title: Pattern Search: A Status Report

Speaker: Virginia Torczon
Department of Computer Science
The College of William and Mary

Date/Time: Thursday, December 13, 2001 2:00 (PST)

Location: Bldg. 921, 137 (Sandia - CA)
Bldg. 980, Room 24 (Sandia - NM)

Abstract: Pattern search methods have seen active use at Sandia National Labs to solve engineering optimization problems. Pattern search methods are a class of methods for solving general nonlinear programming problems. Pattern search methods are represented in at least two Sandia software packages: APPSPACK and DAKOTA.

I will review current progress on two fronts: implementation and analysis. I will discuss what work remains ahead and what the analysis indicates will be significant challenges to the development of effective software.

CSRI POC: Tammy Kolda, 08950, (925) 294-4769

Title: MAGE-ML : A look at Microarray Gene Expression data standards

Speaker: Charles D. Troup
Bioscience Research Solutions
Agilent Technologies

Date/Time: Thursday, September 6, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: DNA Microarrays are becoming an increasingly popular tool for characterization of total cellular messenger RNA. Chip based technologies allow for massively parallel experiments which is exciting for bench and computation scientists alike. The volume of data gathered from any one experiment can be enormous not to mention that of multiple experiments. There is a real need to standardize on a data format for communicating data from experiments using microarrays. MAGE-ML (Microarray Gene Expression Markup Language) is an attempt to capture a large set of possible experiments as a standard representation of the data generated from these experiments. This project is a collaborative effort between several groups on how to standardize data interoperability of microarray data. These groups have attempted to define their own standards in the past. MAGE-ML covers a wide range of data that a scientist or vendor may wish to communicate included array annotation, array manufacturing, sample preparation, feature extraction analysis and clustering analysis. The process used to develop the model is also interesting as we employ the most modern techniques in software model development, such as MDA (Model Driven Architecture).

CSRI POC: Grant Heffelfinger, 9209, 845-7801

Title: A Parallel Direct Solver Package for CoarseGrid Problems[1,2]

Speaker: Dr. Henry M. Tufo III
University of Chicago
Argonne National Laboratory

Date/Time: Tuesday, January 30, 2001, 2:00 pm

Location: Building 980 Rm. 95

Abstract: We have recently developed a direct solver package for parallel solution of coarsegrid problems, Ax = b. Such problems arise when domain decomposition or multigrid methods are applied to elliptic partial differential equations in d space dimensions. The package is currently targeted for the case in which A is sparse, the number of processors, P, is large, the number of degrees of freedom per processor, n/P, is small, and the coarsegrid data and solution before and after the solve stage are distributed. The approach is based upon a (quasi) sparse factorization of the inverse of A. Assuming A is the matrix equivalent of a differential operator discretized with standard finite differences or finite elements having compact support, we show that the method requires O(n^gamma log_2(P)) time for communication and O((n^(1+gamma))/P) time for computation, where gamma=(d-1)/d, and d is the spatial dimension associated with the originating PDE. We review the basic theory and API, present recent results on 1000s of processors, and discuss future extensions.

[1] Department of Computer Science, University of Chicago, Chicago, IL 60637.
[2] Mathematics and Computer Science Division, Argonne National Laboratory,

Argonne, IL 60439.

CSRI POC: David Womble, 09214, 845-7475

Title: A New Dawn
The Advent of Learning Systems

Speaker: Darryl Turner
Turner Technology, Inc., System Services Group

Date/Time: Friday, May 4, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: Intelligence is defined as the ability to learn, reason and think. Consciousness is defined as the ability to maintain an abstract awareness of one's own existence and one's surroundings in real time. Each of these was at one time thought to be the exclusive province of mankind. We now know that various animals are capable of all of the above, though at a much lower level of abstraction. Indeed, it is our ability to function at a uniquely high level of abstraction that actually separates mankind from the other creatures. Now a machine has been developed that does all of the above on about the same level as a human being and subsequent software systems exist that inhabit the internet and, while they do not function on the same level of sophistication, do not require the support of a supercomputer and run on any reasonably strong PC.

Here, we shall review the development of the MAX I, the first artificial person, Project Prometheus, that made it all happen and smaller, simpler Learning Systems that perform useful tasks on the internet. We will briefly look as what it all means and where we are going. Then, there will be a period for questions.

CSRI POC: Jim Otto, 09223, 844-2185

Title: Uncertainty, Plans and Robustness

Speaker: Mark Turnquist
Professor of Civil & Environmental Engineering, Cornell University

Date/Time: Wednesday, July 25, 2001, 1:00 pm

Location: Building 980 Room 95

Abstract: An alternative title for this talk might be: "What to do when you don't know everything you need to know to make the best decision." All of us have experienced this problem as individuals, and organizations both inside and outside the Nuclear Weapons Complex face this problem every day. As part of an effort to advance surety analysis in various arenas of interest to SNL, the Advanced Planning Group in Org. 6515 has been focusing on models applicable to resource-constrained planning under uncertainty. A particular emphasis of this work is on optimization formulations that lead to "robust" decisions. The talk will describe the basic modeling approaches used, and applications to transportation of special nuclear materials, manufacturing of neutron generators, and general project planning.

CSRI POC: Cindy A. Phillips, 9211, 845-7296

Title: Fragmentation as a Strategy for High-Speed IP Networking

Speaker: Todd Underwood
University of New Mexico

Date/Time: Tuesday, May 15, 200l, 9:00 am

Location: Building 980 Rm. 95

Abstract: Commodity hardware, operating systems and network protocols are all important parts of the cluster-based future of supercomputing. This research uses fragmentation to improve the performance of the Internet Protocol (IP) running on gigabit Ethernet network interface cards (NICs). First, in order to determine if fragmentation was suitable for the role, research into the history and an investigation into the current uses of fragmentation was undertaken. The use of path MTU discovery to avoid fragmentation was found to be almost universal and the reordering and loss of fragments was found to be almost non-existent. Second, fragmentation and reassembly were implemented in Alteon Acenic 2 gigabit Ethernet cards. The driver was configured to use a maximum transmission unit (MTU) of 64KB, with the firmware on the NIC fragmenting the datagram to a frame size of either 1.5KB or 9KB. The results demonstrate the viability of the technique by significantly reducing processor utilization without substantially reducing bandwidth. Using 1.5KB frames, the fragmenting firmware was able to perform within 80-90% of the bandwidth of the unmodified firmware with processor utilization reduced by 45% (from about 80% to about 45%). Using 9KB frames the fragmenting implementation was able to achieve comparable bandwidths to the unmodified firmware (at 993Mb/s peak speeds) with a 25% reduction in processor utilization (from 75% to 55%).

CSRI POC: Rolf Riesen, 09223, 845-7363

Title: Nonlinear and linear solution methods

Speaker: Homer Walker
Worcester Polytechnic Institute

Date/Time: Monday-Friday**,** July 23-27, 2001, 9:00 am
Wednesday-Friday**,** August 15-17, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: The ASCI Algorithms Effort and the Computer Science Research Institute (CSRI) at Sandia is offering a series of lectures on numerical solution methods for nonlinear algebraic equations. These lectures will be presented by Homer Walker of Worcester Polytechnic Institute (WPI) who will be visiting the CSRI this summer.

Topic 1. Methods for Problems in One Variable
Topic 2. Newton's Method for Problems in Several Variables
Topic 3. Globally Convergent Modifications of Newton's Method
Topic 4. Quasi-Newton (secant update) Methods
Topic 5. Other Methods (continuation, homotopy, nonlinear CG, fixed point, etc.)
Topic 6. Inexact Newton and Newton-Krylov Methods.
- a. Newton-iterative and inexact Newton methods.
  - i. Formulation and local convergence.
  - ii. Globally convergent methods.
  - iii. Forcing terms.
- b. Krylov subspace methods.
- c. Newton-Krylov methods.
  - i. General considerations.
  - ii. Matrix-free implementations.
  - iii. Adaptation to path following.

Topic 7. Other methods for large-scale problems.
- a. Nonlinear conjugate gradient methods.
- b. Pseudo-transient continuation.

CSRI POC: Roger Pawlowski, 9233, 284-3740; John Shadid, 9233, 845-7876

Title: A MAUI Interface for Linear Solver Packages (Preliminary Report)

Speaker: James Willenbring
St. John's University

Date/Time: Wednesday, October 24, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: There are multiple linear solver packages that are widely available, each of which is individually sufficient for solving many problems. However, numerous benefits can be obtained using a common graphical user interface (GUI) to drive a combination of these packages. For example, many of these packages provide a user interface that requires detailed knowledge of the solver parameters, including which combinations of parameters are valid. Also, mixing and matching of components from different packages, and running many experiments over a range of parameter values is difficult to do without writing detailed, error-prone driver software.

In this presentation we give a preliminary report on efforts to develop a GUI based on MAUI and TSF. MAUI is a package that facilitates the development of a GUI via XML files. TSF is an abstract class hierarchy that can be used to call many general purpose solver libraries. Using this GUI, it is possible to define solver parameters using intuitive point and click techniques. Further, multiple parameter sets for a problem can be specified. Eventually we will be able to mix and match components from separate solver packages to solve a problem.

CSRI POC: Mike Heroux, 09214, (320) 845-7695

Title: A General-Purpose Model for Heterogeneous Computation

Speaker: Tiffani Williams
University of Central Florida

Date/Time: Friday, July 13, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Heterogeneous parallel environments are becoming an increasingly prevalent and important platform for high-performance computing. However, efficient application performance in these environments is often difficult to achieve due to the heterogeneity of the underlying resources. Furthermore, the lack of computational models and programming environments hinders the design and development of heterogeneous parallel applications. In this talk, we will describe the k-Heterogeneous Bulk Synchronous Parallel model, a computational model that encourages the design of software for heterogeneous platforms. The utility of the model is demonstrated through the design and analysis of several collective communication algorithms. Improved performance is not a result of programmers having to account for myriad differences in heterogeneous machines. By hiding the non-uniformity of the underlying system from the application developer, we offer a framework that encourages the design of heterogeneous parallel software.

CSRI POC: Rolf Riesen, 9223, 845-7363

Title: Data Storage Issues

Speaker: John Wilson
University of Colorado at Boulder

Date/Time: Tuesday, May 22, 2001, 1:00 pm

Location: Building 980 Rm. 95

Abstract: Data storage issues pose a potential problem in the performance of turbulence simulations. One tool tried in other domains to address such performance problems is data compression. Unfortunately, data from turbulence simulations does not usefully compress with lossless methods. This has led to the examination of wavelet-based lossy coding techniques for compressing turbulence data. One of the key issues is that a significant portion of the work on lossy coding involves image coding. The error measures used in image coding are not meaningful in the physical context of turbulence data. For this work with turbulence data, error is measured relative to standard deviation and for significant values also as relative error. Depending on the quantities of interest and the evaluation criteria, it is found that compression ratios of 4:1 to 64:1 are achievable in the wavelet-based lossy coding of turbulence data. If only global quantities are of interest, far greater levels of compression are possible.

Studying wavelet-packet-based coding found that finding a best-basis is sufficiently complex that in performance critical applications, wavelet-packet-based coding is not beneficial. The experiments with wavelet-packets and experiments varying the bit allocation used to achieve a given compression ratio have demonstrated the need for cost functions that better match the relevant error in quantities of interest.

In addition to varying the bit allocation among subbands, other experiments altering the coding process were conducted. Some, such as using non-uniform quantization, raise issues for further study, while others are at best of limited usefulness.

CSRI POC: Philip Heermann, 09227, 844-1527

Title: Transform Coding, Wavelet Transforms, and the JPEG2000 Standard

Speaker: Dr. Brendt Wohlberg
Center for Non-Linear Studies, Los Alamos

Date/Time: Wednesday, February 7, 2001 2:00 pm

Location: Building 980 Rm. 95

Abstract: JPEG2000 is a draft image coding standard that will greatly extend the existing JPEG standard. Based on wavelet transforms with sophisticated entropy coding mechanisms, it offers a number of very significant advantages over its predecessor. The JPEG2000 team at LANL is working on a number of different aspects of the standard, including support for scientific data, multi-component imagery, aspects of entropy coding, and extensions to the types of supported wavelet transforms.

The presentation will consist of a tutorial-level overview of the JPEG and JPEG2000 standards, followed by a more in-depth description of the wavelet transform mechanisms in JPEG2000, and will conclude with a brief discussion of the relevant research at LANL.

CSRI POC: Mark D. (Danny) Rintoul, 09235, 844-9592

Title: Robust Bayesian Estimation via Tilting and Data Augmentation

Speaker: William C. Wojciechowski
Rice University

Date/Time: Thursday, January 18, 2001, 9:00 am

Location: Building 980 Rm. 95

Abstract: When conducting a statistical analysis, the practitioner often assumes an idealized model and proceeds as if the conditions under which the model is valid are true. The conditions define the theoretical characteristics of the analysis and determine the interpretation of the numerical summaries produced during the analysis. However, if the conditions are only approximately true, the results can be inaccurate. In actual practice, the assumptions are only an approximation to the true situation. Therefore, methods that are reliable even when the model deviates considerably from the true situation are needed. Because robust statistics are designed to have this type of stability, they are a solution to this problem. We introduce a new general purpose robust estimator. This Bayesian procedure applies Gibbs sampling and data augmentation to achieve robustness by weighting the observations in the likelihood of Bayes' theorem. Because this method does not need to solve a high-dimensional optimization problem, it is particularly useful when there are many parameters. This robust method is computationally intensive, but its implementation is not complex, removing the need for an approximate calculation. Furthermore, there are other robust Bayes methods that achieve robustness by introducing many additional parameters in the model. In this case, there are usually as many or more parameters as observations. In contrast, this new robust Bayes method introduces only one additional parameter. The derivation of the new method will be presented along with examples that use simulated and real datasets.

CSRI POC: George Davidson, 09212, 844-7902

Title: Shortest-Path Network Interdiction

Speaker: Kevin Wood
Operations Research Department
Naval Postgraduate School

Date/Time: Monday, July 2, 2001, 11:00 am

Location: Building 980 Room 95

Abstract: Suppose that an adversary is attempting to move troops or materiel through a road network as quickly as possible from a rear staging area to a battlefront. And, we wish to attack the road network using limited resources in order to maximize the adversary's travel time. We model this problem as a bi-level (max-min) program, convert it to a mixed-integer program and solve it directly or through Benders decomposition. Benders decomposition is enhanced with a new generalization of integer cutting planes called "super-valid inequalities." Additionally, a new set-covering decomposition solves problems when numerical data makes direct solution or standard decomposition unwieldy. The techniques extend to minimizing the maximum output of a general system modeled as a linear or integer program, or even another interdiction model. In particular, "interdicting the interdictor" models the hardening of a system to minimize the effects of interdiction by an adversary.

Kevin Wood is Professor and Associate Chair for Research in the Operations Research Department at the Naval Postgraduate School. He received his PhD in OR from UC Berkeley in 1982 and has published in the areas of network reliability, applications of mathematical programming, integer and stochastic programming, and network interdiction.

CSRI POC: Cindy Phillips, 9211, 845-7296

Title: Implicit Solution of Radiation Diffusion Problems

Speaker: Carol S. Woodward
Lawrence Livermore National Laboratory

Date/Time: Monday, December 10, 2001, 9:00 am

Location: Building 980 Room 95

Abstract: Modeling radiation diffusion processes has traditionally been accomplished through inaccurate and nonscalable solution methods based on decoupling linearized equations. We present an algorithm for the fully implicit solution of radiation diffusion with material energy coupling and discuss numerical issues addressed in its implementation. Our work uses a stiff ODE solver coupled with Newton's method for solving the implicit equations arising at each time step. The Jacobian systems are solved by applying GMRES preconditioned with multigrid. By combining the nonlinear Newton iteration with a multigrid preconditioner, we hope to take advantage of the fast, robust nonlinear convergence of Newton's method and the scalability of the linear multigrid method. Numerical results will be shown comparing our fully implicit method with a semi-implicit technique.

CSRI POC: John N. Shadid, 9223, 845-7876

Title: Data Visualization with VxInsight

Speaker: Brian Wylie and George Davidson, SNL

Date/Time: Monday, January 15, 2001, 12:00 noon

Location: Building 980 Rm. 95

Abstract: The information age allows people to find, extract, and gather information in astronomical proportions. Commercial databases, the World Wide Web, and data warehouses offer the promise of knowledge and insight if one can harness the overwhelming volume of raw information. Currently most tools fall into the 'sorted list syndrome'; all data, no matter how it is processed, eventually gets presented as a sorted list. With a simple list presentation the understanding of complex relationships and broad trends is exceedingly difficult. Our software presents a visual interface that allows users to explore large data collections. The human visual system is extremely adept as spotting trends and patterns, and detecting relationships. The software exploits this capability by presenting the data within an intuitive landscape metaphor.

The presentation will cover the visualization software and current clustering algorithms. Demonstrations will be given on both literature and genomic data.

CSRI POC: Bruce Hendrickson, 09226, 845-7599

Title: Parallel Three-Dimensional Chebyshev Pseudospectral Modeling of Cardiac Tissue

Speaker: Rongqi Yan
Klipsch School of Electrical and Computer Engineering
New Mexico State University

Date/Time: Monday, May 21, 2001, 10:00 am

Location: Building 980 Room 95

Abstract: Monodomain or bidomain modeling has been used widely to study various bioelectric phenomena in cardiac tissues. In most of these previous studies, the finite difference method has been used to solve the governing partial differential equations. An adequate discretization of a tissue with realistic dimensions, however, still often leads to a large model size that is computationally prohibitive. Recently, studies for a two-dimensional monodomain have demonstrated that the Chebyshev pseudospectral method is computationally much more efficient than the finite difference technique. For my research, extension of the pseudospectral method to the three-dimensional case is described. The transmembrane potential is expanded in terms of Chebyshev polynomial trial functions and the monodomain equation is enforced at the Gauss-Lobatto grid points. Spatial derivatives are obtained with fast Fourier transforms and the solution is advanced in time using an explicit technique. To increase the speed of computation, parallelization of the algorithm is being studied, An one-dimensional domain decomposition is used to distribute the computational load among the processors. Initial results on the parallel performance are promising. Details on the parallel three-dimensional pseudospectral technique will be presented together with an analysis of its computational performance.

CSRI POC: Scott A. Hutchinson, 9233, 845-7996

Distribution:

External:

Dr. William Reed, DP-114
U.S. Department of Energy, Defense Programs
1000 Independence Avenue SW
Washington, DC 20585

Dr. Jose Munoz, DP-114
U.S. Department of Energy, Defense Programs
1000 Independence Avenue SW
Washington, DC 20585

Internal:

1	MS 0139	Michael O. Vahle, 09900
1	MS 0139	Robert K. Thomas, 09904
5	MS 1110	David 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