COMPUTER AND COMPUTATIONAL SCIENCES

Director's R&D Fund

Synthesis of High-Performance Algorithms for Electronic and Nuclear Structure Calculations

D. E. Bernholdt,[1] R. J. Harrison,[1] J. B. White III,[1] D. J. Dean,[2] M. R. Strayer,[2] G. Baumgartner,[3] S. Hirata,[4] M. Nooijen,[5] R. M. Pitzer,[6] J. Ramanujam,[7] and P. Sadayappan[3]

[1]*Computer Science and Mathematics Division*
[2]*Physics Division*
[3]*Department. of Computer and Information Sciences, Ohio State University*
[4]*Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory*
[5]*Department of Chemistry, University of Waterloo*
[6]*Department of Chemistry, Ohio State University*
[7]*Department of Computer Science, Louisiana State University*

In order to increase both the productivity of scientific programmers, and the performance of the software they create, we are developing tools and techniques to translate a high-level expression of a computational problem into highly optimized source code, which can then be compiled and run as usual. This approach allows significant optimization of the algebraic expressions, trade-offs between storage space and computation time, and distribution of the data across the parallel computer during code generation. We are applying these techniques to produce a *Tensor Contraction Engine* targeted at a class of problems in computational quantum chemistry and computational nuclear physics. To be able to validate generated code for the nuclear coupled cluster problem, we have also done a hand implementation of such a code for comparison purposes.

Technical Approach

The development of high-performance parallel programs for scientific applications is usually very complicated and is exacerbated by the increasing complexity of modern computer systems. High-level or domain specific languages (HLLs) are one approach to addressing some of these issues. They allow scientists to express their computational problems in a form that is close to how they are derived and might be expressed in scientific publications. They can also be used as inputs for tools that process the high-level problem expression to produce actual computer code. This clearly addresses the "human scalability" problem in modern scientific computing; it also provides an excellent opportunity to address performance issues.

In traditional software development, the implementer makes many decisions, often quite empirically, affecting performance and scalability as they manually translate from the high-level expression of the problem into code. The use of tools to generate the actual code also provides the opportunity to thoroughly and automatically analyze the various possible implementations, taking into account details of the target hardware platform, and even the specific problem. The result is highly optimized and tuned algorithms to solve the problem originally specified in the high-level language.

In this project, we have developed tools and techniques to apply this approach to a class of computations that are both CPU and memory intensive, spanning several scientific domains. Our primary focus has been a class of methods used in quantum chemistry and nuclear physics known as coupled cluster methods. These methods are among the most accurate in routine use in the quantum chemistry,[1,2] while in the area of nuclear physics they are not widely used. They are formulated as contractions of large tensors (the multidimensional analog of matrix multiplication), with a given method involving tens to hundreds of different tensor contraction terms, giving rise to the extreme complexity of both implementation and performance optimization of such codes.

Our approach is to treat the problem from a computer science viewpoint, creating a specialized optimizing compiler, which we call the *Tensor Contraction Engine* (TCE),[3] which takes as input a high-level domain-specific language for describing the tensor contraction equations of coupled cluster theory, and generates code in a traditional programming language (in this case Fortran) which can then be integrated with the rest of the chemistry computational infrastructure for execution. The TCE must make the same design decisions a human would make for a hand implementation, but it has the advantage of being able to more rigorously and exhaustively explore the

options and take into consideration details of the target computer architecture and performance model, as well as details of the particular target problem, if appropriate. Modules of the TCE, depicted schematically in Fig. 1, include the following:

- Algebraic transformations, which transform the input tensor contractions into a mathematically equivalent form that minimizes the number of operations required to evaluate them.
- Memory minimization, which reduces the amount of memory required to store intermediate results during evaluation of the input.
- Space-time transformation, which makes trade-offs between recomputing quantities and storing them on disk.[4]
- Data locality management, which optimizes utilization of storage hierarchies, such as cache/memory or memory/disk.[5,6]
- Data distribution and partitioning determines the best layout of the data for efficient parallel processing.[7,8]

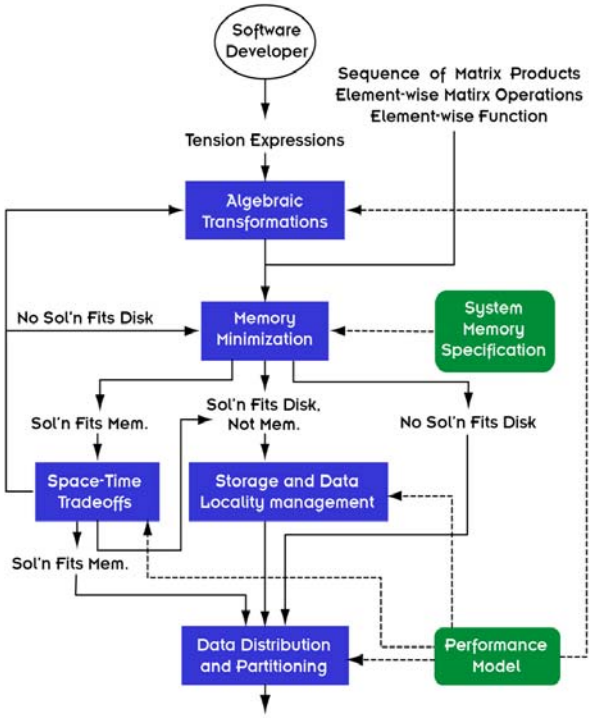

Fig. 1. The general architecture of the Tensor Contraction Engine (TCE). The user expresses their problem in a specially designed high-level language, and it is processed by the TCE to produce Fortran (in this case) source code implementing the input expression. TCE optimizations include algebraic transformations to minimize the overall operation count, memory minimization to reduce storage for intermediates, space-time trade-offs between storage and recomputation of intermediate quantities, data locality management, and parallel data partitioning.

It is worthwhile to note that most of these modules can make use of performance model information for the target computer system to determine the most effective optimizations.

Two versions of the TCE have been implemented. The "prototype" TCE is used primarily for fast extension of the automatic code generation capability and to examine issues around supporting various new types of coupled cluster calculations of interest to chemists. It implements few of the optimizations mentioned above, in the interest of simplicity and extensibility. The "optimizing" TCE is the "production" tool and includes all of the optimizations described above. The framework of the optimizing TCE is designed to facilitate generalization to other scientific domains, where similar needs exist, but the details of methodology and applicable optimizations will differ.

One of the special challenges regarding the use of automatically generated software for a complex problem is verifying the correctness of the code. In the quantum chemistry area, there are a number of existing handwritten implementations of various coupled cluster methods, so validation of the code generation process is relatively straightforward. In nuclear physics, however, coupled cluster methods are not in routine use, so that in order to be able to verify the correctness of the generated code, it was necessary to make a careful hand implementation to compare against. This capability has opened up some additional opportunities for research in the nuclear physics area, which are described below.

Another challenge for automatically generated software is its integration into existing software environments. In its current form, the TCE generates only the code to actually perform the tensor contraction operations, which must then be integrated into an existing chemistry package to provide the convergence checking and iterative control for the method itself, as well as providing various inputs required by the coupled cluster method. While it is possible to extend the high-level domain-specific language to include the necessary flow control and other capabilities required of a complete programming language, this approach leaves the developers of the tool with a task of maintaining an entire programming language rather than a much smaller, specialized language. An alternative solution involves embedding the HLL inside a traditional programming or scripting language, taking advantage of the existing infrastructure for the host language while providing the extra domain-specific capabilities of interest to the user. In anticipation of the need to provide a more complete language environment for the TCE, we undertook a small case study of the embedding of an HLL for expressing multigrid problems[9] in the functional programming language Haskell.[10] The results of the study are described below.

Results

The TCE, though still under very active development by our outside collaborators, has already proven very capable and effective at raising the productivity of software developers. The prototype TCE has been used to produce implementations of more than 20 different coupled cluster-based methods, including energies, lambda equations, dipole moments, MBPT methods through fourth order, CC and CI methods through CCSDTQ, EOM-CCSD, local/AO-based CCSD, and relativistic methods. These capabilities have been included in the latest release of the NWChem computational chemistry package.[11] Figure 2 demonstrates that reasonably efficient parallel code for coupled cluster methods can be generated automatically by the prototype TCE, in this case achieving a speedup of 9 (compared with ideal speedup of 16) between 16 and 256 processors.

The development of the optimizing TCE has been extremely challenging, both from the computer science standpoint and due to the complexity of the domain-specific (chemistry) details, but we have produced a very capable system in a flexible and extensible framework, which we plan to use as the basis for work on HLLs in other domains. From the computer science viewpoint, the optimization algorithms have virtually all been developed for this project, because they go beyond traditional compiler optimizations, resulting in a significant number of papers. We are now working on validating correctness and performance of the optimizing TCE on the simpler CCD and CCSD methods before broadening out. However Fig. 3 shows promise of the approach. This graph shows the effect of just the space-time trade-off optimization for one particular term from a Laplace-factorized CCSD(T) method:

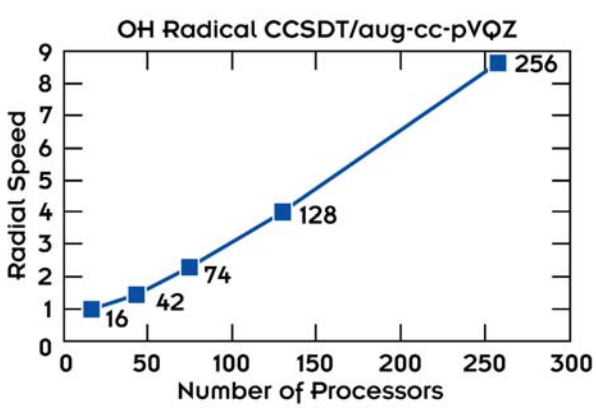

Fig. 2. Parallel scalability of code automatically generated by the prototype Tensor Contraction Engine for a CCSDT calculation on the OH radical using the aug-cc-pVQZ basis. Ideal scaling in this case would be 16 at 256 processors. These results were obtained on the HP Supercluster at PNNL's Molecular Science Computing Facility by So Hirata.

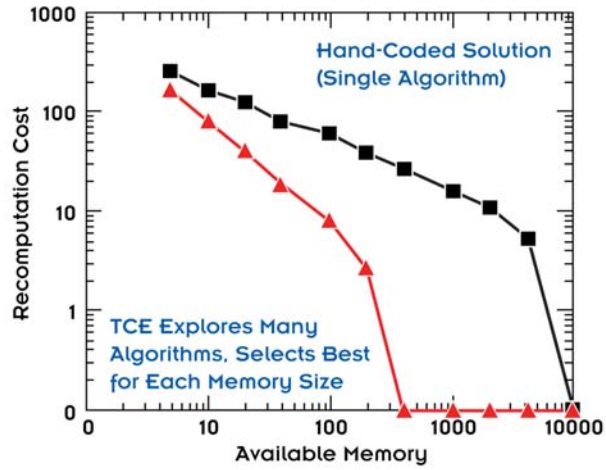

Fig. 3. Performance of the space-time trade-off optimization module compared to a typical hand-coded implementation of this term from the Laplace factorized triples formulation of the CCSD(T) method. The upper curve shows the amount of recomputation required as a function of available memory for a typical hand-coded implementation of the term. The lower curve shows results of the space-time optimization.

$$A3A = \tfrac{1}{2}(X_{ce,af}Y_{ae,cf} + X_{c\bar{e},a\bar{f}}Y_{a\bar{e},c\bar{f}} + X_{c\bar{e},\bar{a}f}Y_{\bar{a}\bar{e},cf}$$
$$+ X_{\bar{c}e,a\bar{f}}Y_{ae,\bar{c}\bar{f}} + X_{\bar{c}e,\bar{a}f}Y_{\bar{a}e,\bar{c}f} + X_{\bar{c}\bar{e},\bar{a}\bar{f}}Y_{\bar{a}\bar{e},\bar{c}\bar{f}})$$
$$X_{ce,af} = t_{ij}^{ce}t_{ij}^{af} \qquad Y_{ae,cf} = \langle ab\|ek\rangle\langle cb\|fk\rangle$$

The upper curve depicts the amount of recomputation required as a function of the system's memory for a typical implementation of this term. The lower curve is the result from the TCE optimization. Because it considers many different "implementations" of the term, the space-time optimizer is able to select variants for each memory size that minimize the recomputation cost, whereas in the hand coded case, only a single algorithm has been selected and implemented based on a more intuitive consideration of the problem.

We have given numerous talks on the TCE in both the computer science and chemistry community, which have been received with great interest in both communities. Many quantum chemists, in particular, have expressed interest in using and contributing to the development of the TCE. We will be holding a workshop in conjunction with the 2004 Sanibel Symposium in February, at which we will make the first formal release of the TCE software to the chemistry community, which we hope will begin a long community-wide collaboration to improve and extend this technology.

Work on nuclear coupled cluster theory has been very productive as well. We carried out calculations on the nuclei ^{4}He and ^{16}O, which are magic in both the neutron and proton shells. We used an oscillator basis with up to

360 single-particle orbits, and the Idaho-A nucleon-nucleon interaction, which is based on chiral perturbation theory. This interaction preserves the underlying symmetries of the QCD Lagrangian and precisely reproduces the nucleon-nucleon experimental phase-shift data. We renormalized this interaction for our model space and then performed the CCSD calculations within the same space.[12] We also performed during this same time calculations of the non-iterative triples corrections for these two nuclei, and for ^{4}He we investigated the EOM-CCSD method (an approach developed in quantum chemistry), which allows us to calculate excited states in nuclei.[13] By comparing with exact diagonalization results in a small model space (consisting of 80 single-particle orbits) with the triples-corrected CCSD results. We found that CCSD and the triples corrected versions compare extremely well with the exact answers obtained from exact diagonalization. The nuclear coupled cluster code used for these calculations has also been developed from scratch as part of this project, in order to serve as a reference to validate generated code in this domain. The code is now running in parallel at over 200 Mflops/processor on the largest jobs (on 256 processors), with the computational loops themselves running at 800 Mflops/processor and the rest of the time due to parallel communications overheads. In the near future, we plan to begin investigating the role of the three-body nuclear potential, which will require the capabilities of the optimizing TCE to produce implementations of this exceedingly complex method in a reasonable amount of time.

Finally, the HLL embedding work demonstrated the effectiveness of the approach by generating an implementation of the NAS parallel multigrid benchmark[14] that is performance competitive with the hand-coded reference implementation. The benchmark algorithm was expressed in just 60 lines of special-purpose high-level language, and another ten lines of code were required to express the stencil operations used. The generated code was 6000 lines of Fortran, which is comparable to the hand-coded version. One of the most interesting conclusions of this work is that there is a relatively limited set of features that are needed in a host language, and that these features can be satisfied by a language such as Python (a widely-used object-oriented scripting language),[15] as opposed to the extensive and complex features of a functional language such as Haskell. A paper describing this work is in preparation.

Summary and Conclusions

We have demonstrated the feasibility and utility of using high-level domain-specific languages together with specialized optimizing compilers to address both productivity and performance issues for complex scientific problems. This work provides a framework that we plan to extend to other problems and other domains. We have also made significant advances in the use of coupled cluster methods in nuclear physics, including introducing new methods developed by the chemistry community into nuclear physics.

Both aspects of this work provide significant benefit to the DOE. The problem of software productivity and the complexity of modern scientific software are issues which are rapidly reaching a crisis level; DOE is particularly effected by the problem because it has the largest concentration of computational science research in the United States. Our work shows the efficacy of one approach to dealing with productivity and complexity issues. Our work in nuclear physics will have the benefit of making it easier and more computationally efficient to compute the properties of nuclei to higher degrees of accuracy.

We are currently working on several proposals to secure follow-on funding. Dean is leading a proposal together with Piotr Piecuch and Karol Kowalski of Michigan State University that would explore the use of some of the coupled cluster-based methods developed by Piecuch's group in nuclear physics simulations. Harrison is leading a proposal, together with Bernholdt and other ORNL researchers to the Basic Energy Sciences Chemistry program that would include TCE-related work focused on quantum chemistry. In addition, Bernholdt, Sadayappan, Baumgartner, and Ramanujam plan to develop a proposal to DOE and/or NSF that would focus on the computer science issues of extending TCE technology into other domains. Finally, Bernholdt, Harrison, White, Sadayappan, and others have begun working with DOE MICS (computer science) program manager Fred Johnson to develop a new research program to address the larger issue of software productivity and complexity.

We are investigating funding opportunities through the NIH Bioengineering Consortium (BECON), in particular, the Exploratory/Developmental (R21) Bioengineering Research Grants program. Our collaborator, Eunok Jung, is visiting Lenhart in February 2004, and the plan of the proposal will be discussed at that time.

The implementation of the control strategy, obtained in our simulations for the optimal chest pressure profile, would have great benefit for saving lives in the civilian and military sectors. This biomedical innovation would be important to DOE and DOD.

References

[1]T. J. Lee and G. E. Scuseria. Achieving chemical accuracy with coupled cluster theory. In S. R. Langhoff (Ed.), *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy,* pp. 47–109, Kluwer Academic, 1997.

[2]J. M. L. Martin. In P. v. R. Schleyer, P. R. Schreiner, N. L. Allinger, T. Clark, J. Gasteiger, P. Kollman, and H. F. Schaefer III (Eds.), *Encyclopedia of Computational Chemistry*, Vol. 1, pp. 115–128, Wiley & Sons, Berne (Switzerland), 1998.

[3]G. Baumgartner, D. E. Bernholdt, D. Cociorva, R. Harrison, S. Hirata, C.-C. Lam, M. Nooijen, R. Pitzer, J. Ramanujam, and P. Sadayappan, " A High-Level Approach to Synthesis of High-Performance Codes for Quantum Chemistry.," in *SC'02* (electronic publication), IEEE, 2002.

[4]D. Cociorva, J. Wilkins, G. Baumgartner, P. Sadayappan, J. Ramanujam, M. Nooijen, D. Bernholdt, and R. Harrison, "Space-Time Trade-Off Optimization for a Class of Electronic Structure Calculations," *Proc. Of the 2002 Conference on Programming Language Design and Implementation,* Berlin, Germany, June 2002.

[5]D. Cociorva, J. Wilkins, C. Lam, G. Baumgartner, P. Sadayappan, and J. Ramanujam, "Loop Optimizations for a Class of Memory-Constrained Computations," in *Proc. 15th ACM Intl. Conf. on Supercomputing,* 2001.

[6]D. Cociorva, J. Wilkins, G. Baumgartner, P. Sadayappan, J. Ramanujam, M. Nooijen, D. Bernholdt, and R. Harrison, "Towards Automatic Synthesis of High-Performance Codes for Electronic Structure Calculations: Data Locality Optimization," pp. 237–248 in *Proc. Of the Intl. Conf. on High Performance Computing,* Lecture Notes in Computer Science, Vol. 2228, Springer-Verlag, 2001.

[7]D. Cociorva, G. Baumgartner, C.-C. Lam, P. Sadayappan, and J. Ramanujam, "Memory-Constrained Communication Minimization for a Class of Array Computations," *Languages and Compilers for Parallel Computing,* 2002.

[8]D. Cociorva, X. Gao, S. Krishna, G. Baumgartner, C.-C. Lam, P. Sadayappan, and J. Ramanujam, *Global Communication Optimization for Tensor Contraction Expressions under Memory Constraints*, Tech. Report, Dept. of Computer & Information Science, Ohio State University, 2002.

[9]P. Wesseling, An Introduction to Multigrid Methods, http://www.mgnet.org/mgnet-books-wesseling.html.

[10]The Haskell Home Page, http://haskell.org.

[11]High Performance Computational Chemistry Group. *NWChem, A computational chemistry package for parallel computers, Version 4.5 (2003)*, Pacific Northwest National Laboratory, Richland, Washington, http://www.emsl.pnl.gov/docs/nwchem/.

[12]D. J. Dean and M. Hjorth-Jensen, "Coupled-cluster approach to nuclear physics," submitted to *Phys. Rev. C* (2003).

[13]K. Kowalsk, D. J. Dean, M. Hjorth-Jensen, T. Papenbrock, and P. Piecuch, "Coupled-cluster calculations of ground and excited states of nuclei," submitted to *Phys. Rev. Lett.* (2003).

[14]D. Bailey, J. Barton, T. Lasinski, and H. Simon, *The NAS Parallel Benchmarks*, Technical Report RNR-91-02, NASA Ames Research Center, Moffett Field, California, January 1991, http://www.nas.nasa.gov/Software/NPB/.

[15]The Python Home Page, http://www.python.org.

Cellular Algorithms for Next-Generation High-Performance Cellular Architectures

G. A. Geist, P. Agrawal, W. A. Shelton, and C. Engelmann
Computer Science and Mathematics Division

The world of high performance computing is rapidly changing, and a new type of highly scalable computer called cellular architecture is emerging as a potential next generation of computer for scientific simulation. In order to position ORNL for such petascale systems, we developed a new class of scientific algorithms that is able to scale to tens of thousands of processors and at the same time be tolerant of failures of some of these processors. We also developed a simulator for 100,000 processor computers that can support C, Fortran, and Java applications. The simulator can mimic single and multiple processor failures. Our research included collaboration with application scientists from materials and biology, as well as the Blue Gene group in IBM.

Introduction

ORNL has experienced tremendous growth in computing power in the past year and a half, going from 150 Gflops to over 5000 Gflops. If we are going to maintain this growth, we must begin preparing for the next generation of computers that will take us from 10 Tflops to over 100 Tflops.

This proposal has three goals:

1. Infrastructure—position ORNL for a 100 Tflops computer.
2. Genomes to Life—give ORNL a unique computational edge in the GTL initiative.
3. Materials Nanoscience—offer the opportunity to push the scale of first principles simulations.

Technical Approach and Results

Algorithms with "natural fault tolerance" have the property that they get the right answer even if one or more of the tasks involved in the calculation fail without warning or notification to the other tasks. In a sense, a naturally fault-tolerant algorithm just ignores failure and continues computing until the solution is found. Natural fault tolerance is an attractive property for cellular algorithms. For machines with 50,000 to 100,000 processors, the failure of a processor is predicted to occur every few seconds. Therefore, many failures will happen during a long scientific simulation.

In this project we have accomplished several important milestones. First, we have shown the existence of algorithms with both natural fault tolerance and the ability to scale to 100,000 processors. Second, we have created a cellular architecture simulator that can simulate a 100,000 processor system with arbitrary communication interconnect and the ability to mimic random failures. Third, we have presented two international and three domestic talks on this research at conferences.

We developed a meshless finite difference algorithm that is scale invariant and naturally fault tolerant. It can be used as a general template to create any number of scientific applications that can be formulated as a "function over some local region." To illustrate this, we took this algorithm and formulated it to solve Poisson's equation. The application gets the correct answer even if a hundred random parallel tasks are killed during the solution. The second challenge was to show the existence of natural fault-tolerant algorithms that are a function over the global region. Global information is often needed to determine convergence or maximum energy or things of this nature in large scientific applications. We were able to create a naturally fault-tolerant global maximum algorithm. At the beginning, all tasks had some random value. At the end of the algorithm, all tasks (that were still alive) knew the largest value and who had it. The algorithm was robust no matter which tasks died during the solution. We then developed a naturally fault-tolerant multigrid algorithm. We explored four different methods to reduce global synchronization within and between different grids. Only two were found to have the fast convergence property of serial multigrid.

During the first year we obtained and installed a Blue Gene simulator from IBM. We found that it simulated one Blue Gene node very accurately but was not designed to simulate hundreds (much less thousands of processors). We began development of a simulator that could simulate at least 100,000 processors and could kill any one or set of simulated processors while an application was running on it. In the first year we completed this simulator, which is itself a parallel application that runs on Linux clusters.

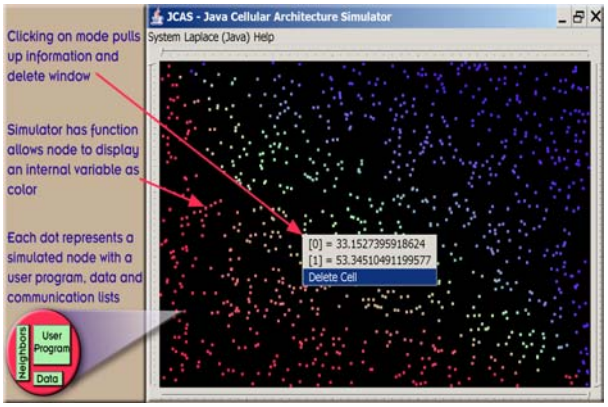

Fig. 1. Basic features of our simulator that can study the effects of faults and application recovery on computers with over 100,000 processors.

Using this simulator we were able to test the Poisson, global maximum, and multi-grid applications on 500,000 processors (see Fig. 1).

In materials nanotechnology, the major accomplishment in cellular algorithms was the development of a new K-space, O(N) first-principles screened Korringa, Kohn, and Rostoker (KKR) electronic structure method that has the potential to take advantage of next-generation architectures based on a cellular topology. This new method is more accurate than the locally self-consistent multiple scattering (LSMS) method. Equally important, we have been able to treat 2048 atoms on a single processor; thus, this new computational technique is capable of treating tens to hundreds of thousands of atoms. Similar to the LSMS method, the screened KKR constructs the so-called screened structure constants in real space by communicating with a local set of nearest neighbor atoms. The screened structure constants are then transformed to K-space via a Fourier transform, providing spectral or pseudo-spectral accuracy. Thus, the communication pattern is similar to the LSMS, which is an ideal algorithm for next-generation architectures based on a cellular topology. On the other hand, the LSMS method is purely a real-space method that lacks spectral accuracy and, thus, the energetics may not be of high enough accuracy to properly describe the underlying physics.

In computational biology, the challenge of folding a protein, through computer simulations, lies in finding the global minimum of the protein conformational energy landscape; sufficient evidence exists that the folded structure corresponds to a region close to the energy minimum. The size of protein conformational energy space increases very rapidly with the size of protein, and the conformational space for a biologically relevant protein is extremely large. It is impossible to exhaustively explore the complete conformational space even for a small protein in a reasonable time frame, given its multidimensional nature. The computational resources required for realistic protein folding simulations and resources available today fall short by several orders of magnitude. IBM's Blue Gene will drastically improve the available computational resources for large simulations. Therefore, we describe here the use of IBM's Blue Gene architecture to address the problem of protein folding by efficiently reducing the multi-dimensionality of the protein conformational energy landscape and matching it to the Blue Gene's architecture topology.

We are developing a new method to rapidly simulate the process of protein folding by reducing the multi-dimensionality of the protein conformational energy space to be searched. This new method is based on the understanding of protein dynamics in solution and using the knowledge of the physical principles which govern the process of protein folding. This method will take advantage of the Blue Gene architecture with thousands of closely connected processors. The method is being designed to be highly scalable with respect to the size of protein molecule and the number of processors.

In the proposed method, the multidimensionality is reduced by describing the modeled protein system by a set of essential dynamics variables. Such a description allows propagation of independent molecular dynamics trajectories to explore distant areas of the conformational space concurrently. Therefore, this approach is suitable for parallel computers with a large number of processors. The essential dynamics information obtained from simultaneous molecular dynamics trajectories are processed to direct the search into new regions of the conformational space. It is important to note that proposed method is different from standard molecular dynamics and Monte Carlo method as it is allows distant areas of the protein conformational space to be searched rapidly. In the current form, this method is implemented by using an iterative scheme. Each iterative step is composed of energy minimizations and molecular dynamics runs, which explore several local regions of protein conformational space. Based on the analysis of regions sampled during each iterative step, a set of new regions is explored in the next iteration, until global minimum is reached.

To take advantage of the Blue Gene's architecture with a large number of processors, the protein conformational space is mapped onto the topology of the machine. This is performed by dividing the protein conformational energy landscape into separate regions using the essential dynamics variables and then independently exploring these regions for the local minimum using a CPU block (see Fig. 2). The size of CPU block will be case dependent; the size of protein and scalability of the modeling program being the main factors. A small calculation at the beginning of each case will be

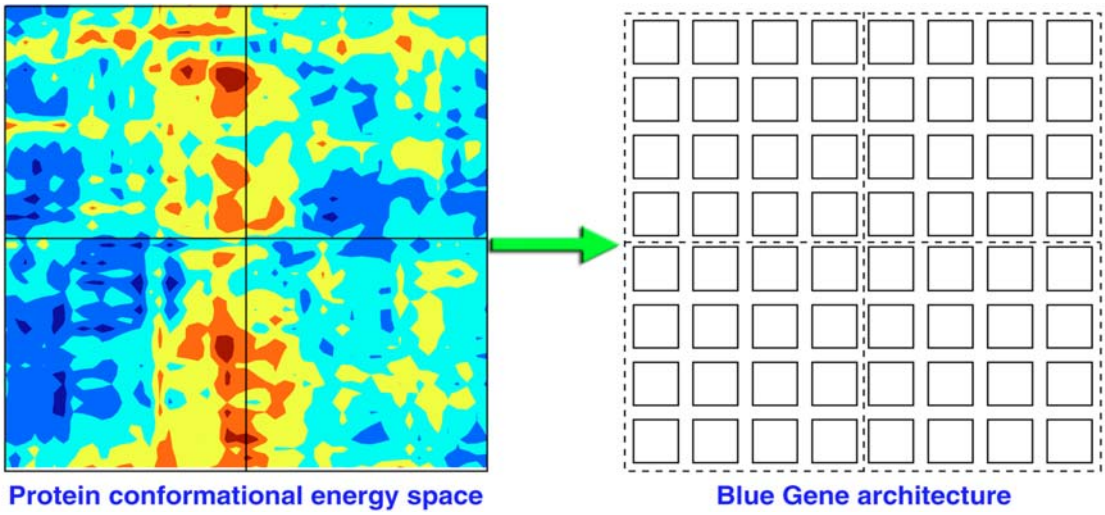

Fig. 2: Mapping the protein conformational energy surface onto the Blue Gene architecture.

able to determine the number of CPUs in each block for optimized performance.

Summary and Conclusions

As noted above, this project has been successful at strategically placing ORNL in position for petascale systems. We have been able to establish the existence of three broad classes of scientific solution methods that are naturally fault tolerant: finite difference, multi-grid, and global operations (such as global maximum). We have presented this work at invited talks at two international conferences and three conferences in the United States. We have also published the findings in two papers.[1,2]

The Blue Gene simulator obtained from IBM was found to be too low level for scalable application development. It could simulate Blue Gene nodes at the bit level but could only simulate up to 32 processors. In contrast, the simulator we developed can simulate up to 500,000 processors using a 64-node Linux cluster. It supports Fortran, C, and Java applications and is able to study the effects of processor failure on these applications.

Feasibility studies for using systems with 100,000 processors in ORNL key areas of biology and nanotechnology were completed and new algorithms developed that can take advantage of large-scale systems likely to exist at ORNL in the next three to five years.

Finally, this project has met its goals of being a catalyst in bringing in new funding to ORNL. One of the objectives of this project was to strategically position the Laboratory for a 100-Tflops computer. A proposal was prepared and submitted to DOE in July 2002 that if funded would place a Cray X1 computer into ORNL by the end of this project and would set the stage for placing a $100M computer into ORNL by 2005. The proposal, which is in response to the Japanese Earth Simulator, leverages the expertise we are gaining in super-scalar algorithms from this LDRD and ORNL's leadership in scalable systems software.

The project also led to a proposal to the DOE Genomes to Life initiative, which was subsequently funded for $19 million, of which $5.7 million will come to ORNL.

The third development activity was getting an IBM CRADA in place to help augment the research into cellular algorithms in general and Blue Gene in particular. This CRADA involves both funds-in from IBM as well as matching funds from DOE and is expected to bring over $2.4 million into ORNL over the past three years.

References

[1]A. Geist and C. Engelmann, "Development of Naturally Fault Tolerant Algorithms for Computing on 100,000 Processors," *Journal Parallel Computing*, December 2002.

[2]A. Geist, "Progress Towards a Petascale Virtual Machines," pp. 10–15 in *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, ed. Laforenza and Dongarra, Springer LNCS 2840, September 2003.

Scalable Tools for Petascale Distributed-Data Analysis

G. Ostrouchov, N. F. Samatova, D. J. Erickson, and R. J. Toedte
Computer Science and Mathematics Division

This project concerned the development of scalable algorithms to perform complex analyses of massive distributed-data sets. The focus was on scalable dimension reduction computations and visualization for astrophysics and climate simulation data, and for biological data. Our results include several new algorithms for dimension reduction that operate on both distributed and streaming data, as well as breakthrough theoretical results that pave the way for faster and more robust dimension reduction computations on massive data sets.

Introduction

The current and future landscape of data analysis is shaped by the relative innovation speed in four basic components of modern computing: processor, memory, disk, and network. Trends recently reported in *Scientific American* show exponential growth in all components with disk capacity outpacing the rest by a factor of 2 per year. The result is that data sets are becoming massive, distributed, dynamic, and essentially immoveable. Nearly all data analysis software fails on such data sets because of excessive memory requirements and the need to bring the data to a central location. Further, the massive distributed-data problem exists not only in the usual sense over a wide area network but also on a single storage device, where data sets are fragmented into smaller files in response to memory constraints. Terabyte data sets from climate and astrophysics simulations are examples of such data, often residing on storage systems like the High-Performance Storage System (HPSS) as a collection of thousands of files. Biology data sets can be massive not only because of the number of genomes and genome components but also because of increasing availability of annotation tools that can produce very-high-dimensional feature sets for the genomes and genome components.

Dimension reduction techniques are fundamental to discovery and visualization of structure in high-dimensional data. These techniques are still underutilized in science even on smaller data sets, and their use will grow as we learn how to pose useful dimension reduction analyses in more applications. We also point out that dimension reduction methods are not only of interest for visualization of structure, but they are also enabling tools for other data analysis methods by reducing the amount of data for further processing. Pushing the frontier of dimension reduction tools to massive distributed data sets enables progress across many data science applications.

Climate model data archives generated at ORNL are approaching 40 terabytes and are increasing exponentially. These model-generated outputs are typically centuries in length, and the time resolution of the archived data sets is 6 hours to 1 month. The spatial resolution is usually 18–70 vertical levels and ~250-km horizontal grids for the global simulations. Several hundred "variables" are computed for each grid point in three dimensions for every 20 min, for centuries. Since general circulation models are highly nonlinear, the need to have scalable dimension reduction is a critical aspect of climate model interpretation.

Under the SciDAC activity lead by ORNL, supernova modeling continues with increasingly sophisticated multidimensional simulations with ever improving physics, particularly multidimensional neutrino (radiation) transport. The ultimate goal is to ascertain the key ingredients that constitute the core collapse supernova mechanism. Dimension reduction methods will be used to uncover spatial structure and relationships in the multidimensional simulation.

Technical Approach

Dimension reduction (DR) techniques are computationally complex but provide the most effective general means of discovering structure in high-dimensional data. For example, the result of an application of principal component analysis (PCA) to climate simulation data is shown in Fig. 1. First, it matters what projections are selected as some have strong structure and some look random. Second, each structured projection has an interpretation. For example, (2,4) in the middle illustrates that global warming affects winter temperatures the most and summer temperatures the least. Winter months are on the left, and summer months are on the right. And third, each component is also visualized as a map, giving insights

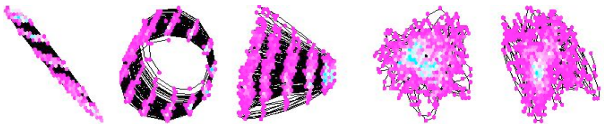

Fig. 1. Monthly surface temperature principal component dynamics of a 120-year climate model simulation under transient CO_2 increase (a 2500-dimensional time series representing 2500 locations in the northern hemisphere). Pair wise dynamics of the top six components are shown as they seasonally rotate through 12 months; all are two-dimensional projections of the same, roughly 2500-dimension object. Components shown are (1,2), (2,3), (2,4), (5,6), and (4,5), where (2,4) in the middle illustrates that global warming affects winter temperatures (left) the most and summer temperatures (right) the least.

into which regions are associated with certain component dynamics. Figure 2 displays the second principal component, where the coefficients are a temperature contrast between land and sea. PCA is by far the most popular and simplest of the DR techniques [known as empirical orthogonal functions (EOF) in climate, Karhunen-Loeve transform in signal processing, and is also related to normal modes in engineering and molecular structure simulation].

Algorithms for parallel matrix computation, and also algorithms for out-of-core matrix computations, provide some insights for organizing distributed DR computations. However, there are important differences. First, we cannot assume that we distribute the data to optimize performance of our algorithm; rather, the data are already distributed according to a given application and it is our task to minimize data transfers required by our computation. This can be considered a disadvantage relative to the parallel and out-of-core computations. Second, DR is a form of approximation and that presents an opportunity and an advantage over exact parallel matrix computation through local DR and reduced data transfer rates. Finally, in DR the matrix factorization itself is of interest rather than the solution of a matrix problem.[1]

We considered data distributed over several locations as well as data streaming from a running simulation experiment. Our approach to analysis of distributed data was driven by the concept of performing local analyses

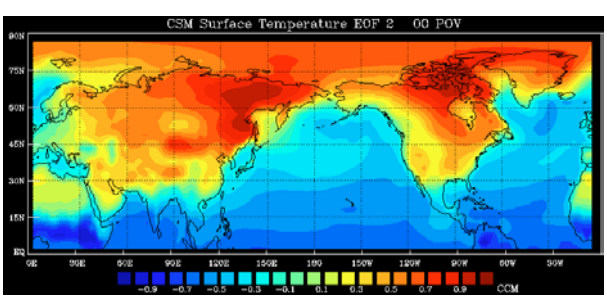

Fig. 2. Second principal component is shown as a map with positive coefficients in shades of red and negative as shades of blue.

on local data that are combined into a global analysis with minimal data transfers. We assume that the data are already distributed according to the needs of a given application and any data transfers must only involve the minimum information needed to compute a required global DR result. Such methods also, in part, address the issues of streaming data. As more data are generated, updating can be accomplished by considering the new data as separate distributed data sets that are sequentially combined into the global analysis. Within our focus on DR techniques, we view the distributed data as components of a massive virtual matrix, where columns are the features defining the dimensionality of the problem, and rows are the items. How such a matrix is constructed is highly varied among applications, although common concepts exist. One such concept is the construction of empirical orthogonal functions in climate simulation.

Results and Accomplishments

Two dimension reduction (DR) methods were selected for the distributed data problem. One is principal component analysis (PCA), which is currently the most popular DR method, and the other is FastMap, a distance-based method that can be viewed as a fast approximation to PCA. We considered the application of these algorithms to data distributed across several locations as well as streaming data. Because fast algorithms are particularly relevant in the massive data set problem, we considered theoretical properties of FastMap. This provided the basis for a new class of adaptive algorithms that are both fast and robust. This class of adaptive algorithms will likely provide the breakthrough performance that is needed for massive petascale data sets. Finally, we also considered very-high-dimensional data sets generated with annotation tools from biological data and developed the statistical basis for extracting relevant dimensions. We describe all these results.

Distributed PCA. We developed an algorithm for the application of PCA to data distributed over several locations.[2] Our algorithm does not require that the local data is transferred to the central location. Only an approximation of each local covariance matrix along with a vector of means is centralized. The basis of our algorithm is that the global data covariance can be partitioned into a sum of "within" locations covariances and a "between" locations covariance. We approximate the "within" local covariances with principal components and compute the "between" covariance centrally. Each local approximation carries the linear optimality properties of PCA. Compared to an algorithm that centralizes the data, our algorithm reduces data transfers from O(np), where n is items and p is features, to O(sp), where s is the number of data locations. The number of items is typically one or more orders of magnitude larger than the number of data

locations. For example, our algorithm computes the top 50 empirical orthogonal functions across *s* climate simulations stored at *s* locations, each with 2400 time steps (200 years of monthly averages) and a spatial resolution of 7000 grid points using data transfers of only (s)×2×50×7000 + 50 data items. Centralizing the data requires transferring (s)×2400×7000 data items, an amount roughly 24 times larger. For large simulations, combining so much data at one location may not be possible.

FastMap Properties and a New Class of DR Algorithms. FastMap is a DR technique that operates on distances between objects. Implicitly, the technique assumes that the objects are points in a high-dimensional Euclidean space. It selects a sequence of orthogonal axes defined by distant pairs of points (called pivots) and projects data onto the orthogonal axes. We showed that FastMap picks all of its pivots from the vertices of the convex hull of the data points in the original implicit high-dimensional Euclidean space.[3] This provided a connection to results in robust statistics, where the convex hull is used as a tool for multivariate outlier detection and for robust estimation methods. The connection sheds a new light on some properties of FastMap, in particular its sensitivity to outliers, and provides an opportunity for a new class of DR algorithms that retain the speed of FastMap and exploit ideas in robust statistics. The main lesson from robust statistics is that the most distant points are often not the best choice for defining a projection axis. The key to new fast and robust methods is a replacement of FastMap's pivot points heuristic by something that considers more than just the maximum distance from a point. Many other properties are available within $O(np)$ complexity, thus retaining FastMap speed. In fact, methods using these properties will be more robust than standard PCA. There are many directions that this methodology can take, and undoubtedly many such algorithms will be proposed, particularly for DR in massive data sets. We have started this with several algorithms for distributed and streamline data.

DFastMap: Distributed FastMap Algorithms. Although FastMap is an order of magnitude faster than PCA, it too requires that the data are in a central location. We have developed two variants of FastMap that perform most of their computation locally near each distributed data set and centralize only some minimal information to compute a global result that is comparable to centralizing the data. The intuition behind our approach is as follows. FastMap tends to select each pair of pivots so that they are widely separated and among the extreme points of a data set. We present two approaches,[4] one is iterative and the other direct. Our connection of the convex hull to FastMap pivots[3] gives an explanation of why an application of DFastMap to distributed data performs as well as the serial FastMap on a centralized data set. The union of local convex hull vertices includes all convex-hull vertices of the centralized data set. DFastMap centralizes the pivots, arguably a very good subset of the local convex hull vertices.

Streamline FastMap Algorithm. This data model—often called data streams—includes scientific simulation output, satellite images, network traffic data, etc. Processing data streams presents both practical and theoretical challenges; it often requires immediate results as streams flow in. We developed a novel dimension reduction algorithm Xmap,[5] based on the FastMap heuristic. Xmap exploits the connection between FastMap and the convex hull of data in Euclidean space. Xmap effectively identifies a small representative subset of past data and maintains well-approximated lower-dimensional map of data received up to any given time.

Streamline Reservoir Sampling Algorithm. Random sampling is an accepted basis for estimation from large data sets that outstrip available computer memory. When the data comes as a stream, its total size is potentially infinite and usually only one pass through the data is possible. Reservoir sampling is a method of maintaining a fixed-size random sample from streaming data. Recently, several methods for reservoir sampling without replacement were introduced. We introduced a new method for reservoir sampling with replacement.[6] We first proved that the proposed method maintains a random sample with replacement at any given time. Then we introduced a refined version that significantly speeds up the overall sampling procedure. Sampling with replacement is relevant for methods such as bagging and bootstrapping, both necessary for uncertainty assessment in complex estimation settings.

High-Dimensional Annotation Data Reduction for Detecting Interactions. Annotation tools in biology produce massive data sets of labels or indicators. Such data are high-dimensional but are not points in Euclidean space as are the data necessary for the PCA and FastMap methodologies discussed earlier. However, such label data can be reduced to a potentially very large multidimensional table of counts. The ways in which this is done can vary depending on a particular application. We investigated such constructions for protein interaction data and developed methodology for reducing such tables to a few significant cells that are key for predicting interaction.[7]

Scalar Field Data Reduction and Reduced Views. We have also developed novel dimension-reduced views of instability propagation in the entropy field of a supernova simulation, shown in Fig. 3. These views provide a compact summary of various characteristics of an entire simulation and provide a new effective means of comparison between two- and three-dimensional simulations.

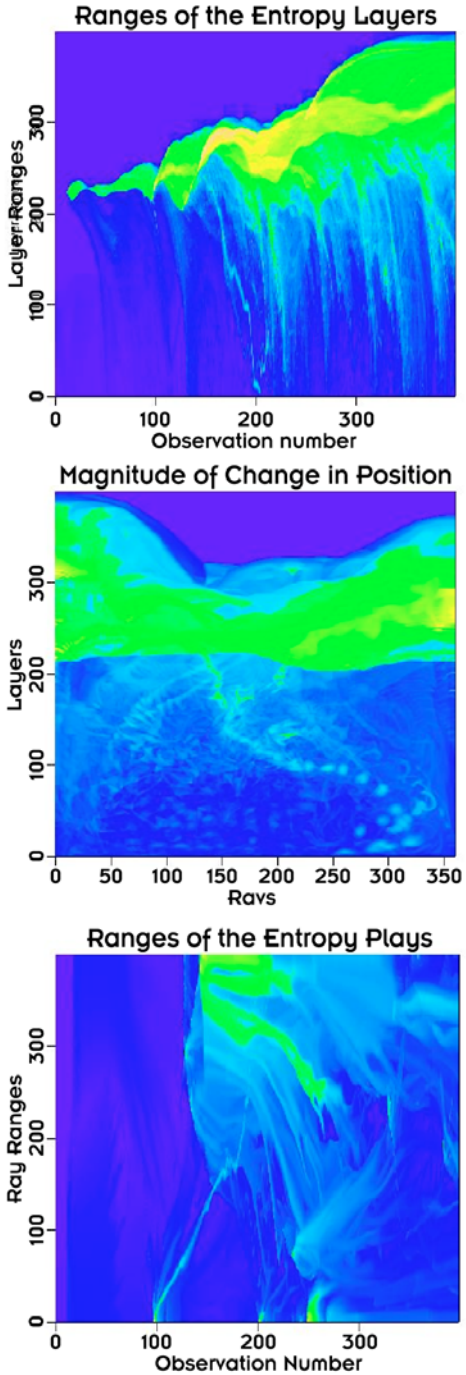

Fig. 3. Three orthogonal views of entropy instability in a 400 time-step 2-d supernova simulation are shown with polar coordinates presented as Cartesian. Spatial and temporal propagation of the instability can be seen from three points of view. These views remain preserved after 30x data reduction based on adaptive principal components.

Summary and Conclusions

The most promising result of this project is the theoretical connection of a distance-based DR to the convex hull of a high-dimensional data set. This provides the basis for a new class of adaptive algorithms that are both fast and robust. This class of adaptive algorithms will likely provide the breakthrough performance that is needed for massive petascale data sets. Strong results were also obtained on several algorithms for complex analyses of massive distributed data sets. Among these is a new distributed version of principal component analysis, which is the most popular DR technique. Other distance-based DR algorithms were developed for distributed as well as streaming data. Dimension reduction methods are not only of interest for visualization of structure but they are also enabling tools for other data analysis methods by reducing the amount of data for further processing. Pushing the frontier of DR tools to massive distributed data sets enables progress across many data science applications, including data from climate and astrophysics simulations.

This research puts the ORNL at the forefront of scalable algorithm development for DR computations. This capability is unique among the DOE laboratories and positions ORNL for future work in complex analysis and visualization methods for massive data sets. This also opens the door to developing useful analysis sequences that involve dimension reduction in applications that produce massive quantities of data.

Our work has resulted in six refereed publications of which three[2,4,5] have already appeared and the other three[3,6,7] are currently under review. In addition, we have co-sponsored the Statistical Data Mining Conference, June 22–25, 2002, in Knoxville jointly with the University of Tennessee Department of Statistics, where we also organized a special session on Distributed Data Mining. Several presentations of most of our results were made at other international conferences, including the Spring Research Conference on Statistics (2002 and 2003) and the Workshop on High Performance Data Mining at the Second (and Third) SIAM International Conference on Data Mining (2002 and 2003).

References

[1]L. Hubert, J. Meulman, and W. Heiser, "Two purposes for matrix factorization: a historical appraisal," *SIAM Review* **42**, 68–82 (2000).

[2]Y. M. Qu, G. Ostrouchov, N. F. Samatova, and G. A. Geist, "Principal Component Analysis for Dimension Reduction in Massive Distributed Data Sets," pp. 4–9 in *Workshop on High Performance Data Mining at the Second SIAM International Conference on Data Mining*, 2002.

[3]George Ostrouchov and Nagiza F. Samatova, "On FastMap and the convex hull of multivariate data: Toward fast and robust dimension reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, submitted for publication, 2003.

[4]F. N. Abu-Khzam, N. F. Samatova, G. Ostrouchov, M. A. Langston, and A. G. Geist, "Distributed Dimension Reduction Algorithms for Widely Dispersed Data," pp. 174–178 in *Parallel*

and Distributed Computing and Systems, Editors: S.G. Akl and T. Gonzalez, ACTA Press, 2002.

[5]B.-H. Park, N. Samatova, G. Ostrouchov, and G. A. Geist, "Fast dimension reduction algorithms for multivariate streamline data," pp. 1–6 in *Proceedings of the 6th International Workshop on High Performance Data Mining: Pervasive and Data Stream Mining*, 2003.

[6]B.-H. Park, G. Ostrouchov, and N. F. Samatova, "Reservoir-Based Random Sampling with Replacement from a Data Stream," in *Third SIAM International Conference on Data Mining*, submitted. April 22–24, 2004, Kissimmee, Florida, submitted for publication, 2004.

[7]B.-H. Park, G. Ostrouchov, G.-X. Yu, A. Geist, A. Gorin, and N. F. Samatova, "PICUPP: Protein Interaction Classification by Unlikely Profile Pair," *Journal of Computational Biology*, submitted for publication, 2003.

Scalable Visualization Tools and Technology

S. P. Dickson
Computer Science and Mathematics Division

There is a critical need for visualization tools and technology that support continued scientific progress in light of the immense growth of scientific dataset size and complexity, as generated by large-scale software simulations. This looming shortfall affects a wide variety of scientific domains, including, but not limited to, astrophysics, biology, climate modeling, material science, and fusion research. The growing emphasis on ultrascale software simulation exacerbates this problem, and visualization paradigms are a key technology for enabling the next generation of scientific discovery. Yet, the ad hoc manner in which visualization has typically been applied is highly inadequate for upcoming DOE Office of Science computational programs. The latest visualization tools and infrastructure must be carefully integrated and applied toward scalably viewing and intuitively interpreting these complex and massive data sets. This project has focused on the evaluation and development of state-of-the-art visualization technologies, including the integration of visualization hardware and software into a functional production-scale user facility. A broad survey of the visualization field has been made, and two distinct scientific visualization facilities have been designed. Several collaborations have also been formed with scientific research groups at ORNL to help educate them in their use of visualization for scientific discovery.

Introduction

Visualization is often applied to scientific datasets to provide an intuitive visual method of presentation and exploration, especially for more massive and complex data collections. Software simulation experiments at ORNL are already poised to generate data in the terabyte (fusion, climate, and astrophysics) to petabyte (biology) range in the near term. An elaborate data analysis and navigation infrastructure is required to enable *any* practical level of scientific discovery for these increasingly difficult scenarios. Yet, such an infrastructure consists of many challenging subproblems, including large-scale data storage, networking, data reduction, analysis and mining, and of course, visualization tools and algorithms and scalable high-resolution graphical displays. Unfortunately, adequate visualization support is often "assumed" when considering the needs of large-scale simulation experiments. Nonetheless, the technology and systems required to efficiently visualize such increasingly immense amounts of data are far from "turn key" solutions.

ORNL has a respectable and growing competence in many computer-related research areas, including high-performance scientific simulation, parallel and distributed computing, and large-scale data mining and analysis, but there are serious technology gaps with respect to scalable visualization tools and the associated technology. While many useful visual images of science have been generated over the past several decades, the visualization technology and facilities at ORNL were antiquated and in serious need of improvement. Further, most local scientific research teams have not been using visualization, even where there could be great benefit from doing so. Other research groups that have applied visualization technology have typically done so with simple, low-resolution techniques, using an informally organized infrastructure.

A vital and unified research program must therefore be systematically initiated at ORNL in several areas, including application-specific scientific visualizations, parallel rendering, scalable visualization and analysis pipelines, interactive data exploration, and large-scale display management. The purpose of this LDRD activity has been to promote such growth in state-of-the-art visualization at ORNL. The specific goals of this work were to (1) explore visualization ideas and platforms that support scientific application areas and (2) apply emerging technologies toward the development of new visualization hardware/software facilities at ORNL.

Technical Approach and Results

The two main thrusts of this work relate to opposite ends of the scientific visualization spectrum – exploring application-specific visualization solutions for key application areas and developing a general-purpose scalable visualization facility hardware design. The former requires interactions with domain scientists to determine

their general needs and how visualization might assist their efforts. The latter involves concrete design of a production visualization venue where scientists of all domains might come to explore their large datasets. Both of these efforts required an evaluation of the existing visualization technology, in both hardware and software.

Application-Specific Visualization

A variety of scientific application domains would benefit from visualization that is specially catered to the features and processes unique to associated physical models. Obvious examples include rendering molecules for chemical or materials simulations, but this idea could be extended to other fields such as biology, genomics (protein folding), climate, high-energy physics, and fusion science. Highly accurate and literal visual depictions can be employed to effectively communicate specific phenomena being modeled. Many new science research programs are also including a visualization component in their funding – a new concept gaining credibility in many SciDAC (DOE's Scientific Discovery through Advanced Computing) application areas, such as the Terascale Supernova Initiative (TSI) astrophysics center.

To determine the best tools and technologies to apply for application-specific visualization, and to avoid duplication of existing work in this area, a brief survey of several prominent visualization systems was performed. The following systems were explored: the VMD[1] system (University of Illinois) for visualizing the results of Molecular Dynamics, DIVERSE[2] (Virginia Tech) for visualizing data on multi-tile and immersive display systems, the Visualization Toolkit[3] (VTK, KitWare) for general-purpose scientific visualization, and ASPECT[4] (ORNL) for applying a variety of data-analysis functions in visualizing large-scale datasets. To explore the applicability of these tools to the variety of common research platforms used by scientists in different application domains, these visualization tools were installed and tested on both MacOS and Linux workstations. Testing was also performed using a prototype parallel visualization cluster dedicated to driving various display facilities at ORNL. Several experiments were performed to combine and integrate sets of these tools as would likely be required for production visualization scenarios; for example, VMD was combined with Chromium[5] (University of Virginia), a pixel compositing software system for parallel rendering and display.

Several meetings were arranged with various domain scientists at ORNL to determine their current uses of and potential needs for visualization. Based on interactions with the Computational Biology Group, several approaches were explored for visualizing the molecular models in the Protein Data Bank file format. Similarly, volume datasets from various mathematical models, and experimental data sources such as from X-ray computer tomography (CT), were visualized using the VTK system. An improved visualization was prepared of the Virtual Human work done by researchers in ORNL's Computational Sciences and Engineering (CSE) Division.

In collaboration with researchers in the Computational Biology Group, a view was developed of the *E. Coli* 1ML5 molecular model using the VMD software package on both Linux and MacOS-X systems (Fig. 1). Similar methods were explored to visualize the 1ML5 molecular model using VTK on a Linux system. As a result of this view development, several software bugs were identified in VTK's Protein DataBank (PDB) and VRML file readers; these problems were subsequently resolved in cooperation with VTK developers at Kitware. Protein Data Bank models were also visualized in VMD using Chromium on the prototype visualization cluster and PowerWall. Several other datasets were visualized, including a CT dataset of a porcine thorax using VTK, and data from the Virtual Human effort at ORNL. Much of this work was collected into a 20-min high-resolution movie to showcase ORNL science programs at the IEEE Supercomputing 2003 Conference.

Several external conferences, tutorials and workshops were attended to exchange information with peer researchers on the currently available visualization tools and determine the fastest growing areas of visualization research: the ACM SIGGRAPH and IEEE Visualization conferences, tutorials and workshops; the UTK Conference on Conformal Geometry; and the DOE Computer Graphics Forum.

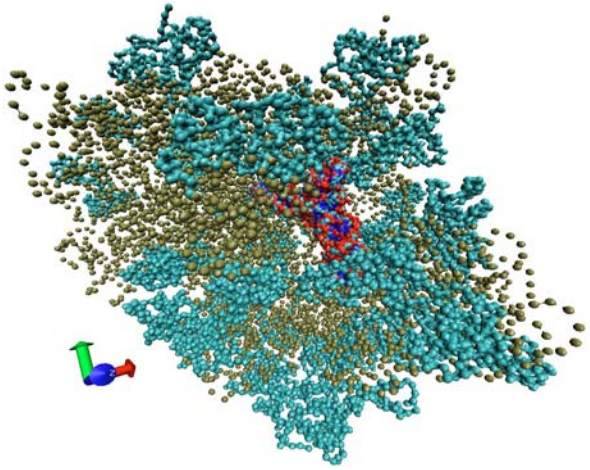

Fig. 1. View of the E. Coli 1ML5 *protein model.*

High-Resolution Visual Display Facility

The many powerful visualization tools are useless without a proper visualization facility infrastructure within which to execute them. A proper venue is needed for bringing together datasets from various application domains, to render and visually display the data for interactive exploration. When manipulating and visualizing such massive data sets, it is crucial to maximize the capabilities of the display hardware to provide the best opportunity for scientific discovery, by revealing fine-grained features and relationships in the scientific data. While ORNL has supported several useful visualization laboratories over the years, the current equipment was antiquated and unable to adequately handle the high-resolution needs of the large-scale data being produced. To this end, a new high-resolution visual display wall, known as the "Science Exploratorium," was designed.

Initially, a literature search[6,7] was performed to identify the proper hardware, software and facilities infrastructure necessary to construct a large "PowerWall" visualization theater. In addition to research papers, several conference sessions were attended regarding the principles of designing, constructing, and maintaining such a large-scale display. In conjunction with an ORNL Visualization Task Force, various display technologies, including screens, projectors, and mechanical mounting systems, along with rendering cluster and network infrastructure, were researched to design a production visualization facility for ORNL's new Computational Sciences Building.

The resulting design consists of three large adjacent glass screens, each 8 ft high and 10 ft wide, driven by an array of 27 powerful (high-lumen) rear-projection projectors (9 projectors per screen). A large 64-node computational cluster is dedicated to parallel rendering to drive the projectors for this "tiled" display wall, wherein each projector fills in the pixel illumination for one of many small rectangular areas, or "tiles," on a given screen of the display. The collection of projectors must be carefully synchronized, aligned, and color-matched to produce a smooth and uniform image across the entirety of the display wall (Fig. 2). The nature of these subtle adjustments is surprisingly complex and is still an active area of significant research.[8]

During the development and acquisition of this multi-million-dollar visualization facility, a simple prototype system was constructed using three large screens with only one projector per screen. This scaled-down version, at one-ninth the resolution, enabled much of the software infrastructure to be explored in advance of the final system installation. Many of the visualization tools explored in the first phase of this project were applied to drive the prototype PowerWall facility using a small cluster test bed. Several datasets from different application domains were selected for these initial experiments, as discussed in the following section, already enabling useful exploration of several scientific results (Fig. 3). In recognition of this effort, the Visualization Task Force received a Significant Event Achievement award for their contributions to the development of the new facilities.

Fig. 2. Projector array for tiled display walls.

In conjunction with the primary "Exploratorium" theater design, a smaller facility was designed for use in a conference room in the new Research Office Building at ORNL. This system consists of 6 projectors that drive a display in tiles on a 12-ft-wide by 8-ft-high glass screen. Based on the experiences from these two visualization facility designs, recommendations were also offered on the design of audio-visual facilities for the new Joint Institute for the Computational Sciences (JICS) Building at ORNL.

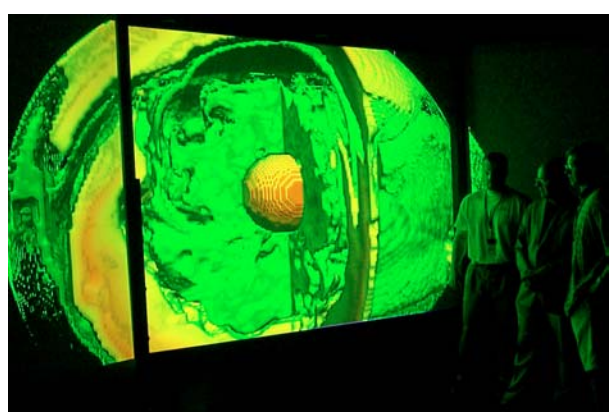

Fig. 3. Prototype powerwall visualization display.

Summary and Conclusions

Visualization provides an intuitive representation of the results of scientific computation, leading to useful insight into the science being explored. Visualization is often the point of reference through which scientists on an investigative team discuss their individual contributions toward the final results. The knowledge gained by this project, through research into PowerWall design and construction, has guided the planning and implementation of new visualization facilities in the Computational Sciences and Research Office Buildings. These facilities, even prior to full completion, have become a focal point for a variety of scientific meetings and collaborations. This result alone is significant.

The efforts of this project have also highlighted several scientific achievements at ORNL via visual presentation of their results. Observations of the activities in the Computational Sciences Building point to a broader use of visualization in exploring science and sharing project results. The expertise gained in scientific visualization will be of benefit to many existing and future scientific projects at ORNL and our collaborators.

References

[1]Visual Molecular Dynamics (VMD), Theoretical and Computational Biophysics Group, University of Illinois, http://www.ks.uiuc.edu/Research/vmd/

[2]DIVERSE Cross-Platform Open-Source Interface for Developing Virtual Reality Applications, Virginia Tech., http://diverse.sourceforge.net/

[3]The Visualization Toolkit (VTK), Open Source Software System for 3-D Graphics, Kitware, Inc., http://public.kitware.com/VTK/

[4]Scalable Data Management (SDM) Center, Office of Science, Scientific Discovery through Advanced Computing (SciDAC) Initiative, ASPECT Tool for Exploratory Analysis and Data Mining, http://sdm.lbl.gov/sdmcenter/

[5]G. Humphreys, M. Houston, Y.-R. Ng, R. Frank, S. Ahern, P. Kirchner, and J. T. Klosowski, "Chromium: A Stream Processing Framework for Interactive Graphics on Clusters," ACM Transactions on Graphics, Vol. 21, no. 3, (SIGGRAPH 2002), http://www.cs.virginia.edu/~gfx/pubs/cr/

[6]NCSA "Display Wall-in-a-Box," July 19, 2001, National Computational Science Alliance, http://www.ncsa.uiuc.edu/News/Access/Stories/IAB/

[7]Princeton University, "Scalable Display Wall." http://www.cs.princeton.edu/omnimedia/

[8]Raskar Ramesh, Paul Beardsley, "A Self Correcting Projector," Mitsubishi Electric Research Laboratories TR-2000-46, Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR) 2001, January 2002.

Scaling Climate Models for Future Computer Architectures

A. S. Bland, J. B. Drake, P. H. Worley , and J. B. White III
Computer Science and Mathematics Division

The 40-TFLOP/s Japanese Earth Simulator (ES40) has challenged the U.S. climate-modeling community and computer industry to accelerate the development of both the models and the computers required to run them, to keep our country competitive in both climate science and policy decisions. This project will characterize the performance of current U.S. models and suggest research areas for improvements. Researchers will use representative kernel codes to understand the factors that limit performance and scalability and then work to mitigate these factors. The second phase will be to instrument and evaluate the latest full climate model, CCSM2, to improve its performance. The objective is to prepare ORNL to assist DOE in developing and deploying an ultra-scale computer system for the U.S. science community.

The objective of this project is to prepare ORNL to assist DOE in developing and deploying an ultra-scale computer system for the U.S. science community. The approach to this is to begin scaling the components of the Community Climate Systems Model (CCSM) up to the resolution and physical fidelity that is expected to be used on this new class of computer systems and to port these codes to a new generation of high-end computer architectures.

This project has three major technical thrusts:

1. Benchmarking performance of a dynamical kernel on a high-resolution model to identify hardware and software issues with current and planned architectures.
2. Evaluation and implementation of performance enhancements for global spectral atmospheric models.
3. In-depth scientific performance study of the CCSM2 at high resolution on IBM p690 and Cray X1 systems.

The performance benchmark targets a spectral model of atmospheric dynamics. We will match the resolution of T1279L96 used in the ES40 benchmark that achieved 26 TFLOP/s. As currently configured, both the IBM p690 and Cray X1 at ORNL are currently insufficient to run this problem resolution. However, we have run experiments using problem sizes with large horizontal resolutions (up to T680 on the Cray and T1279 on the IBM) with small numbers of vertical levels and problem sizes with smaller horizontal resolutions and up to 92 vertical levels, determining the processor and memory requirements necessary to run the target problem size.

The experiments on the Cray were preceded by extensive performance tuning of the kernel code PSTSWM to improve its vectorization. We have improved the performance by a factor of 10 over the original version of the code on a single processor, achieving approximately 50% of peak for the largest horizontal resolution. Experiments indicate that maximum serial performance is achieved by computing on a few vertical levels at a time. This is not the approach used in the current production spectral model (nor in the Earth Simulator code) and has motivated investigating alternative parallelization strategies in the SciDAC project concerned with development of the production atmospheric model. The next step will be to run the full T1279L96 problem resolution on the upgraded Cray X1. Preliminary optimizations of the parallel algorithms for this large run have already been completed using the current system.

A high-resolution ocean configuration has been targeted using the POP code. The ES40 is showing performance of hundreds of Gflops in a 60-node configuration for POP at 1/10-degree resolution. The Cray X1 exceeds the ES40 performance for the low-resolution case, and we expect to find competitive performance with the high resolution case as well.

Results to date show that the Cray X1 architecture is superior to the IBM p690 in both performance and scalability for the CCSM. Future work will focus on completing scaling to T1279L96 resolution for benchmarking and maximizing performance of the full model for high-resolution configurations.

Advanced Computational Methods

R. J. Harrison, D. E. Bernholdt, and E. D'Azvedo
Computational Science and Mathematics Division

We are developing, evaluating, and deploying novel numerical methods for the solution of non-linear integral-differential equations on the next generation of massively parallel computers. Initial application is to computational chemistry, but the resulting methods and tools are expected to be widely applicable. In computational chemistry, we are developing methods that are both free of the basis-set error that plagues conventional methods and scale correctly with the system size. In a complementary effort, we are also developing a new and general framework for composing programs for the next generation of massively parallel computers with $O(10^5)$ processors.

In principle, all of chemistry can be understood and predicted from solution of the Schrödinger and/or Dirac equations. In practice, we cannot solve these 3-N dimensional (N, the number of electrons) differential equations and must resort to a variety of approximations. These approximations limit both the accuracy that may be obtained and the size of system that may be studied. We aim to remove several of these limitations with the objective of enabling more reliable and accurate computations on larger systems.

Conventional computational chemistry adopts the linear combination of atomic orbitals (LCAO) approximation and constructs effective one-electron molecular wave functions from a superposition of atomic wavefunctions. This gives very compact wave functions with desirable physical properties, since molecules are indeed composed of atoms. However, the LCAO approximation is not suitable for high-accuracy calculations and, because the atomic wave functions have extended support, the cost of calculations scales non-linearly with the size of the system. Many-electron wave functions are conventionally constructed from linear combinations of products of one-electron molecular wave functions (orbitals); however, this expansion converges very slowly. By using multiresolution methods in multiwavelet bases, we have already constructed fast algorithms for the solution of the density functional equations to arbitrary finite precision. Novel separated representations for integral operators make this approach feasible in three and higher dimensions. The use of multiwavelet bases with disjoint support enables high-order convergence even in the presence of singularities. This near-spectral convergence may be extended to many-body wave functions in several ways, and we plan to explore several of these options.

In collaboration with Prof. R. Butler (Middle Tennessee State University), we are also designing and implementing parallel tools for the support of irregular and hierarchical distributed data structures using one-sided access mechanisms on massively parallel computers. These will be built upon the Global Array (Harrison and Nieplocha) and ARMCI (Nieplocha) libraries that have already been optimized on essentially all current high-performance computers and workstation clusters. They both already exploit shared-memory within SMP nodes. We will add a primitive but efficient distributed hash table that can support a wide variety of data structures from tuple spaces to nested trees.

Creating New Climate Drivers and Interactions in Global Climate Models

D. J. Erickson,[1] M. Branstetter,[1] M. Sale,[2] W. M. Post,[2] A. King,[2] L. Gu,[2] F. Pan,[2] and S. Hadley[3]
[1]*Computer Science and Mathematics Division*
[2]*Environmental Sciences Division*
[3]*Engineering Science and Technology Division*

This project involves the identification of important climate drivers not already present in current climate models. The drivers to be perturbed are (1) carbon-cycle feedbacks, (2) hydrological-cycle feedbacks and (3) economic feedback responses to climate change. The scientific analysis of the significant feedbacks in the Earth system is the key to reliability and uncertainty estimation of future climate prediction and is critical in making meaningful policy decisions. The programmatic objective is to make ORNL the central DOE institution for computational climate modeling and to leverage the significant present and future ORNL computational resources to obtain future funding in the burgeoning field of numerical climate, carbon and water cycle modeling, and economic-climate feedbacks.

The technical R&D objectives of this LDRD are the evaluation and creation of new climate drivers/feedbacks in the climate modeling system. The installation of state-of-the-art global climate models and an assessment of feedbacks and critical biogeochemical physical parameterizations in the global climate models operating on ORNL computational platforms are central to this LDRD. The deliverables will be several refereed journal articles, meeting presentations, and a clear delineation of future (2–5 years) funding opportunities. This LDRD is intended to position ORNL climate science at the forefront of the U.S. climate modeling efforts.

We have a wide range of accomplishments to report at this juncture.

We have ported the U.S. state-of-the-art global Climate Model (CCSM2) to the CCS/CSM high-performance computing center. We have evaluated the control run of the CCSM2 (Branstetter and Erickson, 2003) and have performed a variety of sensitivity tests of the CCSM2.

An additional milestone is that we have, with colleagues at LLNL, ported the only fully coupled climate-carbon cycle model in the United States to ORNL computers. The effect of acclimation has been coded in to IBIS terrestrial biosphere submodel, and an experimental run has been staged. We have three additional test runs of the CCSM2 staged to explore new scientific characteristics of the CSM2.

In the area of hydrologic cycle feedbacks, we are developing new modeling procedures to improve how soil moisture dynamics and lateral transport of water is represented in the CCSM2 code. Our initial computational experiments completed this year indicate that the code for soil wetting and drying is not realistic and that there may be errors in the land surface modules of CCSM2. We are now in the process of recoding the important parts of the land surface model, including a new simplified algorithm and supporting global data for estimating TOPMODEL baseflow parameters and soil profile characteristics. This new approach is based on work completed by Pan this year that evaluated a number of different methods for calculating topographic index and flow direction.

We have several refereed publications published, in press, and in preparation.

Our program development strategy is initially to build an obvious expertise/group that is active in global climate modeling science. At the national CCSM2 workshop in Breckenridge, Colorado, in June 2003, ORNL had 14 scientific/computational scientists representing our efforts. We had several lectures, posters, and workshop participation on a variety if climate modeling issues.

Extensive team/partnership building with DOE program sponsors, NASA sponsors, sister DOE National laboratories, and university colleagues has been a focus of this LDRD. Erickson has initiated and taken responsibility in contacting and visiting scientists at LANL (January 2003), LLNL (December 2002), and DOE Headquarters [1.25-hour presentation to Patrinos, Elwood, Farrell, Amthor (May 2003)]. As the funding increases on the computer hardware side continue, we are positioned to secure funding from DOE HQ as the climate science budget grows in concert with the hardware applications. King is pursuing collaborations with Jon Foley, University of Wisconsin, author of the IBIS terrestrial biosphere model.

We have, at present, five proposals submitted to NASA, six proposals submitted to DOE HQ, and one submitted to USAID.

Biomolecular "Locks and Keys"—High-Performance Computing for Investigation of Recognition Principles in the Complexes of Biological Macromolecules

A. Gorin,[1] R. Day,[1] V. A. Protopopescu,[1] Y. Li,[1] D. Xu,[2] and M. Mascagni[3]
[1]*Computer Science and Mathematics Division, ORNL*
[2]*Life Sciences Division and University of Missouri*
[3]*Florida State University*

We propose to develop and implement a biomolecular modeling platform with new computational capabilities in three tightly linked areas: (1) modeling of diverse and heterogeneous biological macromolecular complexes; (2) algorithm design that will utilize a new generation of supercomputer architectures; and (3) incorporation of novel types of experimental and database information. None of the planned capabilities currently exist, and successful realization of at least some of them is likely to dramatically benefit the whole field of computational simulations in biology. The developed platform will be prototyped and applied to the study of two problems, which are known to be notoriously complex but of critical importance in modern biological research: structure of membrane protein assemblies and recognition principles in DNA-protein complexes. The proposed effort is complementary to other projects in the Computer Science and Mathematics, Life Sciences, and Environmental Sciences divisions and fits well into the framework for developments in nanotechnology and high-performance computing at ORNL.

The objectives of this project are to design new capabilities in the computational modeling of biological complexes in three areas: (1) biomolecular modeling platform, (2) exploration of Monte-Carlo methods, and (3) incorporation of NMR residual dipolar couplings (RDC) measurements into biomolecular modeling.

Here we report current progress and the FY 2004 research plans, including two additional research tasks. During the first year, we implemented a software library for collective variable algorithms and explored several Markov Chain Monte-Carlo schemas, proposing a new one—the Accelerated Simulated Tempering. In the application area we developed the combinatorial assignment procedure (CAP), which is a new method to utilize RDC for robust automatic assignment and structural characterization of biological complexes.

Task 1. Implementation of the Collective Variable Algorithms

We have developed and implemented (both in C and in Perl) algorithms for constructing molecular models using collective variables. Specifically, we have completed implementation of several types of helical parameters (local, global, symmetric, and linked to the single rotation axis), dihedral angle variables, including pseudo rotation angles, and routines for the manipulation of the rigid sub-fragments (such as secondary-structure elements). Algorithms and subroutines for the force field parameter files (reading, writing, calculating energy components) are progressing well and are expected to be finished by the end of the fiscal year. Plans for the next fiscal year include work on the closuring algorithms and command language for the modeling platform. With the completion of these components (end of FY 2004), we will be in a position to implement our platform on Cray X1 supercomputer (in collaboration with Cray team) and apply it for the studies of protein-protein recognition.

Task 2. Monte Carlo Sampling Methods and HPC implementations

During FY 2003, we have implemented and tested four major types of Markov Chain Monte-Carlo algorithms, namely, (a) Metropolis Monte-Carlo, (b) Simulated Annealing Monte-Carlo, (c) Simulated Tempering and (d) Weighted Monte-Carlo. The most promising results in terms of convergence speed were obtained with Simulated Tempering. Based on our studies, we have proposed a new method—"Accelerated Simulated Tempering Scheme." We construct an alternative ladder that combines temperature and size of step, (t_i, s_i), where $1 = t_1 < t_2 < \ldots < t_m$, and $s_1 < s_2 < \ldots < s_m$. Then, we adjust the original Greyer and Thompson's algorithm by "leaning: the temperature ladder. To this end by set at each ladder asymmetrical transition function, forcing the system to have a lower probability of visiting the higher-level temperature ladders. Plans for the next fiscal year include

an implementation of the conformational stack algorithm (to accumulate together results of the several parallel Monte-Carlo processes) and testing of the implemented algorithms on two real-life biological applications: our molecular modeling platform (models of protein-protein complexes) and Rosetta program (models of single-domain proteins).

Task 3. Integration of the Residual Dipolar Couplings data

Residual Dipolar Couplings (RDC) measurement is one of the most promising new developments in NMR solution techniques. Recently, we have recognized that RDC-based technologies assisted by high-performance Computing can be applied for solving hard problems in structural ribonomics and proteomics. Briefly, the CAP assignment method works in the following way. RDCs are first measured on a target protein molecule. This leads to the assignment of chemical shift values with a corresponding RDC value. Unlike chemical shifts, the magnitude of a given RDC value reflects the average orientation of the bond vector and hence can readily be related to a given structure. For a given assignment guess, the RDC values will either agree or disagree with the given structure. For the ideal case, the best agreement, measured as the root mean square deviation (rmsd) between measured and "best-fit" calculated RDCs, will occur for the correct set of assignments. To enhance this unique mapping, one can also work with a group of RDCs (backbone NH and $C\alpha H\alpha$) that have been correlated using 3D experiments [HNCA, HN(CA)HA]. All possible RDC assignment permutations can therefore be explored for consistency with a given target structure. During next fiscal year, we will apply our algorithm on a wide variety of biological systems including large RNAs, single- and multi-domain proteins.

Toward Common Components for Computational Nanoscience

T. C. Schulthess,[1] M. E. Summers,[2] D. E. Bernholdt,[1] and W. Elwasif[1]
[1]*Computer Science and Mathematics Division*
[2]*Computational Science and Engineering Division*

An open-source repository of computational materials science codes that are portable, extensible, and well maintained on modern high-performance computing architectures is needed to meet the challenges and take advantage of the opportunities posed by nanoscience. The costs of building and maintaining such a repository will be prohibitive, unless new approaches are found to integrate legacy code systems in such a repository. Here we propose to start building a suite of software analysis and modification tools to semiautomatically integrate legacy codes into modern software environments for computational material science. Recent developments of an object-oriented and generic toolset for computational materials science have proven an efficient avenue for building future applications codes for nanoscience. We propose to leverage this fact to develop versions of toolset codes that are compliant with the Common Component Architecture, an emerging approach to component-based software development, which explicitly recognizes the requirements of high-performance computational science codes, including those in nanoscience. It will also demonstrate how the efforts to develop software for computational nanoscience can be integrated with other software projects of the SciDAC initiative.

In order to meet future expectations, computational materials science (CMS) must make efficient use of the most advanced high-performance computing (HPC) infrastructure that is available. The strategy for advancing CMS within ORNL's Center for Computational Science (CCS) is through the participation of domain scientists in the development of computational methods, algorithms, and simulation software and early evaluation of new computer hardware. To enable this, a Materials Research Institute within CCS (CCS-MRI) has been created with the mission to build and maintain a community of leading computational materials scientists and to integrate them into the development of high-performance computing infrastructure of the CCS. One of the tasks of the CCS-MRI is to develop a repository of computational materials science codes that are optimized and maintained on CCS's high-performance computing infrastructure. Besides legacy electronic structure codes, the repository project includes the *ψ-Mag* toolset, a highly extensible generic library of data structures and algorithms that are commonly used in materials simulations. Furthermore, an important part of the SciDAC initiative is the development of a Common Component Architecture (CCA), which aims at providing specification for the development and use of components as basic building blocks of high-performance simulation codes. CCA is designed to make the incorporation of existing code and libraries into components straightforward.

The vision pursued by this project is to combine automated source-code analysis and manipulation tools (SCAMT) with the paradigms of the CCA and the generic software design techniques used in the *ψ-Mag* toolset, in order to build and maintain the software repository of the CCS-MRI. The SCAMT that are developed here will be used for the maintenances of existing codes as well as the transformation of legacy codes into CCA-compliant components. Similarly components of generic libraries, such as *ψ-Mag*, BOOST (www.boost.org), or ALPS (alps.comp-phys.org), will be made CCA compliant. Future software development in CMS that makes use of these tools will be more efficient and well tailored to high-performance computer architectures. Furthermore, the task of integration and maintenance of external codes and libraries into component-based frameworks is representative of widespread issues faced by CCA in other domains as well. At present this project is the only one in the CMS domain using CCA, presenting a unique opportunity to establish ORNL leadership in this area while establishing the groundwork for the development of CMS community codes.

In the first year of this project, we focused on developing the basis for SCAMT and with making the generic *ψ-Mag* tools CCA compliant.

Developing the basis for SCAMT, A F77/F90 Parser and user interface for CMS source code analysis. We followed our original plan and developed an open source

parser for Fortran 77 and Fortran 90 with which abstract-syntax tree (AST) representations of the source codes are built and subsequently used in the analysis tools. The specifications for the AST—there is no commonly accepted standard for ASTs—are being developed along with the analysis tools to optimize the parser systems for the particular requirements of CMS code development and maintenance. Our parser is presently complete enough to generate ASTs for the codes that are presently used in the CCS-MRI (the particular codes we tested are LSMS, SIC-LSD, and LKKR). For scientists who are working on the CMS codes but are not interested in the bare AST, we have developed a Web-based interface which displays source code along with the information provided by the AST in terms of pop-up menus and hyper links between source code segments.With this tool it will be straightforward to isolate parts of legacy codes that are to be used in CCA-compliant components. Making these tools Web-based also simplifies collaborative development of CMS software for the CCS-MRI repository over the internet.

Making the ψ-Mag Tools CCA Compliant. The *ψ-Mag* toolset is representative for numerous generic libraries based on C++ templates that are gaining importance in CMS. CCA uses the Scientific Interface Definition Language (SIDL) to interface with such libraries. The major issue in the present task was to delineate the parts of the *ψ-Mag* toolset that can be mapped onto the CCA-compliant components using SIDL in its present form from those parts that will require modification of SIDL. We have found that it is possible to turn generic classes and functions in the *ψ-Mag* toolset into SIDL objects by (1) creating an instance of the *ψ-Mag* object inside the SIDL implementation and (2) providing a C++ wrapper for each SIDL interface that exactly matches the corresponding *ψ-Mag* concept. Automation of this process appears straightforward, either through a user-generated specification file that extends SIDL to describe C++ typdefs and templates, or through a parser which is able to extract the appropriate information from the C++ source code. However, a wholesale export of *ψ-Mag* into SIDL is not desirable because generic libraries use templates to provide information about objects they implement at compiler time—CCA presently follows the object-oriented model that uses primarily run time polymorphism. We are examining several solutions on how to handle such problems.

During the second year of this project, we will

(1) develop example SCAMT such as an automatic precision conversion tool,
(2) combine components with competing parallelization strategies in one code, and
(3) based on task 1.1, develop CCA-compliant CMS components from legacy codes.

Intelligent Consequence Management for Energy Assurance

J. H. Sorensen,[1] O. Franzese[2], F. Southworth[2], B. Bhaduri[3], and C. Liu[3]
[1]*Environmental Sciences Division*
[2]*Engineering Science and Technology Division*
[3]*Computational Sciences and Engineering Division*

This project proposes to develop a prototype system to demonstrate the feasibility of intelligent consequence management for energy security. An intelligent consequence management system will utilize state-of-the-art sensor technology and wireless communications coupled to real-time simulation and decision support models that will greatly enhance management of crises in rapidly unfolding events. This system will have major impacts on how we can reduce the consequences of threats to our infrastructure. The concept applies to many energy security issues including power disruption from natural threats such as hurricanes and ice storms, energy emergencies such as blackouts, power outages or natural gas disruptions, and terrorist-induced damage against energy facilities or buildings. Specifically we will develop a real-time evacuation model for simulating the evacuation of an area threatened by the release of a hazardous material from an energy system and couple the model with a protective action expert system. This will integrate key ORNL strengths including emergency management, transportation, building technology, sensor technology, simulation, and high-performance computing.

We are developing components of an intelligent consequence management (ICM) system that utilizes state-of-the-art sensor technology and wireless communications coupled to real-time simulation and decision-support models that will greatly enhance management of crises in rapidly unfolding events. Among the properties of intelligent consequence management are the following:

- New sensor networks or links to existing sensor networks designed to detect and monitor the threats of concern.
- High-speed communications and data exchange.
- Real-time simulation models running on high-speed machines.
- Advanced decision support tools that can process data and simulation outputs into a format useful to decision-makers.

The chosen application to demonstrate ICM is a real-time evacuation management system. The problem of determining optimal evacuation paths out of an area at risk can be modeled as a network flow problem. The objective of the network flow evacuation problem is to route a given amount of people from a set of source nodes to a set of exit nodes in the least time, without violating the capacity constraints. Current evacuation models use static traffic assignment and assume the conditions of the network at the beginning of the simulation prevail throughout the evacuation. A better representation of the problem of determining the optimal a priori evacuation paths would be a dynamic network flow model.

A prototype dynamic assignment model was developed at ORNL prior to the development of real-time traffic monitoring for vehicular networks. An experimental application to use traffic data was also developed. This code was integrated with an existing evacuation simulation code ESIM to produce a dynamic evacuation code called DSIM. In each time period of the simulation, new traffic assignments are calculated that are based on the distribution of traffic in the previous time period. Predictions of future traffic behavior are based on current traffic patterns, which are continuously updated at discrete time periods.

The DSIM code more closely simulates driver behavior in congested traffic than a static model. If people encounter stalled traffic, they will take an alternative route based on their knowledge of the area, maps, or traffic guides. Seeking alternative routes that represent underutilized capacity decreases the time needed to evacuate. Initial tests of DSIM show that evacuation times are predicted to be shorter than when modeled with a static assignment. The dynamic assignment makes it feasible to use real-time traffic sensor data to drive the model. The real-time data would be supplied by sensors that assess travel times and flows along the different links on the network as well as queues at chocking points.

Distributed Intrusion Detection and Attack Containment for Organizational Cyber Security

S. G. Batsell,[1] N. Rao,[2] and M. Shankar[1]
[1]*Computational Sciences and Engineering Division*
[2]*Computer Science and Mathematics Division*

The sheer number and sophistication of cyber attacks are making our nation's critical information infrastructure increasingly vulnerable. Important techniques for cyber security include the use of intrusion detection and distributed containment. Most available intrusion-detection schemes are either single host based or local network based and consequently are unable to protect systems and networks effectively when the attacks propagate autonomously in a distributed fashion. The goal of this project is to address the growing demands on intrusion detection and containment by developing an integrated cyber-security framework to identify and contain attacks within large-scale distributed organizational networks. The framework will be distributed, autonomous, and capable of detecting new and rapidly spreading attacks.

Summary

Federal agencies such as the DOE and other large nongovernmental organizations rely heavily on information infrastructures, especially through connection with the Internet to conduct their operations. For example, the entire energy production and transportation industry uses networked computer systems to monitor, control, manage, and barter information and goods. Cyber attacks can significantly impact their operations and potentially create denial-of-service scenarios affecting millions of Americans. A cyber attack jeopardizes the ability to provide services, enforce privacy, and protect information; it undermines the level of trust one has in both the infrastructure and in an agency's ability to perform its mission.

While recent years have seen significant advances in cyber security research, the task of combining diverse detectors of intrusion remains challenging and open from an analytical and foundational perspective. In this project, we employ known results from data fusion to correctly weight the signals of the intrusion sensors based on their detection and miss probabilities. An advantage of the technique of data fusion is that it enables us to build upon and reuse preexisting sensors by incorporating them into new detection schemes without sacrificing accuracy. Following the detection of an intrusion at a host, the system must take preventive steps and distribute this information to other hosts so that they can proactively take action against a spreading intrusion. Containment measures such as closing router ports and isolating compromised hosts have high overheads, thus requiring that fusion and information distribution steps are timely.

In our first year, we completed modules that utilize network-level sensors to trace the physical paths of packets within a domain by using simulated sensors. By building propagation graphs that capture attack paths, and by utilizing graph-search algorithms to perform packet-level content search, we isolate the attack sources. When feasible, such tracing leads us to the attacker. In particular, this method detects if the attacks are from within or from outside the organizational network by suitably collecting the sensor data. We are planning to integrate these methods with actual network sensors in the future.

We have also developed a new approach of fusing intrusion data from diverse sensors to improve detection performance.[1] We determined, through quantitative analysis, the response-time parameters required of the distributed system implementation we are developing. Based on peering and fusing information from distributed sensors of intrusion, the detection and containment architecture can deter rapid self-propagating attacks such as worms. We calculate the real-time constrained responses needed to stifle the spread of an intrusion when the protection is enforced locally within an intranet. Our analysis also suggests the response-time characteristics required in a wider-area distributed network. We are currently developing our research results for further publication and continuing development on the prototype implementation.

Reference

[1]M. Shankar, N. Rao, and S. Batsell, "Fusing Intrusion Data for Detection and Containment," in IEEE *MILCOM 2003*, Boston, October 13–16, 2003.

Image to Intelligence Archive: Intelligent Agent–Based, Large-Scale Spatial-Data Management and Analyses

T. E. Potok,[1] B. L. Bhaduri,[1] K. W. Tobin,[2] P. J. Palathingal,[1] E. A. Bright,[1]
S. S. Gleason,[2] and T. P. Karnowski[2]
[1]*Computational Sciences and Engineering Division*
[2]*Engineering Science and Technology Division*

Enormous volumes of remote sensing image data are being produced on a daily basis throughout the world. This geographical data is analyzed to create scientific, military, and intelligence information. There are two significant challenges that exist in producing this information: (1) managing the vast amount of available data which is steadily increasing and (2) automating the manual processes that are currently needed to produce and search this type of information. We have developed a prototype that addresses these issues by combining novel approaches from three distinct research areas, namely, software agents, georeferenced data modeling, and image analysis.

Image and remote sensing information is of critical importance to the scientific community and for national security purposes. Staggering increases in the volume of remote sensing image data are being produced and made widely available. For example, the United States Geological Survey (USGS) has been archiving enormous volumes of data and faces exponential near- and long-term growth in digital data.[1]

We have focused on two key issues in processing this type of data: (1) managing the vast amount of increasing data and (2) automating the manual processes that are currently used to produce and search this image information. The current work in the field has a strong focus on managing and performing analysis on a closed set of images.[2–4] Many of the proposed techniques do not scale well and do not consider how new images will be found and incorporated into the system. Likewise, the ability for a scientist or an image analyst to effectively use these systems to find geographic information is also limited. What is needed is the ability to select a geographic location for search, then the ability to find regions within that location, and then to find features within that region. For example, an image analyst may be looking over the country of Iraq for regions that contain small cities and then truck convoys within those cities. Developing tools that support this type of analysis is the objective of this project.

We have developed a prototype system that merges capabilities from three separate areas: software agents, georeferenced data modeling, and image analysis. Our system uses software agents to gather new images from previously unknown sources, update image archives using "better" images, and parallelize aspects of the image region and feature extraction process. The georeference data modeling is able to normalize the images placed into the archive, while the image analysis technology provides the ability to segment an image into geographic regions, then within those regions, to find various features. With this system, a scientist or an imagery analyst is able to determine changes in regions of images based on items of interest, for applications such as damage assessment, automatic target recognition, and a wide variety of intelligence, security, and monitoring purposes.

The architecture of our system is an integration of the three components, namely (1) software agent technology, (2) geoconformance modeling, and (3) image analysis. A high-level outline of the architecture is shown in Fig. 1. The agent subsystem is responsible for searching for and retrieving spatial data. Intelligent software agents search for images on the internet in places where they are likely to be found. They store these images in a repository that also maintains metadata about the image in XML format. The images in the repository are modeled and normalized by the geoconformance modeling subsystem. Finally the image analysis module takes the images and generates regions of interest. A graphical user interface (GUI) interacts with the image analysis module to enable users to effectively search and retrieve regions of interest from these vast amounts of spatial data.

With our prototype system, we have demonstrated the ability to manage vast amounts of image data and automate the manual processing of this data. This is a significant step forward in image analysis capabilities. We have presented this novel approach to a number of potential sponsors who are very interested in the concept, including the Department of Defense, International Atomic Energy

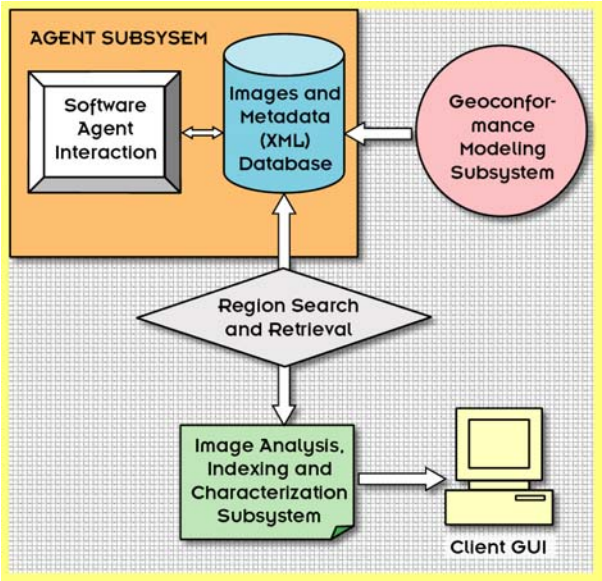

Fig. 1. An outline of the I2IA Architecture.

Agency, and the Intelligence Community. We are planning on furthering the capabilities of the prototype to include a demonstration using ORNL reservation data and to publish the results of our work.

References

[1]United States Geological Survey, "http://vangogh.cse.buffalo.edu:8080/digital/digital/node4.html," *Multimedia Research Group, Department of CSE, SUNY at Buffalo* (2000).

[2]B. Zhu, M. Ramsey, T. D. Ng, H. Chen, and B. Schatz, "Creating a Large-Scale Digital Library for Georeferenced Information," *D-Lib Magazine* **5**, 7/8 (1999).

[3]T. M. Lehmann, B. Wein, J. Dahmen, J. Bredno, F. Vogelsang, and M. Kohnen, "Content-Based Image Retrieval in Medical Applications: A Novel Multi-Step Approach," *in Proceedings SPIE* **3972**, 312–320 (2000).

[4]O. Derrick, S. Raha, J. Kerievsky, and K. Kaila, http://htmlparser.sourceforge.net/.