Parallel and Distributed Computing: Systems and Application Development Infrastructure

PIs: Lori Pollock, Gagan Agrawal, Sandra Carberry, Guang Gao, Ashfaq Khokhar, Errol Lloyd
AIs: Gonzalo Arce, Daniel Chester, Keith Decker, Chandra Kambhamettu, David Saunders, Adarsh Sethi

Department of Computer and Information Sciences
Department of Electrical and Computer Engineering
University of Delaware

1 Introduction

The University of Delaware, marking its 256th year as the only research university in the State of Delaware, is a state-assisted, privately governed institution with approximately 15,000 undergraduate students, 3,000 graduate students, 3,000 students enrolled in credit courses through Continuing Education, and 900 faculty members. The Department of Computer and Information Sciences (CIS) is one of 25 departments in the College of Arts and Sciences, and offers BA, BS, MS, and PhD degrees. There are 16 full-time tenure-track or tenured faculty, 2 visiting faculty, and 5 research faculty in the department. There are 455 CIS undergraduate majors and 85 CIS graduate students. CIS courses are also taken by non-major students both as requirements in their degree programs and to earn a CIS minor. The Department of Electrical and Computer Engineering (ECE) is one of 6 departments in the College of Engineering, and offers a Bachelor of Electrical Engineering, Bachelor of Computer Engineering, MS and PhD degrees. There are 5 full-time tenure-track or tenured faculty in the Computer Engineering portion of ECE, and there are 110 undergraduate Computer Engineering majors, along with 20 Computer Engineering graduate students.

Delaware has a core group of faculty who specialize in compiler optimization, runtime systems, architectural and algorithm design for high performance computing systems. Several faculty work in developing software for improved services on a variety of distributed applications, utilizing the new generation of high-speed network switches. Further, a number of application projects underway at Delaware are aimed at parallel and/or distributed environments. These applications include parallel symbolic analysis, video-image processing, real-time multi-agent systems, gesture recognition, and multimodal applications.

Over the past several years, the faculty research focus has been on theoretical approaches that model the key aspects of their research problems, and on simulation studies that are implemented versions of the theoretical models. These approaches, while extremely valuable in elucidating the fundamental structure of, and approaches to, a problem, cannot include all of the many and varied aspects that arise in practice. In this context, there has been an urgent need for this ongoing research to incorporate realistic and robust experimental studies. The NSF CISE Research Infrastructure Award described here is aimed at establishing a general infrastructure in the University of Delaware CIS and ECE departments for experimental research in parallel and distributed system software tools, and in parallel and distributed applications development. Prior to this grant, there were *no* permanent experimental resources available to CIS and ECE faculty. As a result of the developments of faster workstations and networking switches, the areas of systems, networking, and parallel applications can all share identical infrastructure on which to perform experimental research. In particular, the presence of a local, changeable, crashable computing resource of substantial size is a particular benefit. This grant provides the necessary funds to construct a research infrastructure based on a cluster of high-performance, multiprocessor workstations connected via high-speed networks. Three primary goals of the grant are: (1) For this infrastructure to provide the necessary exclusive access to a dedicated system used to validate and measure the effectiveness of novel compiler, runtime, and architectural design techniques. (2) To allow researchers to load experimental and prototype software by enabling access to crashable resources such as hosts, interfaces, and network switches. (3) For the applications researchers to be able to use the infrastructure as a parallel and distributed computing testbed that functions as the primary processing engine for performing experimental studies.

However, the goals of the project go beyond the expansion of the experimental component of these ongoing research projects. Until recently, faculty focused primarily on individual research projects in their particular subdiscipline. With our current expertise and the recent addition of new faculty, we are now poised to take another major step in research productivity by embarking on large-scale projects that involve multiple subareas of computer science. Faculty in the systems area recognize the importance of working closely with application developers to demonstrate the usefulness of their research and for validating their research in realistic scenarios. Here, the varying demands of parallel *and* distributed applications provide an interesting benchmarking testbed. Likewise, faculty working in application areas will benefit from particular aspects of the systems research projects, with the goal of developing and re-engineering applications that benefit from the effective use of workstation clusters. We are excited by the collaborations and interactions between systems and applications faculty that are beginning to grow out of the infrastructure.

Strong collaborations between faculty in CIS and faculty in ECE have been the norm at Delaware since the award of an NSF equipment grant in the early 80's that established a joint research laboratory. Even though the two departments reside in different colleges, the faculty share in NSF and DoD research grants, co-advise MS/PhD students, and sponsor joint weekly seminars. The result is a thriving environment that is mutually beneficial to CIS and ECE students at all levels. The acquisition of this infrastructure has already begun to further expand and strengthen the CIS and ECE research collaboration, particularly among faculty in the systems area, most of whom have been hired within the past five years (these recent hires are Khokhar and Gao in Computer Engineering, and Agrawal in CIS).

Lastly, the infrastructure will provide a more practical life experience for graduate and undergraduate students working on research in networking, compilers, runtime systems, and architecture. Prior to this grant, there was no laboratory equipment that could be "torn down" for experimental studies on various systems issues. This experience of being able to actually "tinker" with a system of this kind is a valuable and exciting learning experience, that blends research and education.

The official start date for this grant is September 1, 1997, so the current year covered by this report is the third year of the grant. The URL for this project is: http://www.eecis.udel.edu/nsfri.

2 Accomplishments for the Past Year

Our third year goals for this project have focused on: (1) further expanding the research infrastructure laboratory, and (2) expanding existing projects in new directions based on the existence of the new research infrastructure.

Expansion of the research infrastructure laboratory consisted primarily of adding memory and disks to enable research involving large data sets. We purchased and installed enough memory to take 8 of the cluster nodes up to 1024 MB 4-way interleaved from 512 MB with no interleaving. We rearranged the remainder of the nodes to have 512 MB 4-way interleaved instead of no interleaving. An 18GB disk was added to 12 of the nodes, and an 18BG disk was added for home directories.

The National Science Foundation also supported the third year's salary of the staff person hired to set up and maintain the hardware and software of the cluster.

The *expansion of existing projects* based on the existence of the new infrastructure is demonstrated through additional research grants, including two NSF CAREER awards, and the many papers from this RI project. The research described in each of the proposals listed below will involve use of the research infrastructure provided by this grant.

Funded Proposals during Third Grant Year:

- Gagan Agrawal, Joel Saltz, and Alan Sussman, "Compiler and Runtime Support for Data Intensive Computing", NSF, 01/01/00 to 12/31/02, $426,000.
- Guang Gao, "Configurable Off-the-shelf Multithreaded Experimental Testbed (COMET)", NSF-Equip-99.
- Chandra Kambhamettu, "Methods for Nonrigid Motion Analysis," NSF (CAREER award), $273,358, 6/00 to 5/04.
- Chandra Kambhamettu, "Analysis of Tongue Contours from Ultrasound Image Sequences for Speech Disorders," University of Delaware Research Foundation, $29,948, 6/01/00-5/31/01.
- Ashfaq Khokhar, "Multithreaded Algorithms, Models, and Runtime System Tools for Multimedia Applications," NSF (CAREER award), $205,000, 1999-2002.
- Ashfaq Khokhar, "High Performance Data Mining for US ARMY TECOM Data Repositories - PHASE II," US ARMY Test and Evaluation Command, Aberdeen Proving Grounds, $75,000, 1999-2000.
- Lori Pollock, "Problem-based Learning in CIS: Prototype," (parallel computing course), University of Delaware CTE Instructional Improvement Grant, $4250, 9/1/99 - 9/ 1/00.

Some of the projects that are using the equipment include a mapping of fine-grain multithreaded program execution models onto a cluster, a compiler for translating data-intensive applications from a data-parallel dialect of Java to run on a cluster, advanced optimizations for compiling on multithreaded architectures, experimental research on algorithms for large integer matrix computations, testing the scalability of multi-agent systems, debugging and development of parallel algorithms for data mining applications; secure multimedia communications; nonlinear signal processing; and an implementation and study of the performance of a distributed system to track tongue surface from ultrasound images compared to performance of this application on stand-alone workstations.

Here, we highlight some of the projects which have actively used the infrastructure over the past year. The first few projects are really new projects that have just started using the infrastructure in this past year, and either did not do experimental work previously, or are new research directions for the teams. One project has been an experimental study of benchmarks in order to measure the performance of the cluster. As the cluster is used for work which is very computationally intensive, it is important to accurately measure the performance capability of this machine. Based on the results, the conclusion can be drawn that the Myrinet network is far superior to the Fast Ethernet network, as they are implemented and used on the NSFRI cluster. Based on this conclusion, the Myrinet network was chosen to provide the internode communication for the benchmarking of the entire cluster. This benchmarking has led to the generation of a performance graph that summarizes the computational power afforded by the NSFRI cluster. All of these results show that the NSFRI cluster is a very powerful MIMD parallel computer system.

Another project is the development of a compiler which processes data intensive applications written in a data-parallel dialect of Java and executes them on a cluster of workstations with associated disk farms. This project particularly focuses on applications which process and analyze large volumes of disk-resident data. A dialect of Java with data-parallel extensions is used for specifying collections of objects, a parallel for loop, and reduction variables as our source high-level language. The compiler analyzes parallel loops and optimizes the processing of datasets through the use of an existing run-time system, called Active Data Repository (ADR). Loop fission followed by interprocedural static program slicing can be used by the compiler to extract required information for the run-time system. This compiler has been evaluated by comparing the performance of compiler generated code with hand customized ADR code for three templates, from the areas of digital microscopy and scientific simulations. The experimental results show that the performance of compiler generated versions is, on the average, 21% lower, and in all cases within a factor of two, of the performance of hand coded versions. This team is also working on a number of novel optimizations on the runtime system. They have successfully ported the runtime system to utilize multiple processors on the same node.

One research team has been conducting experimental studies of mapping sparse matrix computation on a parallel computing platform based on a fine-grained multithreaded program execution model. Such sparse MVM computation, when parallelized without performing graph partitioning, offers a very high communication to computation ratio, and is well known to have a very limited scalability on traditional distributed memory machines. The particular multithreaded system we use is the EARTH model, which can be implemented from off-the-shelf processors without extensive special hardware support. With the NAS class B input sparse matrix (with 75,000 rows), we observed an absolute speedup of 90 on 120 nodes of a distributed memory configuration. This is achieved without using *inspector/executor* or *graph partitioning*, or any communication minimization phase, which means that same performance can be achieved on adaptive problems as well. High scalability is achieved because of a number of characteristics of the EARTH architecture: asynchronous execution, low communication overheads, ability to overlap communication and computation and low context-switching costs. This experimental study was conducted using a detailed simulator of the EARTH architecture, called SEMi. Extensive simulations performed for the above study required availability of several dedicated high speed workstations. The individual workstations from the NSF Cluster were used for this purpose.

Several projects are now making significant use of the cameras. The project on generating appropriate utterances in negotiation dialogues is videotaping subjects interacting with one another in negotiation subdialogues via the computer and analyzing them. An electroglottograph is also being used to capture intonation in negotiation subdialogues. Currently, this equipment is borrowed, while we determine which model to purchase with our RI grant money this coming year. The cameras also play a critical role in the tongue-tracking research project. This team has implemented on the cluster a system that tracks tongue surface from ultrasound images. Evaluating the performance of numerous runs of this system on the cluster has shown that very high performance can be achieved, which is an important hurdle f or dealing with these images. In addition to continuing the use of the cluster for the research on tongue motion analysis and tracking, the team intends to implement their system for tracking facial movements on the high performance cluster with the object ive of gaining performance to allow for more detailed analysis.

The rest of the projects mentioned here are ongoing projects, which have increased their use of the infrastructure for more substantial experimental studies. One set of projects is the EARTH and EARTH-2000 projects in collaboration with the University of Southern California and Cornell University. This project is ongoing with the implementation and performance evaluation of the mapping of the EARTH multithreaded design to the SUN SMP cluster. A first implementation of the team's scheme (EARTH-SMP) was completed for the SMP cluster. Performance measurements show that inside an SMP node EARTH-SMP outperforms Solaris threads in both thread creation and synchronization. In addition, EARTH-SMP dramatically outperforms EARTH-S, another implementation which does not exploit the shared memory in a node in thread synchronization and communication. Performance results of Fibonacci number, Prefix Sums, Ray Tracing and N-queens have shown the advantage of EARTH-SMP over EARTH-S.

Simultaneously, we have one team developing and implementing compiler optimization techniques in the EARTH-C compiler, which takes C programs and automatically generates Threaded-C programs, removing the difficult task of thread partitioning from the programmer. This team is targeting the issues that arise from compiling irregular codes for a multithreaded architecture which is mapped onto an SMP cluster. At present, the optimization techniques are being tested using a pentium cluster (beowulf design). When the mapping of EARTH to the SMP cluster is more stable, we will begin evaluating the compiler optimizations on the SMP cluster.

One research team has been running numerous matrix codes to compute linear algebra results. The matrix entries are from finite fields in some cases and (exact, bignum) integers in others. The matrices are very large and sparse (e.g. several hundred thousand rows, near a million non-zero entries). The problems solved are matrix rank, minimal polynomial, and integer Smith form. The compute times range from a few minutes to 30 days. The 30 day computations were run sequentially. Some parallel computations last up to 3.5 days (with excellent parallel speedup). The specific matrices are from a mathematician, Volkmar Welker, in Berlin. The methods and codes developed here at Delaware enable the computation, not previously possible, of homology groups of large simplicial complexes, thus advancing the research in homology. These computations have attracted some further interest and, by invitation, the local research group is working on making the code available as a plug-in to GAP, a software system for group theory based at St. Andrews, Scotland. A projected product of this work is a software library called "Linbox", which has the goal of solving sparse linear algebra problems exactly (computer algebra style) by adaptation of numeric iterative methods (but, in some cases using elimination methods). None of the collaborators on this project from Genoble, NCSU, or Western Ontario have a resource like this cluster; the Linbox software system development project would not be possible without this cluster.

Another application which is using the cluster extensively is the development of DECAF. DECAF [Distributed, Environment-Centered Agent Framework] is a Java-based software toolkit for the rapid design, development, and execution of "intelligent" agents to achieve solutions in complex software systems. The cluster is being used to perform extensive testing of DECAF. This includes both internal agent measurements (each agent is multi-threaded) and external, multi-agent system measurements. Initially, this team is interested in architectural overhead, for example, average delays in handling communications under various agent task loads; average delays in resuming suspended tasks; average matchmaker and broker response times; etc. These experiments are being carried out with both synthetic task loads and actual task loads as they become available.

Several products have already resulted from the research performed with this infrastructure: (1) a prototype compiler for a data-parallel dialect of Java to translate data-intensive applications for execution on a cluster (2) a tongue tracking and extraction system from ultrasound images (3) a profiling tool for Threaded-C and EARTH-C multithreaded codes (4) an EARTH multithreading architecture on the SMP cluster (5) a partially developed software library Linbox for linear algebra

Another indication of success of this RI grant thus far is the current involvement of 34 students (primarily graduate students) along with the 13 faculty, in these projects. Eight of the students are undergraduates supported under REU grants. Four of the researchers are female, while two are Hispanic, and one is a person with disabilities. So far, there have been 2 PhD's, 5 master's thesis, and one senior honor's thesis result from the research in the first 2 1/2 years.

3 Evaluation

In the first three years of this grant, we have been very successful in achieving our overall goals for this grant: building the research infrastructure laboratory, increasing collaboration among the participating researchers, and expanding existing projects in new directions based on the existence of the new research infrastructure.

This third year has resulted in several new research projects enabled by this infrastructure which go beyond the original proposed research utilization.

This RI award has already had significant impact on the CIS and ECE departments. Clearly, it has provided the necessary resources for researchers in parallel and distributed computing systems and applications to embark on realistic large-scale experiments to validate and measure the effectiveness of their techniques. Many of the research teams are now using the infrastructure regularly, while others are learning more about parallel and distributing computing and designing experiments and implementing their ideas within this new infrastructure. Further, and very important to the long term success of the CIS and ECE departments, faculty have been able to include significant experimental components in their research proposals, providing them additional leverage in the proposal review process. Researchers who previously have used simulation studies are now designing hands-on experiments that complement their simulation work. Students are exposed to both theory and hands-on experimentation in their research experience.

An increased number of hands-on systems projects and experimental studies have been incorporated into courses and independent studies. An example is a course on high performance computing taught by Ashfaq Khokhar that involves class projects in MPI.

Finally, the RI grant has played an important role as leverage in obtaining additional funding for both educational purposes and research. We now have 4 funded CAREER proposals by investigators involved in this grant, and several other funded NSF research grants which are using this equipment.

15 Selected Publications from Third Grant Year:

- Gonzalo R. Arce, Jose Paredes, and John Mullan, ”Nonlinear Filtering for Image Analysis and Enhancement”, Image and Video Processing Handbook, pp 81-100, May 2000. Edited by Al Bovik, Academic Press, San Diego, CA.
- Sandra Carberry, Chandra Kambhamettu, and Leah Schroeder. Gestural Evidence and the Recognition of Attitude and Intention. *Proceedings of the UM99 Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction*, 1999.
- Bruce Carter, Chuin-Shan Chen, L. Paul Chew, Nikos Chrisochoides, Guang R. Gao, Gerd Heber, Antony R. Ingraffea, Roland Krause, Chris Myers, Demian Nave, Keshav Pingali, Paul Stodghill, Stephen Vavasis, and Paul A. Wawrzynek, “Parallel FEM Simulation of Crack Propagation - Challenges, Status and Perspectives”, Proceedings of the ”Irregular 2000”, April 2000, Lecture Notes in Computer Scien ce.
- J-G. Dumas, B. D. Saunders, and G. Villard Integer Smith Form via the Valence: Experience with Large Sparse Matrices from Homology Accepted for proceedings of ISSAC’00 (International Symposium on Symbolic and Algebraic Computation).
- Renato Ferreira, Gagan Agrawal, and Joel Saltz, “Compiling Object-Oriented Data Intensive Applications”, in Proceedings of International Conference on Supercomputing (ICS), May 2000.
- Graham, McHugh, Mersic, McGeary, Windley, Cleaver, Decker. ”Tools for developing and monitoring agents in distributed multi-agent systems”, in Proc. Agents-2000 Workshop on Agent Infrastructure and Scalability, Intl. Conference on Autonomous Agents, Barcelona, June 2000
- Graham, Decker. ”Towards a distributed, environment-centered agent framework”. In N. Jennings and Y. Lesperance, editors, Intelligent Agents VI: Agent Theories, Architectures, and Languages, pp. 290–304, Springer-Verlag, 2000
- A. Khokhar, E. Albuz, and E. Kocalar, “Scalable Color Image Indexing and Retrieval using Vector Wavelets,” IEEE Transaction on Knowledge and Data Engineering, 1999.
- A. Khokhar, P. Thulasiraman, G. Heber, and G. Gao, “Load Adaptive Algorithms and Implementations for the 2D Discrete Wavelet Transform on Fine-Grain Multithreaded Architectures,” IPPS/SPDP, 1999.
- Jose Paredes and Gonzalo R. Arce. ”Recent developments in stack filtering and smoothing”, Advances in Imaging and Electron Physics. Edited by P. W. Hawkes, Academic Press, Burlington, MA. In preparation.
- Leah Schroeder and Sandra Carberry. Realizing Expressions of Doubt in Collaborative Dialogue. *Proceedings of the 18th International Conference on Computational Linguistics*, August 2000.
- Kevin Theobald, Rishi Kumar, Gagan Agrawal, Gerd Heber, Ruppa Thulasiram, and Guang Gao, “Developing a Communication Intensive Application on the EARTH Multithreaded Architecture”, Europar 2000.
- Ye Zhang and Chandra Kambhamettu, ” Integrated 3D Scene Flow and Structure Recovery from MultiviewImage Sequences ”, IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, June 2000.
- Lin Zhou and Chandra Kambhamettu, ”Hierarchical Structure and Nonrigid Motion Recovery from 2D Monocular Views ”, IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, June 2000.
- Gary M. Zoppetti, Gagan Agrawal, Lori L. Pollock, Jose Nelson Amaral, Xinan Tang, and Guang Gao, “Techniques for Improving Thread Granularity for Dynamic and Irregular Multi-threaded Applications,” ICS’00, May 2000.