COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 01-2	FOR NSF USE ONLY
NSF 00-67 04/02/01	NSF PROPOSAL NUMBER
FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.) OIA - SCI & TECH CTRS (INTEG PTRS)	**0120977**

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
				050299031	

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)	SHOW PREVIOUS AWARD NO. IF THIS IS	IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY?
741109620	☐ A RENEWAL ☐ AN ACCOMPLISHMENT-BASED RENEWAL	YES ☐ NO ☒ IF YES, LIST ACRONYMS(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE	ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE
William Marsh Rice University	William Marsh Rice University 6100 Main Street, MS-16 Houston, TX. 772511892
AWARDEE ORGANIZATION CODE (IF KNOWN) 0036046000	
NAME OF PERFORMING ORGANIZATION, IF DIFFERENT FROM ABOVE	ADDRESS OF PERFORMING ORGANIZATION, IF DIFFERENT, INCLUDING 9 DIGIT ZIP CODE
PERFORMING ORGANIZATION CODE (IF KNOWN)	

IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions) ☐ FOR-PROFIT ORGANIZATION ☐ SMALL BUSINESS ☐ MINORITY BUSINESS ☐ WOMAN-OWNED BUSINESS

TITLE OF PROPOSED PROJECT **Center for Grid Application Development Software \(CGrADS\)**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PREPROPOSAL NO., IF APPLICABLE
$ 19,434,370	60 months	08/01/02	0095508

CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW

- ☐ BEGINNING INVESTIGATOR (GPG I.A)
- ☐ DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C)
- ☐ PROPRIETARY & PRIVILEGED INFORMATION (GPG I.B, II.C.6)
- ☐ NATIONAL ENVIRONMENTAL POLICY ACT (GPG II.C.9)
- ☐ HISTORIC PLACES (GPG II.C.9)
- ☐ SMALL GRANT FOR EXPLOR. RESEARCH (SGER) (GPG II.C.11)
- ☐ VERTEBRATE ANIMALS (GPG II.C.11) IACUC App. Date ____
- ☐ HUMAN SUBJECTS (GPG II.C.11) Exemption Subsection ____ or IRB App. Date ____
- ☐ INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES INVOLVED ____
- ☐ HIGH RESOLUTION GRAPHICS/OTHER GRAPHICS WHERE EXACT COLOR REPRESENTATION IS REQUIRED FOR PROPER INTERPRETATION (GPG I.E.1)

PI/PD DEPARTMENT	PI/PD POSTAL ADDRESS
Department of Computer Science	6100 South Main Street Houston, TX 77005 United States
PI/PD FAX NUMBER 713-348-5136	

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address
PI/PD NAME Ken W Kennedy	Ph.D.	1971	713-348-6009	ken@rice.edu
CO-PI/PD Francine D Berman	Ph.D.	1979	858-534-6195	berman@sdsc.edu
CO-PI/PD Jack J Dongarra	Ph.D.	1980	865-974-8295	dongarra@cs.utk.edu
CO-PI/PD Ian Foster	PhD	1988	773-702-3487	foster@cs.uchicago.edu
CO-PI/PD Daniel A Reed	Ph.D.	1983	217-333-3807	reed@cs.uiuc.edu

Electronic Signature

Received :04/02/2001

CERTIFICATION PAGE

Certification for Principal Investigators and Co-Principal Investigators:

I certify to the best of my knowledge that:

(1) the statements herein (excluding scientific hypotheses and scientific opinions) are true and complete, and
(2) the text and graphics herein as well as any accompanying publications or other documents, unless otherwise indicated, are the original work of the signatories or individuals working under their supervision. I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if an award is made as a result of this proposal.

I understand that the willful provision of false information or concealing a material fact in this proposal or any other communication submitted to NSF is a criminal offense (U.S.Code, Title 18, Section 1001).

Name (Typed)	Signature	Social Security No.*	Date
PI/PD **Ken W Kennedy**	**Signature Not Required**	**SSNs are confidential and are not displayed** *ON FASTLANE SUBMISSIONS*	
Co-PI/PD **Francine D Berman**	**Signature Not Required**		
Co-PI/PD **Jack J Dongarra**	**Signature Not Required**		
Co-PI/PD **Ian Foster**	**Signature Not Required**		
Co-PI/PD **Daniel A Reed**	**Signature Not Required**		

Certification for Authorized Organizational Representative or Individual Applicant:

By signing and submitting this proposal, the individual applicant or the authorized official of the applicant institution is: (1) certifying that statements made herein are true and complete to the best of his/her knowledge; and (2) agreeing to accept the obligation to comply with NSF award terms and conditions if an award is made as a result of this application. Further, the applicant is hereby providing certifications regarding debarment and suspension, drug-free workplace, and lobbying activities (see below), as set forth in Grant Proposal Guide (GPG), NSF 01-2. Willful provision of false information in this application and its supporting documents or in reports required under an ensuring award is a criminal offense (U. S. Code, Title 18, Section 1001).

In addition, if the applicant institution employs more than fifty persons, the authorized official of the applicant institution is certifying that the institution has implemented a written and enforced conflict of interest policy that is consistent with the provisions of Grant Policy Manual Section 510; that to the best of his/her knowledge, all financial disclosures required by that conflict of interest policy have been made; and that all identified conflicts of interest will have been satisfactorily managed, reduced or eliminated prior to the institution's expenditure of any funds under the award, in accordance with the institution's conflict of interest policy. Conflict which cannot be satisfactorily managed, reduced or eliminated must be disclosed to NSF.

Debarment Certification

(If answer "yes", please provide explanation.)

Is the organization or its principals presently debarred, suspended, proposed for debarment, declared ineligible, or voluntarily excluded from covered transactions by any Federal department or agency? Yes ☐ No ☒

Certification Regarding Lobbying

This certification is required for an award of a Federal contract, grant, or cooperative agreement exceeding $100,000 and for an award of a Federal loan or a commitment providing for the United States to insure or guarantee a loan exceeding $150,000.

Certification for Contracts, Grants, Loans and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, ''Disclosure Form to Report Lobbying,'' in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements and that all subrecipients shall certify and disclose accordingly.

This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

AUTHORIZED ORGANIZATIONAL REPRESENTATIVE	SIGNATURE	DATE
NAME/TITLE (TYPED) **Jordan Konisky**		**........**

TELEPHONE NUMBER	ELECTRONIC MAIL ADDRESS	FAX NUMBER
713-348-4002	**konisky@rice.edu**	**713-348-4906**

*SUBMISSION OF SOCIAL SECURITY NUMBERS IS VOLUNTARY AND WILL NOT AFFECT THE ORGANIZATION'S ELIGIBILITY FOR AN AWARD. HOWEVER, THEY ARE AN INTEGRAL PART OF THE INFORMATION SYSTEM AND ASSIST IN PROCESSING THE PROPOSAL. SSN SOLICITED UNDER NSF ACT OF 1950, AS AMENDED.

Received :04/02/2001

3 Project Summary

Advances in networking technologies will soon make it possible to use the global information infrastructure in a qualitatively different way—as a *computational* as well as an information resource. As described in the recent book "The Grid: Blueprint for a New Computing Infrastructure," this "Grid" will connect the nation's computers, databases, instruments, and people in a seamless web of computing and distributed intelligence, that can be used in an on-demand fashion as a problem-solving resource in many fields of human endeavor—and, in particular, for science and engineering.

The availability of Grid resources will give rise to dramatically new classes of applications. Three dominant themes will drive the development of *21st century Grid applications*: computing resources are no longer localized, but distributed, heterogeneous, and dynamic; computation is increasingly sophisticated and multidisciplinary; and computation is integrated into our daily lives, and hence subject to stricter time constraints than at present. The impact of these new applications will be pervasive, ranging from new systems for scientific inquiry, through computing support for crisis management, to the use of ambient computing to enhance personal mobile computing environments.

Although ambitious research programs in ubiquitous computing and Grid middleware are targeting key challenges at the infrastructure level—security, resource discovery, resource management, power management, and the like—*one fundamental problem is not being addressed: the programming of these highly complex and dynamic systems.* Because the Grid is inherently more complex than individual computer systems, programs that execute on the Grid will reflect some of this complexity. Hence, making Grid resources useful and accessible to scientists and engineers will require new software tools that embody major advances in both the theory and practice of computation and program development.

The goal of the *Center for Grid Application Development Software (CGrADS)* is to simplify distributed, on-demand, heterogeneous computing in the same way that the World Wide Web simplified distributed, on-demand access to heterogeneous, remote information sources. CGrADS will explore the scientific and technical problems that must be solved to make the development of effective Grid applications a routine task. To this end, it will conduct research in five key areas:

- programming models, languages, compilers, environments, and tools to support the creation of Grid applications and problem-solving environments;
- mathematical and data structure libraries to facilitate construction of Grid applications, including numerical methods for adaptivity, control of accuracy, and latency tolerance;
- Grid software architectures that facilitate information flow and resource negotiation among applications, libraries, compilers, linkers, and runtime systems;
- base software technologies, such as scheduling, resource discovery, and communication to support development and execution of performance-efficient, adaptive Grid applications; and
- innovative new science and engineering applications that can exploit these new technologies to run effectively in Grid environments.

CGrADS researchers will incorporate research advances into *GrADSoft*, which will become a comprehensive and integrated suite of Grid middleware and tools. They will work systematically with application collaborators to explore, demonstrate, and characterize the effectiveness of the technologies developed in two major research testbeds. The *MicroGrid* testbed provides tools that use a combination of simulation and direct execution to produce a repeatable, observable testbed for Grid experiments. The *MacroGrid* testbed integrates computational and network resources at CGrADS participating sites to provide a realistic, although less configurable, experimental Grid. The MacroGrid will provide a more controlled environment, and a much higher degree of instrumentation and data capture, than is possible in typical Grid environments. These research testbeds will augment the evolving national Grid "cyberinfrastructure," which will be used for full-scale experimentation and demonstration.

The proposed research and education program will be conducted by a team of thirteen principal investigators from Rice University (*Ken Kennedy, Keith Cooper, John Mellor-Crummey, Richard Tapia, and Linda Torczon*), University of California San Diego (*Fran Berman and Andrew Chien*), University of Chicago

(*Ian Foster*), University of Houston (*Lennart Johnsson*), University of Illinois at Urbana-Champaign (*Dan Reed*), University of Southern California Information Sciences Institute (*Carl Kesselman*), and University of Tennessee (*Jack Dongarra and Rich Wolski*). The team we have assembled is uniquely qualified to conduct the proposed research program. It includes internationally-recognized researchers in all the areas of activity within the scope of this proposal, including distributed system services (*Chien*, *Foster*, and *Kesselman*), application scheduling and resource performance forecasting (*Berman* and *Wolski*), performance analysis and monitoring (*Reed*), programming language implementation (*Cooper, Kennedy, Mellor-Crummey*, and *Torczon*) and scientific libraries and programming environments (*Dongarra* and *Johnsson*).

Under the direction of Richard Tapia, CGrADS will initiate a variety of new education programs that build on the past successes of the CGrADS team members in education and human resource development. These programs will address graduate education, undergraduate education, and training for both K–12 teachers and parents. At the graduate level, our efforts will focus on the development of new courses that incorporate Grid-oriented problem solving, and on an active long-term exchange of graduate students between research groups. At the undergraduate level, our efforts will include new courses that bring Grid-enabled computing into the classroom, programs that involve students directly in the research programs of the PIs, and programs to build support communities for student participants to increase retention of those students in technical degree programs. At the K–12 level, our effort will focus on improving teacher training, through improvement and expansion of existing efforts, and on a pilot program in parent education to encourage students, especially under-represented minorities, to pursue careers in science and technology. The pilot parent education program will bring parents together in a safe, non-competitive environment to learn how to use computers, and enable them to better communicate and share computer experiences with their children.

Knowledge transfer in CGrADS will include programs aimed at a broad audience, programs that target narrow segments of the technical community, and programs aimed at students. Broad-based knowledge transfer will focus on an extensive web presence and a newsletter that would be published and widely distributed three times a year. Knowledge transfer to the scientific community will be carried out through two principal mechanisms. First, we will work closely with a group of industrial collaborators to encourage the adoption and standardization of system software technologies that arise from the center's research. In this respect, the work of the Global Grid Forum, as well as the Consortium for Open Grid Software being established by Foster and others, should prove helpful. Second, we will work directly with individual application developers—we have already established collaborations with the developers of Cactus and GriPhyN—and with application groups through the two NSF PACIs, NPACI and the Alliance, and through the NASA IPG, ASCI ASAP programs, and ASCI Institutes. CGrADS investigators are already affiliated with these programs. Knowledge transfer to students will also be an integral part of the education programs described earlier.

The management of the project will follow the successful distributed management model pioneered by the Center for Research on Parallel Computation (CRPC), which included several of the CGrADS principal investigators. Direction of the Center will be the responsibility of the Executive Committee consisting of all the principal investigators, with the advice of a strong and diverse External Advisory Committee and an external Industrial Council consisting of representatives from each of the industrial collaborators. Given the integrated nature of the project and its geographic distribution, the principal investigators will conduct three internal all-hands workshops per year to review and plan ongoing research. These meetings will be augmented by workshops on special technical topics as appropriate. At least one annual workshop will be devoted to interaction with application developers, usually at an all-hands meeting of one of the participating PACIs.

If we are successful, CGrADS will foster research, education, and technology transfer programs that will contribute to revolutionary new ways of utilizing the global information infrastructure as a platform for computation, changing the way scientists, engineers, and business professionals solve their everyday problems.

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.C.

Appendix Items:

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

5 Focus, Research Plans, and Relevance

Imagine remote, biodegradable sensors in the ocean, monitoring temperature, biological materials, and key chemical concentrations, transmitting the measurements via wireless technology to digital libraries of oceanographic data, mining and visualizing this data directly to derive new insights, using the refined data in large scale predictive models, redeploying the sensors to refine the system as a result of the predictions, and finally, triggering nanoactuators to remove inappropriate concentrations of effluent or other non-native materials.

Imagine an earthquake engineering system that integrates "teleobservation" and "teleoperation" to enable researchers to control experimental tools—seismographs, cameras, or robots at remote sites. By combining real-time, remote access to data generated by those tools, along with video and audio feeds, large-scale computing facilities for coupled simulation, data archives, high-performance networks, and structural models, researchers will be able to improve the seismic design of buildings, bridges, utilities, and other infrastructure.

Imagine a personal digital assistant integrated into eyeglasses, powered by body heat, and capable of calling upon ambient computing, information, and network resources so that when you enter a building, your personal information space is available to you, local computing power offloads tasks such as face recognition, translation, and navigation, and you can be simultaneously monitoring your latest earthquake engineering experiment—or your stock portfolio.

21st Century Computing These examples illustrate what we believe will be three dominant themes in 21st Century computing: computing resources are no longer localized, but distributed—and hence heterogeneous and dynamic; computation is increasingly sophisticated and multidisciplinary; and computation is integrated into our daily lives, and hence subject to stricter time constraints than at present.

None of these examples is that far-fetched: revolutionary changes in broadband communications and wireless networking, as well as relentless miniaturization, provide the necessary technical underpinnings. Furthermore, ambitious research programs in ubiquitous computing and Grid middleware are targeting key challenges at the infrastructure level: security, resource discovery, resource management, power management, and the like. However, *existing efforts are not addressing one fundamental problem: the programming of these highly complex and dynamic systems.* This challenging problem is the focus of our proposed *Center for Grid Application Development Software.*

The Grid Programming Problem and CGrADS Our use of the term "Grid" in the title of our proposed center is inspired by a recently published volume entitled "The Grid: Blueprint for a New Computing Infrastructure" [16], which established a compelling vision of a computational and information resource that will change the way that everyone, from scientist and engineer to business professional, teacher, and citizen uses computation [46, 16]. Just as the Internet defines fundamental protocols that ensure uniform and quasi-ubiquitous access to communication, so the Grid will provide uniform access to computation, data, sensors, and other resources. Grid concepts are being pursued aggressively by many groups and are at the heart of major application projects and infrastructure deployment efforts, such as NASA's Information Power Grid (IPG) [28], the NSF PACI's National Technology Grid and Distributed Terascale Facility, the NSF's Grid Physics Network (GriPhyN), and the European Union's European Data Grid and Eurogrid projects. *These and many other groups recognize the tremendous potential of an infrastructure that allows us to conjoin disparate and powerful resources dynamically to meet user needs.*

Despite the flurry of recent work in the area of Grid computing, significant challenges remain for making Grid computing broadly accessible. Few software tools exist. Our understanding of algorithms and methods suitable for Grid environments is extremely limited. Middleware exists, but its suitability for a broad class of applications remains unconfirmed. Impressive applications have been developed, but only by teams of specialists [34, 10, 16, 20, 32, 43, 6]. *Entirely new approaches to software development and programming will be required for the potential of Grid computing to be realized.*

It is this combination of a fundamentally important problem and a need for multidisciplinary research advances that leads us to propose a project of broad scope and significant duration in this area:

> The goal of the *Center for Grid Application Development Software* (CGrADS) is to conduct fundamental research leading to the development, in prototype form, of technologies needed to make the Grid usable as a computationally-rich, problem-solving system.

Collectively, the challenges that must be overcome to achieve this goal can be summarized in a single requirement: *We need application development technologies that make it easy to construct and execute applications with reliable [and often high] performance in the constantly-changing environment of the Grid.*

As we pursue this goal, we can draw upon a significant body of knowledge and technology in distributed computing, the Internet, and Grid middleware. However, while traditional distributed computing technologies provide critical building blocks and frameworks for Grid application development, distributed computing is not concerned with, and does not address the large-scale and dynamic resource sharing, frequently stringent performance requirements, large resource needs, and the multidisciplinary nature of Grid applications. Although emerging Internet, peer-to-peer, and Grid middleware technologies are meeting the need for large-scale resource sharing, they do nothing to simplify application development.

CGrADS will develop the knowledge and technology base required to support application execution in this new computing environment, along with application development strategies to make it accessible to ordinary scientists, engineers, and software developers for problem solving. To do this, we will pursue research in four major areas: (1) collaboration on the design and implementation in prototype form of important scientific applications for the Grid; (2) the design of programming systems and problem-solving environments that support the development of configurable Grid applications by end users in high-level languages close to the notation of their application domain; (3) the design and implementation of execution environments that dynamically match configurable applications to available resources in order to provide consistent, reliable performance; and (4) the design and construction of hardware and software testbeds for experimentation with the GrADS program preparation and execution system and the applications developed to use them.

We anticipate that the successful completion of this research program will lead to revolutionary new ways of utilizing the global information infrastructure as a platform for computation, data sharing, and collaboration.

5.1 Applications

Just as the emergence and usability of the World Wide Web has ushered in new paradigms in application development and access to information, the maturing of the Grid and its natural extension to peer-to-peer platforms, wireless endpoints, remote instruments, and sensors will engender innovative new application paradigms and new environments for application development and execution. Such environments will support application adaptivity, portability, ubiquity and performance. Emerging Grid applications will provide the driving force behind the architecture, research and prototypes that will be developed by researchers in the Center for Grid Application Development Software. Over the next decade, Grid applications will address a wide variety of critical challenges in science and engineering. New applications in computational biology, bioinformatics, genomics, high energy physics, crisis management, and other domains will require real-time data collection, mining and analysis, simulation, and visualization of results.

There are several key challenges that must be addressed for Grid computing to be effective. Applications must be able to nimbly adapt to a dynamic set of target resources and to incorporate huge amounts of information from heterogeneous sources and distant endpoints (sensors, target computational resources on peer-to-peer networks). Moreover, the Grid software infrastructure to which the applications themselves are targeted is complex, heterogeneous, and dynamic. Globus, NetSolve [2], Condor [48], Legion [22], and commercial infrastructure systems have different levels of robustness and operate on different resource subsets. Over the next ten years, applications will need to be able to adapt and perform with respect to the infrastructure provided by ambient resources. The design of development environments and run-time systems for such adaptable and "ultra-portable" applications constitutes an extremely challenging and comprehensive set of problems.

The success of the Grid as a computing platform is dependent on development of performance-efficient applications that can effectively exploit a wide range of cooperating resources. Software that supports development and execution of such applications is critical to making Grid programming tractable. During the research associated with the Center for Grid Application Development Software, a collection of emerging Grid-enabled applications will focus our research goals, and help set research priorities. These applications provide a means for critical evaluation and assessment of the Grid application development software resulting from our research. The following examples are representative of major classes of a new generation of Grid applications.

On-Demand Applications A critical aspect of computational Grids is their ability to concentrate the massive computational and information resources required for real-time, on-demand applications. To understand the importance of on-demand application development for the Grid, consider the problem faced by a crisis manager after a major disaster such as an earthquake. Although the component operations that are essential to crisis management are known in advance, each crisis presents unique requirements. The crisis manager must integrate information from many different sources to determine the actions needed to respond to the particular crisis at hand. For example, she must be able to understand the state of the current infrastructure, possibly by aggregating sensors in buildings, power lines, and utility conduits into a network that can report the changing state of the basic infrastructure. She must be able to integrate reports from emergency crews with patient information to ensure that emergency treatments are consistent with the needs of each patient. Finally, she needs to be able to access mesoscale weather models, fueled by information from a grid of Doppler radars, to identify weather patterns that may exacerbate the crisis. There may also be a need to simulate the flow of groundwater contaminants through the soil.

Ubiquitous Applications During the next decade, an increasing number of users will develop applications for execution on a platform where the user does not know (or care) where the application might be executed. Current examples of such software platforms include SETI@home and Entropia (which target largely embarrassingly parallel applications to compute on "throw away" endpoints), Condor (which targets individual, migratable and largely embarrassingly parallel jobs on workstation clusters) and APST (middleware that targets parameter sweep applications on a wide variety of grid environments). Such systems demonstrate the potential for the Grid, but the research community must improve application programming models for Grid execution. As part of CGrADS research, we intend to develop more sophisticated (and dependent) programming models that can execute ubiquitously and reliably in large-scale Grid environments.

Robust, Portable Applications Much as the World-Wide Web has catalyzed the creation of immense collections of private data, and supports easy accessibility to large quantities of public data, we anticipate that the emergence of the Grid will spawn public Grid resources (computing, storage, etc.). Already, we see significant development underway for production Grid systems (e.g., the NASA IPG, NSF GriPhyN and NEESgrid, and European Data Grid) based on a existing software infrastructures (e.g., Globus in the four examples just cited). To capitalize on such resources, a new generation of portable, Grid-aware applications is needed.

The user community's desire to exploit Grid resources and to protect their software development investment is important driver for developing portable, standard services that support robust Grid applications. In the long run, convincing the developers of Grid applications—users, scientists, and third party commercial organizations (ISV's)—must depend on the development of standard interfaces for critical Grid services that enable both *portability* forward to new software infrastructures and platforms and *access* to very large numbers of resources. However, current Grid software infrastructures lack the capabilities to support flexible, robust Grid applications in a world of heterogeneous systems, unreliable networks, and asynchronous resource revocation. The core research proposed under the aegis of this center includes support for adaptivity, resource negotiation, and performance contracts. These capabilities will help applications operate effectively in an ever-changing Grid environment. In addition, the proposed research will develop the understanding that enables definition of standard shared libraries that export these capabilities.

Integrated Data Analysis and Simulation Data-oriented applications will constitute one of the most active and critical areas for the next decade in science and engineering. Many research communities collect, analyze, and mine immense amounts of data in collections that are often not co-located with the computational servers. For example, there is considerable effort currently being devoted to the development of parallel and distributed applications that use genomic data to assess, evaluate and develop structural models and to answer fundamental questions about life.

In addition, the GriPhyN [23, 5] project is developing a distributed analysis environment for physics experiments that will serve thousands of users. A crucial concept being pioneered by GriPhyN is virtual data, i.e., derived data products that are defined by the computations to produce them. Given a set of virtual data definitions, a user request for data value(s) can be translated into computations and data movements. We anticipate collaboration with GriPhyN in two areas: estimation of the computational requirements of virtual data computations, along with scheduling of computations and data movement based on compiler-detected query profiles.

Another important area in which integrated data assimilation from distributed resources is becoming more important is in weather forecasting, such as for instance the Integrated Forecasting System (IFS) developed by European Center for Medium Range Weather Forecasting (ECMWF) and in climate analysis such as the ERA-40 project covering the time period from 1957 - 2001 pursued jointly between NCAR, NOAA, NESDIS, ECMWF and several other organizations. The IFS has real-time aspects and uses a wide range of sensor and network technologies and is an excellent application for GrADSoft technologies.

The Center for Grid Application Development Software will use both developing applications and mature Grid exemplar codes to guide our design and development efforts. We will work with developers of exemplar applications in each of these application classes to prototype program development software that meets the needs of current Grid applications as well as the new generation of applications that evolve to reap the benefits of the Grid. During our research into application development software, we expect to gain new insights into how to design and implement grid applications. Thus, our research agenda includes the study of new types of applications, as well as new approaches to application design and implementation.

5.2 Vision for Software Development

For the Grid to become a really useful computational environment—one that will be routinely employed by ordinary scientists, engineers, and other problem solvers—it must be relatively easy to develop new applications. Currently applications must be developed atop existing software infrastructures, such as Globus, by developers who are experts on Grid software implementation. Although many useful applications have been produced this way, it is too difficult for Grid computing to achieve widespread acceptance.

In our vision, the end user should be able to specify applications in high-level, domain-specific problem-solving languages and expect these applications to seamlessly access the Grid to find required resources when needed. In these environments, users would be free to concentrate on how to solve a problem rather than on how to map a solution onto the available Grid resources.

To realize this vision we must solve two fundamental technical problems. First, we must understand how to build programming interfaces that insulate the end user from the underlying complexity of the Grid execution environment without sacrificing application execution efficiency. Second, we must provide an execution environment that automatically adapts the application to the dynamically-changing resources of the Grid. Our overall approach to addressing these challenges is described in the next section.

5.3 The GrADS Approach

CGrADS will address the fundamental challenge of program development for Grid environments via a coordinated and far-reaching program of research, prototyping, and technology transfer aimed at the central problems of programming models, algorithms, programming systems, and applications.

Underlying and unifying our diverse investigations is a basic assumption: that *effective application development for Grid environments requires a new approach to the design, implementation, execution, and optimization of applications.* A new strategy is needed because the traditional development cycle of separate code, compile, link, and execute stages assumes that the properties of underlying resources are static and relatively simple. In the Grid, this assumption is not valid. (Needless to say, the alternative approach, frequently adopted in distributed computing, of handcoding applications with socket calls or remote procedure calls is not viable either.) *We require a software development environment that enables the effects of dynamism to be mitigated and controlled.*

Figure 1 presents the new program development structure that we believe is required. In what we refer to as the GrADSoft architecture, the discrete steps of application creation, compilation, execution, and post-mortem analysis are replaced with a continuous process of adapting applications to a changing Grid and to a specific problem instance. Two key concepts are critical to the working of this system. First, an application must be encapsulated as a *configurable object program*, which can be optimized rapidly for execution on a specific collection of Grid resources. Second, the system relies upon *performance contracts* that specify the expected performance of modules as a function of available resources. Our research program will target the various elements of this architecture. The rest of this section summarizes the key ideas; the rest of the proposal, explains, in detail, the technical challenges to be addressed and the approach to be followed in each area.

GrADS Program Preparation System. The left side of Figure 1 depicts the tools used to construct configurable object programs. We expect that most application developers will use high-level problem solving

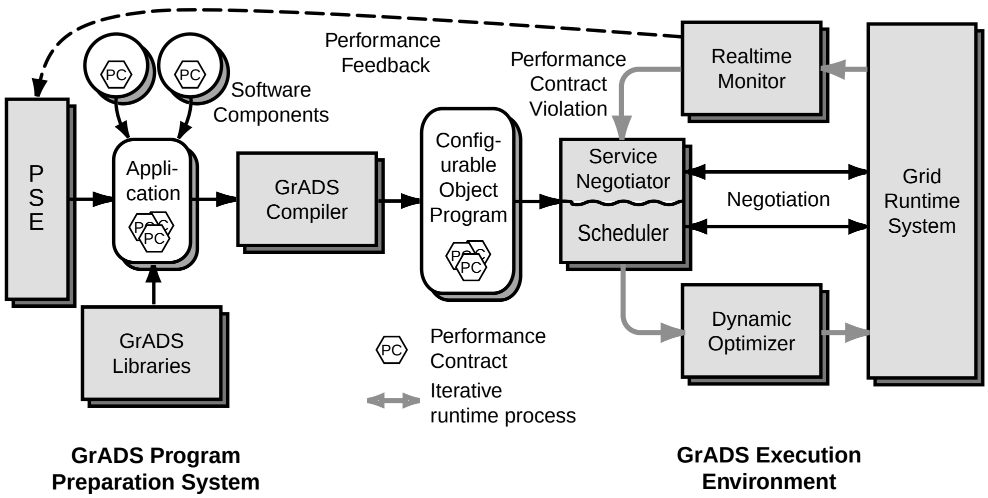

Figure 1: GrADS Program Preparation and Execution Architecture

environments (PSEs) to assemble Grid applications from a toolkit of domain-specific components. Another path allows developers to build the specialized components that form these PSE toolkits (*e.g.,* a library for solving PDEs on computational grids) or to create new modules for their specific problem domain.

In either scenario, modules are written in derivatives of standard languages with Grid-specific extensions (*e.g.,* data or task distribution primitives). They are bound together into larger components, libraries, and applications with a coordination language. This process creates malleable modules, annotated with information about their resource needs and predicted performance for a wide variety of resource configurations.

The goal is to build tools that free the user from many of the low-level concerns that arise in programming for the Grid today, and to permit the user to focus on high-level design and performance tuning for the heterogeneous distributed computing environment.

GrADS Execution Environment. When a configurable object program is delivered to the execution environment, the GrADSoft infrastructure must first determine what resources are available and secure an appropriate subset for the application. Using annotations from performance contracts and results from compiler analysis, service negotiators will broker the allocation and scheduling of module components on Grid resources. Next, the infrastructure will invoke the dynamic optimizer to tailor the reconfigurable object program for good performance with the available resources. This step will also insert sensors and actuators to help the performance monitoring system control application execution.

During program execution, a real-time monitor tracks program behavior and validates observed behavior against performance contract guarantees. Should a performance contract be violated, the monitor will respond by interrupting execution through an actuator, leading to several possible actions. The actuator can invoke the dynamic optimizer with more information (from performance monitoring) to improve program behavior in the current execution context, negotiate a new execution context where the existing executable is more likely to satisfy the old contract, or do both by negotiating a new context and tailoring the executable for it. Dynamic forecasts of resource performance and Grid capacity will be used to reduce renegotiation overhead. The goal of this closed loop system is to ensure that execution of the application proceeds reliably, meeting the specifications of its performance contracts, in the constantly changing Grid environment.

5.4 Program Preparation System

Developing a parallel program for efficient execution on the Grid currently requires a level of expertise that few possess. Unless Grid programming can be greatly simplified, the power of Grid computing will be inaccessible to many. CGrADS research on program preparation systems will focus on the design and construction of software that simplifies building and running Grid-enabled applications.

To simplify development of Grid-enabled applications, we propose that most users will work in a domain-specific language, composing pre-existing components. This approach abstracts away the Grid-level details and lets the user express the computation in terms that make sense to an expert in the application domain.

Underneath the domain-specific language, and supporting it, will be a layer of software that manages the complex task of computing on the Grid. This software, embedded in a collection of libraries, will include not only the base algorithms, but also composable performance models and dynamic mapping strategies for each method. To manage heterogeneity and the late binding of resources without sacrificing performance, we will design and build a dynamic optimizer that does load-time code optimization. It will also insert the sensors and actuators needed by the runtime system.

Ultimately, we believe that developers will want to construct Grid-enabled applications by using high-level scripting languages to compose Grid-enabled components written by experts. For this approach to succeed, we must address the following scientific and technical problems: (1) developing programming models and compiler technology to support efficient high-level programming; (2) developing programming models that help library writers cope with properties of the Grid such as variation in latency and performance, or even failure; (3) designing composable performance models for use in selecting resources, in tailoring code to runtime resources and to detect performance problems; (4) using partial evaluation in the GrADS compiler to support rapid runtime tailoring for efficient execution; and (5) understanding the impact of essential activities such as checkpointing, reporting, and monitoring on overall performance and devising strategies to mitigate these effects. These problems cannot be solved at a chalkboard. Finding the appropriate solutions will take extensive experimentation and exploration, in the context of the CGrADS applications effort. By repeatedly moving solutions into prototype tools and using these tools in the next generation of applications, we will refine our approaches and improve both their effectiveness and usability.

A Framework for Grid Application Development A Grid programming system should make it easy for end users to build applications that execute efficiently on the Grid. Such a systems should provide several ways to construct applications. We expect that the most common approach will be to compose applications from pre-written, domain-specific, library components as is done with CCAT [21], Khoros [31], SciRun [2] and NetSolve [2]. These systems allow users to "script" an application by configuring and composing software components or services that run elsewhere into a single distributed application. Scripts are either composed graphically, or they are written using a high level scripting language like Python [33] or Matlab [26].

To bridge the gap between these comfortable, high-level, scripting languages and an efficient, Grid-enabled executable, we must develop a novel and effective compilation system. Our strategy has two novel components: an implementation technique for the domain-specific languages that we call *telescoping languages* [30], and a tool for load-time tailoring that we call the *dynamic optimizer*.

Telescoping Languages The telescoping languages approach makes extensive use of whole-program analysis and optimization to automate the construction of extensible languages. We will build a system called the *TeleGen* compiler, shown in Figure 2. TeleGen will read an annotated, domain-specific library, analyze it, and produce a customized optimizer that understands the library entry points (including their execution properties) as if they were native primitives in the base language. It will also create a version of the library that includes optimized versions of some entries. TeleGen's compilation approach builds upon work in high-level axiom-driven optimization [36, 35], call-site analysis and library routine implementation selection [25], and interprocedural optimization of high-level languages [11, 4].

The resulting optimizer behaves as a compiler for a new language—the base language (*e.g.,* C) augmented with the functions of the domain-specific language. In this scheme, a domain-specific scripting language can be implemented as a preprocessor that translates scripts into base language programs that call the library components. The optimizer should produce highly optimized code from such an input. This implementation strategy can be applied iteratively to several different levels of libraries—telescoping them into one translator.

The success of the telescoping languages strategy depends upon the existence of sophisticated component libraries that are specially prepared for Grid execution by professional library developers. (The challenges of developing such libraries are detailed below.) These libraries will be annotated by the developers with mapping strategies, performance models, hints on how to optimize calls to individual components in various contexts, and algebraic specifications of library operations (*e.g.*, commutativity, transitivity, associativity [52]) to facilitate high-level optimization.

A critical issue for this work is constructing TeleGen so that the optimizers it generates can, themselves, produce configurable object programs suitable for use in the GrADS execution system. The optimizer must understand all of the components of a library—code, performance model, and mapping strategy—and must manipulate them to create the configurable object program.

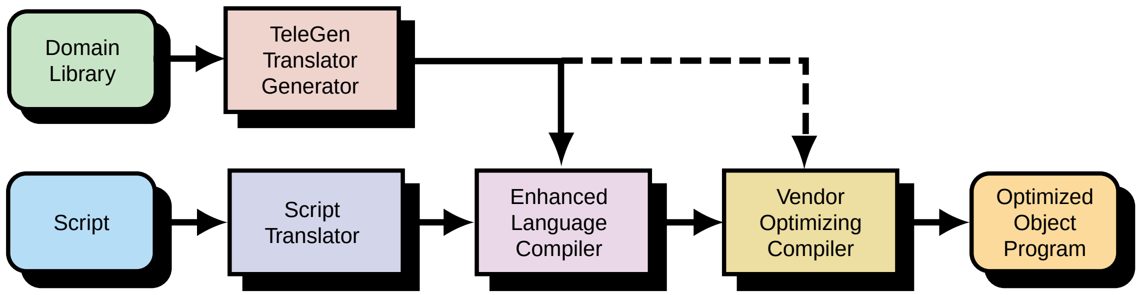

Figure 2: Telescoping Languages

Research issues include the design of high-level optimizers for the Grid, methods for selecting the right code variants for a given collection of Grid resources, mechanisms for generating and managing the myriad variants that the system will need, and the design of tools to help the library designer build useful library annotations.

Dynamic Optimizer The dynamic optimizer is a component of the program preparation system that lives in the execution environment. It is invoked at load time to tailor the configurable object program to the actual runtime environment. The dynamic optimizer queries the target machine for configuration data, inserts the sensors and actuators needed by the runtime system, and rewrites the object program into an executable that will run efficiently on the target machine. Deferring code generation into the dynamic optimizer, should simplify configurable object programs, reduce their size, and provide more consistent optimization.

Libraries and Algorithms Modeling, simulation, and data intensive computing have become staples of scientific research. This has exposed the difficult aspects of scientific computing to a broader audience of scientists and engineers. While access to computing has improved dramatically over the past decade, efficient scientific computing still requires specialized knowledge in numerical analysis, computer architectures, and programming languages. Many working researchers do not have the time, the energy, or the inclination to acquire such expertise. Scientists expect their computing tools to serve them, not the other way around. Unfortunately, the growing desire to tackle interdisciplinary problems with more realistic simulations on increasingly complex computing platforms, will only exacerbate the problem. The classic solution to this problem was to encode the requisite expertise into easily used libraries. While traditional numerical libraries (e.g. LAPACK, Ellpack, ScaLAPACK, PETSc) have brought immense benefits, the radical changes occurring in scientific computing are creating new challenges that these libraries, in their current form, cannot meet. that despite the immense benefits such traditional numerical libraries

To address this challenge, we will develop a new generation of *Self-adapting Numerical Software (SaNS)* systems. These SaNS will not only meet today's challenges; they should address future changes in scientific computing as well. We will design and build a framework for SaNS for numerical libraries and algorithms. This system will operate as "black box" software; intended for use by domain scientists who need not understand the algorithmic and programming complexities it encapsulates. To manage the complexities of the Grid and to adapt in ways that maximizes their effectiveness, SaNS must encapsulate far more intelligence than standard libraries. The work described below will make it possible to produce a SaNS system that: (1) automatically analyzes the logical and numerical structure of the data to let the library choose the best algorithmic strategy for solving the problem; (2) embodies a set of rules, progressively self-tuned over time, for choosing the appropriate algorithm for a given linear system, based on its analysis of the data and any hints provided by the user. (3) encodes metadata about the user's data, about its own characteristics, and about the known implementations of the algorithm it selects, so that the system can schedule the computation effectively on the available resources. (4) uses a scripting language that generalizes the decision procedure that the SaNS follows and enable scientific programmers to easily make use of it.

Using SaNS libraries should improve the ability of computational scientists to solve challenging problems efficiently—without requiring much extra-domain expertise. As these innovations become generally available, they will create a dynamic computational environment that automatically selects and integrates the most effective library components for a given problem, data set, and collection of resources. The SaNS metadata scheme will let us capture this self-adaptive process in databases, creating an indispensable resource for future library developers. Current numerical libraries, whose limitations are increasingly obvious, are now threatened with obsolescence. This investigation will lay the foundation needed to meet the challenging demands of computational science over the next decade.

5.5 An Execution Environment for Grid Applications

As we explained above, our research seeks to allow future Grid applications to operate in highly dynamic environments, adapting their resource demands and behavior to the environments in which they find themselves—and also, when possible, adapting the environment to fit their requirements.

The realization of this overall goal requires the development of new mechanisms for information and control flow between program preparation system, program, and environment, so that (1) information about the environment, and program behavior in that environment, can be discovered and communicated to program components in meaningful terms, and (2) program requirements can be communicated to the environment, and to program components, in ways that admits to effective control.

These two goals define, collectively, the purpose of the GrADSoft execution environment. They lead us to focus our research in this area on three key issues, namely the protocols, services, and methods required to (1) discover and disseminate information about the dynamically changing structure and state of Grid resources (*Grid information service*); select, allocate, and control collections of Grid resources, and communicate requirements among resource providers and consumers (*resource management service*); and (3) monitor and, as necessary, control adaptively an executing program. These protocols and services represent what is sometimes called middleware [1]: code that executes in the network in support of applications. As in other areas of the Grid, we are concerned with achieving a separation of concerns between *resource protocols* that must be broadly deployed and *collective protocols and services* that can be localized in more application-specific code [19].

The following scenario illustrates some of the issues that arise in the Execution Environment. We imagine the Program Preparation System generating an executable image, a set of performance requirements, and a budget for executing the application that is expressed in some Grid currency. These latter two abstractions serve as the basis of a "performance contract" between the application and the resources it uses. The Execution Environment then launches the program by submitting it to the *application monitoring and adaptive control service* (service (3)). To do so, this system consults the *Grid information service* (service (1)) to determine what resources are available and appropriate, and the *resource management service* (service (2)) to ensure that those resources are allocated for the execution, subject to demand and supply respectively.

In pursuing these goals, we will work initially within the context of the Grid architecture defined by the Globus Toolkit [17], due its widespread adoption within the scientific community, and significant experience and code base. At the Connectivity and Resource levels in a Grid architecture, the Globus Toolkit defines standard authentication and authorization protocols, information service protocols, and resource management protocols. At the Collective level, it provides resource discovery, brokering, and co-allocation functions. (Other relevant protocols and services are being discussed within the Global Grid Forum, for example for event delivery.) The adoption of this framework allows us to focus our attention on the central problems (for us) of how to obtain, organize, and exploit monitoring and control information, problems that can be expressed in terms of interactions among cooperating services and resources. Issues of security, resource access, and the like can be relegated to Globus—and/or to complementary industrial standards and trends such as Jini [50] and the emerging peer-to-peer technology base (as is being pioneered by companies such as Entropia, CDDB, Napster, and Parabon [14, 3, 37, 39]). The result of this work will be the definition of both a middleware architecture and specific new middleware services designed to support the concerns of adaptive Grid computations. We expect that this work will result in useful feedback to the Grid protocols and services R&D community.

In the following, we expand upon each of the three points noted above, indicating in each case the nature of the primary research challenges.

Grid Information Service To provide the functionality needed for negotiation and scheduling, the Execution Environment must be able to obtain information about the resources available for application use. A wide variety of information can conceivably be of interest: for example, hardware configuration, measured load, access control policies, application performance data, power consumption [38], estimates of nonobservable system properties, and predictions of future states [53, 12, 45, 13, 38, 29]. Our goals in the CGrADS project is to first to develop an integrated framework in which these many different types of information can be used in a coordinated and uniform fashion, second to conduct a broad exploration of how different sorts of information can be produced and used, and hence, third, to produce a set of effective techniques for information collection, analysis, and application.

In previous work, we have established frameworks for providing uniform access to, and indexing of, diverse information sources (the Globus MDS [15, 7], for collecting experimental data and using this data to generate forecasts of future state (the Network Weather Service: NWS [53]), and for structuring networks of sensors and transformers to support adaptive control (Autopilot [41]). Each of these systems has been proven effective in various experimental and (in some cases, e.g., MDS), large-scale deployments. We will build on this infrastructure in this work, integrating these diverse elements and extending them in major ways. We will develop new services, including distributed event management, new methods of measuring and predicting components of system state, methods for discovering and maintaining relevant information about resources of interest in the execution environment, robust and scalable publication methods, methods that can deal effectively with both measured and dynamically derived data, and methods for information service discovery in widely distributed, dynamic environments [49, 9, 24, 27]. We will also address the question of how to represent our degree of confidence in data and security concerns relating to dissemination of data.

Grid Resource Management Service The Execution Environment must also provide the ability to reserve, allocate, configure, and manage collections of resources that match an application's needs. Building on elements of the Globus resource management architecture [8], which provides secure remote access and reservation [18] mechanisms, we will develop new co-reservation and co-allocation algorithms capable of dealing with resources with dynamic and probabilistic properties, integrate performance contracts (see next paragraph) into resource reservation and resource operations, integrate traditional quality of service methods into resource management frameworks, and map compiler-derived and library-derived performance information into global resource reservation and allocation services.

A major goal of our work in this area will be to explore and understand the nature of the language that should be used to share complex, multi-dimensional requirements and performance data among resource providers and consumers. We will investigate the design of a language of *performance contracts* to enable dynamic negotiation among resource providers and consumers. A performance contract maps a set of resources and a set of application resource needs to a specified performance level—to satisfy the contract, the assigned resources and the application must behave as specified.

Our approach to performance contracts derives them from a performance model provided by the configurable object program and a set of resource performance characteristics culled from the Grid information service. The *service negotiator* (which is logically part of the *application monitoring and adaptive control service*) brokers performance contracts between applications and resources. It uses the information and reservation services to find available resources, select a set that matches the predicted needs of the application, and make any needed reservations. As a part of our research, we will develop a theory of performance contracts and service negotiation that can be adapted to the varying behavior of the Grid. Matchmaking techniques may be relevant here citeMatchMaker; see also [47].

Building on this framework, we will investigate more dynamic resource brokering mechanisms based on the use of economic models (*e.g.,* bidding, cost negotiation, and dynamic pricing) as a basis for arbitrating between competing resource demands. We will study both auction-based and commodity-based formulations of the performance economies. Auction-based systems are attractive because of their scalability, but it can be shown that commodity-based (but not auction-based) economies achieve both equilibrium and stability [51, 44]. Since Grid applications must adapt to changing performance conditions, overall system stability will be an important concern.

Application Monitoring and Adaptive Control Service Work in the two areas just listed will provide a powerful, extensible framework for communicating requirements, various information, and control functions among applications, intermediate brokering functions, and resources. The third area in which we will work builds on this framework to construct a a closed-loop control system that uses various dynamic performance information sources to guide an application to completion despite performance variations in the underlying resource base, via a process of adaptive control of both application behavior and resource demands.

To enable such adaptation, the execution monitor depends on the *dynamic optimizer* (which will be developed in conjunction with the Program Preparation System) to insert the sensors and actuators that let it manage the execution. The dynamic optimizer, invoked just prior to execution, also instantiates the final performance contract according to the rules of the resource economy that is in place.

The Autopilot system [42, 40, 41] embodies several of the ideas on which we will build our distributed monitoring systems. Atop this substrate, the key research issue is developing techniques to decide how and when a performance contract has been violated (*e.g.,* managing temporal variation and distributed contract testing) and how to respond to the violation in order to maximize application performance. To carry out this plan, we need new strategies that let the compiler, scheduler, runtime system, and other components cooperate to extract, non-intrusively, pertinent information from the running application.

5.6 Understanding Grid Software Behavior

The long-term success of our Grid software research agenda requires that we develop design methodologies that allow systematic design and evaluation of dependable, robust, and scalable Grid services and applications software. Unfortunately, such design methodologies are currently totally lacking. It is no exaggeration to say that Grid services and software are designed and characterized today largely based on the designer's intuition and on ad hoc experimentation with little knowledge of *when* they will fail catastrophically. We view this as completely unsatisfactory and adopt as our long-term research goal the development an experimental methodology for characterizing grid software that allows us to evaluate and predict the performance, fault tolerance, and scalability of middleware services.

As an important first step towards the development of such design methodologies, we will develop and deploy two major testbeds and associated tool suites designed to provide both soft (configurable) and hard (fixed) environments for exploring dynamic Grid behaviors. Our goal in this work is to enable systematic study and ultimately understanding of the dynamic behavior of Grid resources, middleware, and applications.

These tool testbeds and tool suites are what we call the *MicroGrid* and *MacroGrid* testbeds. Both share in common the use of Globus services as a unifying computational environment. They differ in terms of the degree of configurability and realism they offer. The use of a common environment means that programs can be run without change on both testbeds, hence allowing comparative studies.

The MicroGrid testbed provides tools that use a combination of simulation and direct execution techniques to produce a repeatable, observable testbed for Grid experiments. Major challenges here include:

- *Fidelity in Grid Resource Modeling* The modeling of computation, storage, and networking resources faithfully, across a range of resource requirements and execution speeds. Scalable online network simulation is a critical challenge—and differs from the offline simulation efforts generally studied by the networking research community.
- *Representative Background Load Modeling* Understanding what interaction of background and foreground load is critical to representative behavior. This is critical for all aspects of resource modeling, including computation, storage, and network systems.
- *Efficiency and Scalability* Achieving efficient simulation to enable study of long periods of behavior, and scalability to achieve the study of large systems—which often exhibit different behavior.

As part of the CGrADS effort, we will build and experiment with a number of generations of the MicroGrid tools, developing a succession of greater capabilities. These efforts will integrate our novel research efforts as well as relevant efforts developed in the community.

The second major infrastructure, the MacroGrid, integrates computational and network resources at CGrADS sites to provide a realistic (although less configurable) experimental testbed. This testbed will provide a more controlled environment, and likely a much higher degree of instrumentation and data capture, than is possible in typical Grid environments. This testbed will be used to validate MicroGrid simulations.

Subsequent efforts will focus on developing an experimental methodology for characterizing grid software in a manner that allows accurate evaluation of the software's behavior before deployment. A further goal of this work is understanding how to characterize a regime of behavior and also to identify those regimes for which behavior is poor, or at least uncharacterized. Possible approaches include statistical sampling, perturbation analysis, and enforcement of behavioral constraints (e.g., linearity) on software.

6 Education and Human Resource Development

The Principal Investigators are firmly committed to integrating education, human resource development, and knowledge transfer programs into CGrADS. More than 30% of the Center's NSF budget and 28% of the cost share will be allocated to supporting education and human resource (EHR) development. While this is a significant percentage, our experience in running successful EHR programs suggests that, as our programs mature and expand, they will require additional support from sources other than CGrADS. Thus, we will use the STC (NSF) budget in three ways: (1) to fund new pilot programs on a local level; (2) to provide support for students participating in CGrADS research; and (3) to continue participation in successful programs already being managed by the PIs, where we can obtain leverage on our expenditure. Where the pilot programs show particular promise, we will seek additional funding to scale them to a larger community.

The PIs have a history of success in EHR activities. Several members of the CGrADS team have implemented strong education and outreach programs, under the management structure of center-style programs (e.g. CRPC, NCSA, NPACI). These efforts have included programs for undergraduate and graduate students, training programs for K–12 teachers and students, and programs aimed specifically at increasing the number of women and minority students in the computational sciences. Richard Tapia, who will direct the CGrADS EHR efforts, has been recognized with several national awards for his work in this area.

As part of the CGrADS effort, we will initiate new programs that address graduate education, undergraduate education, and training for both K–12 teachers and parents. At the graduate level, our efforts will focus on the development of new courses that incorporate Grid-oriented problem solving, and on an active long-term exchange of graduate students between research groups. The former activity will create educational material suitable for adoption at other sites; the latter will broaden the experiences of the students at the seven CGrADS institutions. At the undergraduate level, our efforts will include new courses that bring Grid-enabled computing into the classroom, programs that involve students directly in the research programs of the PIs, and programs that build support communities for student participants to improve retention of those students in technical degree programs. At the K–12 level, our effort will focus on improving teacher training, through improvement and expansion of efforts like GirlTECH, and on a pilot program in parent education to encourage students, especially under-represented minorities, to pursue careers in science and technology. The pilot parent-education program will bring parents together in a safe, non-competitive environment to learn how to use computers, and enable them to better communicate and share computer experiences with their children.

The next section describes our plan for leveraging programs that the PIs have already established. It is followed by a discussion of the new pilot programs that we envision for CGrADS.

6.1 Leveraging Existing Programs

GirlTECH, one of the CRPC's most successful EHR programs, is a professional development course for K–12 teachers that combines instruction on the classroom use of information technology with a discussion of issues that have discouraged women from pursuing technical careers. Our direct goal is to equip teachers, counselors, and teacher educators with knowledge and strategies that encourage girls and underrepresented minorities to participate more fully in science and engineering activities, with an emphasis on computer technology and scientific computing. The problem of gender and ethnic inequality is still significant. Males continue to dominate attendance in advanced computer courses. Seldom will a school have an advanced computer science course with over 20% female enrollment. This problem has staggering repercussions. The Bureau of Labor Statistics lists computer scientists, computer engineers, and systems analysts as the three occupations with the fastest expected growth in employment for the period 1996–2006. As science, mathematics, and engineering become more and more dependent on computing, women's lack of engagement with computers must not become another impediment to achieving equity in the field. Regarding ethnic participation, it is important to note that approximately 56% of white families owned a computer in 2000, whereas only about 34% of Hispanic families and 33% of African American families owned a computer. Internet access from home was approximately 46% for white families and 24% for Hispanic and African American families (see `http://www.ntia.doc.gov/ntiahome/fttn00/charts00.html#f16`).

GirlTECH has trained roughly 140 teachers in the past five years. GirlTECH graduates conduct in-service training for other teachers, both in their local district and beyond. GirlTECH graduates have trained over 2,000 other teachers, who have, in turn, reached hundreds of thousands of students. The program's Web site (`http://www.crpc.rice.edu/CRPC/Women/GirlTECH`) includes math and science teaching material developed by GirlTECH participants; these lessons have been accessed more than 2 million times. With PACI support, GirlTECH is beginning to expand to PACI partner institutions. CGrADS will support this expansion of GirlTECH through the PACI infrastructure by supplying new course materials, management support, and additional participant funding. With this support, it is expected that the number of teachers being trained over the next 5 years will more than triple, and the number of students reached will approach one million.

Another CRPC program, Spend a Summer with a Scientist (SaS), demonstrates how a Science and Technology Center like CGrADS can enhance a university's culture while tackling one of the nation's most daunting educational challenges. SaS was designed to increase the number of women and underrepresented minorities in the computational sciences through enhanced recruitment and retention in graduate school. It provided opportunities for undergraduate and graduate students to work during the summer under the guidance of researchers and to develop within a supportive, nurturing community that valued diversity. The extensive success of SaS was a key factor in Rice's receiving one of the NSF's Minority Graduate Education grants in 1998 (now known as Alliance for Graduate Education and Professoriate (AGEP)). AGEP expanded SaS across all of Rice's science and engineering departments and to the University of Wisconsin-Madison. In 1997, the Learning through Evaluation, Adaptation, and Dissemination (LEAD) Center conducted an evaluation of SaS to aid in its replication at other universities and to uncover the essential elements contributing to its success. As to its effectiveness—LEAD found that of the 30 graduate SaS participants in the program at Rice, only 1 left without a graduate degree. In addition, 89% of the graduate and undergraduate participants studied reported that the program increased their desire to attend or remain in graduate school. As to the reasons for its effectiveness—the three program elements identified as having the greatest impact on participants' desire to attend or remain in graduate school included: interactions with Dr. Tapia, being in the company of other minority students, and being in the company of other students. CGrADS will incorporate this critical community-building component into its undergraduate and graduate programs as it strives to increase participation of domestic (American) students, especially minorities and women.

6.2 New Pilot Programs

Programs focused on Graduate Students Graduate education is one of the most important products of a research effort and a particularly effective way to transfer technology. Therefore, CGrADS will place special emphasis on graduate education. This is especially important for an emerging area like application development for the Grid because the topic has emerged only recently. Producing Ph.D.'s from underrepresented groups is a particular priority for the PIs, who remain committed to the principle that neither gender nor ethnicity should be a bar to achievement in this field.

One of the important lessons learned during our pilot GrADS project, undertaken as a part of the NSF Next Generation Software Program, is the extent to which researchers, staff, and students across all the institutions involved must collaborate closely to build pilot applications. Early in the lifetime of CGrADS, it will be critical to develop a corps of graduate students with the experience to bring up new applications in the emerging framework for Grid execution. To build this expertise, we will design and offer projects courses at all the participating institutions that will involve an introduction to Grid programming and use inter-institutional collaborative projects to construct applications. This grid-oriented, problem-solving course will be designed in a way that makes it suitable for export to other institutions. To solidify the collaborations, students in this course will present their work in seminars that are presented over the Internet.

As we learn more, we plan to develop an academic course on Grid architectures and programming systems that can be taught at both the graduate and undergraduate levels. Eventually, we expect that this course will be offered at each of the participating CGrADS institutions.

Because we have different concentrations of expertise at each of the CGrADS institutions, it will also be important for graduate students to spend extended periods of time at least one other CGrADS institution or to work with a collaborating application project. This experience will broaden each student's background and will make them better able to understand and reason about the complex interactions that arise in the development and use of Grid applications.

Programs for Undergraduates At the undergraduate level, we plan to make knowledge of the Internet and Grid application development more readily available to undergraduates. We envision two target audiences for this educational initiative: computer science and engineering majors, who need more technical educational activities, and the nontechnical student body, for whom a strong intuition about the Internet is becoming more important.

For students specializing in technology, CGrADS will focus on promoting graduate study among domestic (American) students at our own institutions (in our own backyard), targeting women and underrepresented minorities in particular. Frequently, "second tier" students at first rate institutions may not be encouraged to attend graduate school when, in fact, they could be quite successful. By second tier, we do not mean second rate. The second tier consists of individuals who are certainly talented and capable, and can succeed given proper guidance and direction, but who have not been properly developed. High school preparation plays a key role in how well a student will do as an undergraduate. Contrast the experiences of a student who might have been the valedictorian of a weak urban or rural high school with a student from an academically elite school. The first student will begin woefully behind and most likely will stay behind. Yet students like this have the talent to succeed. Had they attended a less academically-competitive university, they would have been among of its best students, and, possibly, would have been counseled and encouraged to attend graduate school.

This program will apply what we learned from the SaS experience to develop a supportive community of learners among undergraduates that promotes graduate studies. The program will involve activities that include building close relationships with caring, supportive faculty and graduate students; providing opportunities for undergraduate research; attending professional conferences; and receiving directed, personal, academic counseling. The PIs will collaborate to develop and test this model at several institutions so that, if successful, it can be adapted and disseminated more broadly. As with other CGrADS programs, if successful, we will seek leverage funding (e.g. from REU programs) to supplement the initial pilot project funding from the CGrADS budget.

CGrADS activities for the nontechnical student body will focus on developing general education courses in technology. As an example, Ken Kennedy has already begun to develop a general education course, entitled *Information Technology Architectures*, based on the hypothesis that you can understand the high-level principles used to design and build information technology systems without having to know how to write an object-oriented program. The course will cover a variety of topics including computer architecture, networks, principles of software design, and a survey of topics in software systems, such as operating systems, databases, client-server design, and computer security. It will take a functional approach to the subject, providing insight into how the systems work, rather than a constructive approach, requiring students to build them. Thus, it should let students go beyond the rudiments of how to use computers and the Internet to gain a deeper grasp of how to understand and manage information technology. This course supports the International Society for Technology in Education's national Technology Foundation Standards for Students that describe what students should know and be able to do with technology. Standard one, for example, says "Students demonstrate a sound understanding of the nature and operation of technology systems."

This course will be a cornerstone of Rice's general education offerings in Computer Science. Once this course is tested at Rice University, we will work to establish similar courses at other CGrADS sites.

Programs for K–12 Teachers and Parents A cornerstone of our K–12 program will expand on the GirlTECH program for teachers by targeting parents of urban minority K–12 students who do not have access to computers, training on their use, reliable telephone service, or Internet services. We believe that parents play an influential role in a child's decision to pursue a career in science and engineering, and that exposure to technology training and career counseling will enhance their understanding of the opportunities available to their child through science and engineering. To tackle the issue directly, we will design and implement a series of Parent Technology and Career Nights at local schools. These activities will be designed and overseen by Richard Tapia, and implemented by teachers in the GirlTECH network.

These evening classes will target the parents of underrepresented minority students. They will bring the parents together in a supportive, non-competitive environment to learn about computing, network resources, and their role in education. We will seek corporate, state, and foundation funding to make computers and network connections affordable for parents who complete the program. It is clear to us that implementing this program will be difficult, however, we firmly believe in the principles of the program and we are strongly motivated to make it succeed. This approach can have a strong and far-reaching impact—encouraging more

	Total Degrees	Time To Completion	Gender		Nationality	
			Male	Female	US	non-US
Masters Students	24	2.4	19	5	8	16
Ph.D. Students	20	5.9	14	6	13	7

	Ethnicity				
	White	African American	Hispanic	Asian	Other
Masters Students	10	0	3	11	0
Ph.D. Students	9	1	5	5	0

Table 1: Summary Chart of M.S. and Ph.D. Degrees During the Past 3 Years

minority students to follow careers in the computational sciences. If we succeed, we will export the program to other CGrADS sites, and other places where we can find appropriate leadership (e.g. through the NSF PACI EOT program sites), using the experiences and contacts developed in expanding GirlTECH.

We will also create a version of the general education course on information technology architectures for for K–12 teachers in the GirlTECH program, with the goal of providing them with a deeper understanding of the Internet, computers, and the administration of networked systems. This will prepare them to integrate information technology into their own schools more effectively and to teach others the same lessons. Programs such as GirlTECH provide the CGrADS PIs with enormous leverage as they try to make the technology developed in the laboratory accessible in K–12 schools.

6.3 PI's Record with Recent Students

As part of this solicitation, we were required to collect data from the project PIs on the number of students that they have advised through the graduate degree process in the past three years, and on the demographics of those students with regard to gender, ethnicity, citizenship, and time to complete their degree. Table 1 summarizes this data. The average time to finish a Ph.D. was 5.9 years. The average time to finish a Master's Degree was 2.4 years. We feel that these numbers are reasonable, given the number of students involved directly in experimental research, and the involvement of many of them in mentoring programs for underrepresented students. The table shows that the number of underrepresented minorities produced by the group is higher than the national average.

In our experience, a graduate program that aggressively recruits and retains underrepresented minorities often accepts and includes some students who have less rigorous backgrounds than their peers in graduate school. As with many students from non-traditional backgrounds, these students have the talent and ability to succeed, but they need extra time in their graduate programs to strengthen their background knowledge before completing their degrees. Thus, we see a tension between our goal to increase the number of women and underrepresented minority students in our program, and the expressed desire to shorten the time that students spend obtaining a graduate degree. If CGrADS is funded, the PIs will:

1. be aggressive in recruiting and retaining domestic students from underrepresented ethnic and gender groups, even when their backgrounds indicate that they may require a longer time to graduate.
2. be aggressive in monitoring our time to completion and ensuring that we do not keep students in the graduate program longer than necessary.

We recognize the need to balance time to completion against accepting, training, and graduating students from non-traditional backgrounds. We will work with our departments and our institutions on both of these issues.

6.4 Summary

The CGrADS team has a strong track record of leadership in developing new programs in Education and Human Resources. Richard Tapia, who will direct the CGrADS programs in these areas, brings deep experience and skill in addressing the inclusion of underrepresented groups and in working with the K–12 community. Both Tapia and Ken Kennedy, the CGrADS director, have spent years building bridges to the Houston Independent School District (HISD) through direct service and through training programs like GirlTECH. Tapia has served as Advisory Board Chair of HISD's NSF Urban Systemic Initiative program since its inception. As a result, they have immediate access to HISD administrators at both the central office and in the schools. HISD, the seventh largest school district in the U.S., serves an enrollment of over 200,000 students, with Hispanic and African American students comprising 87% of the students (`http://www.houstonisd.org/Pubs/AboutHISD/Enrollment.htm`). The close ties with HISD will help CGrADS members initiate pilot projects with HISD schools and HISD teachers.

The CGrADS PIs have a history of commitment and participation in successful programs that enhance education and human resource development, with a particular emphasis on improving the representation of women and underrepresented minorities in science and engineering. NSF funding, supplemented by a significant portion of the cost sharing, will be used to support existing programs where we feel we can successfully add value. It will also be used to implement a number of pilot programs where, if they show promise of success and scalability, we will seek additional funding to expand the programs to a larger community. The CGrADS management structure will formalize EHR development activities as a regular part of the Center's focus and review. As pilot programs scale up through new funding support, or are cancelled because of lack of success, the funds become available for new programs. One thing is certain—led by Richard Tapia, this group of principal investigators will generate and pursue many new ideas for improving technology education at all levels and increasing the participation of underrepresented groups.

7 Knowledge Transfer

Knowledge transfer is an important part of the CGrADS mission. The principal investigators have had significant success transferring the knowledge and technology created in their research programs to others. Their efforts have addressed a broad variety of target audiences, including the broader research community; industrial partners in both established companies and start-ups; application scientists in academia, the national labs, and industry; and students at the secondary, undergraduate, and graduate levels. CGrADS will use the experience, the contacts, and the established knowledge-transfer activities of the PIs and the organizations that they represent to leverage its own knowledge transfer programs and to transfer the knowledge and technology from this project to a variety of target audiences.

The CGrADS efforts in knowledge transfer will attack the problem with different degrees of focus. CGrADS knowledge transfer activities will include programs aimed at a broad audience, programs that target narrow segments of the technical community, and programs aimed at students.

7.1 Broad-based Knowledge Transfer

CGrADS will produce a variety of publications that aim for the broadest possible audience. The Center will build and maintain a Web site that provides access to technical reports, software, vision documents, and technical presentations by the PIs. The Web site will provide a conduit for someone actively seeking information about CGrADS. Unfortunately, a Web site is, essentially, a passive tool. The Web site will not actively bring the Center to the attention of the target community. To provide an active outreach to the broader community, CGrADS will produce a regular newsletter and distribute it widely. When the newsletter appears in people's mailboxes, it will call their attention to CGrADS activities and provide them with pointers to the Web site, to publications, and to contacts for personal interaction. We will also seek out venues, like HPCWire, for wider online distribution of CGrADS newsletter articles.

Many CGrADS projects will produce artifacts that embody the research results—publications, presentations, and software systems. The PIs intend to distribute these artifacts widely, using whatever mechanisms are appropriate. Several of the investigators have played critical roles in the establishment of well-known, Web-based software repositories (NETLIB, SoftLib, NCSA's software repository, and NPACI's software repository). The PIs will make software and technical reports available through such repositories, as well as through the CGrADS Web site.

Another important avenue for disseminating CGrADS research results will be via inclusion in high-quality toolkits. For example, the Globus Toolkit is already in use by numerous groups worldwide for Grid deployment and application projects. CGrADS research will contribute to the further development of this technology base. Globus developers and collaborators in the U.S. and Europe are currently working to establish the Consortium of Open Grid Software to further the overall goal of open architecture, open source Grid software. We will participate in this initiative as and when appropriate.

7.2 Targeted Knowledge Transfer Activities

Many of the results of CGrADS research should have a strong direct impact on smaller groups: application scientists (both individual projects and broader coalitions of users), companies that either use large-scale computation or build products that support it, and the community of other computational scientists. CGrADS will use different strategies to establish and maintain ties with each of these groups.

7.2.1 Individual Projects

To advance the CGrADS research agenda, the PIs will work closely with a small number of application groups to develop Grid-enabled versions of their applications. This is a fundamental part of the research process. By helping to port emerging applications to the Grid, the PIs will develop a deeper understanding of the problems that affect Grid performance and of the challenges that confront users of their tools and systems. These collaborations establish a two-way relationship: CGrADS researchers gain valuable insights and application groups gain knowledge and experience.

The PIs have already established working collaborations with the Cactus and GriPhyN application groups, among others. The Cactus group, led by Ed Seidel, is currently incorporating Grid-related technology developed by CGrADS PIs directly into the Cactus system. The GriPhyN project, led by Paul Avery and involving CGrADS Co-PIs Foster and Kesselman, aims to use CGrADS technology for performance estimation and scheduling in large-scale data grids. Also, the NEESgrid project, led by Tom Prudhomme and also involving Foster and Kesselman, provides a technology transfer avenue to the engineering community. The research effort, described in Section 5, will involve initiation of new collaborations, as the PIs seek out emerging applications with diverse characteristics to broaden their own base of expertise.

7.2.2 Broader Coalitions of Users

The PIs already have ties with several large groups that are natural consumers of the knowledge and technology that CGrADS will produce. The PIs will work through those organizations to find additional customers. Each of these organizations has periodic technical meetings where CGrADS PIs already participate. As the CGrADS research matures, the PIs will offer talks and tutorials at these meetings to inform their colleagues and attract new users and collaborators.

NSF PACI Programs CGrADS has deep existing ties to the two NSF-sponsored Partnership in Advanced Computational Infrastructure (PACI) programs (NCSA and NPACI). Many of the PIs are involved with the PACIs; some PIs are in both programs. Dan Reed (Director of NCSA) and Fran Berman (Director of NPACI) are both PIs on this proposal and have expressed strong interest in being able to integrate CGrADS results into their PACI activities.

CGrADS and the PACIs should have a natural synergy. The PACIs have funding for technology development and technology transfer, but not research. CGrADS researchers will work closely with application teams and application developers involved in the PACI programs. This will ensure that application scientists evaluate CGrADS ideas, designs, and prototypes, and that CGrADS research and knowledge will be incorporated into the full-scale implementations built by PACI developers.

Los Alamos Computer Science Institute Several of the PIs are involved in the Los Alamos Computer Science Institute (LACSI), a program of research focused on the computational problems that confront researchers in the DOE laboratories, and at Los Alamos National Laboratory in particular. This program provides direct contact between CGrADS researchers and another community of application scientists. Again, this collaboration should foster a bidirectional flow of knowledge, expertise, and opportunity. CGrADS research results should help the DOE application scientists build more effective implementations; experience helping those scientists will shape the PIs' vision of the problems that face developers of Grid-enabled applications.

Accelerated Strategic Computing Initiative Several of the CGrADS PIs have projects currently funded by the Accelerated Strategic Computing Initiative (ASCI) of the DOE. This provides them with access to another group of researchers who should be natural consumers of CGrADS results. The PIs will work within the ASCI community to find customers for CGrADS research results; they will also bring the problems that arise in the ASCI community back into the research discussions of the CGrADS community. Of course, this activity is contingent on continued ASCI funding for individual CGrADS PIs.

7.2.3 Companies

The PIs have a long and active history of collaborating with industry and of transferring research results into practical products. This includes a broad range of products and technologies: compilation techniques, numerical libraries, MPI, PVM, and High Performance Fortran. Their efforts have involved companies across the spectrum in information technology and its application to large-scale computing. The CGrADS effort has already begun to attract industrial interest. We have discussed potential collaborations with companies that include Compaq; IBM; SUN Microsystems; Cray, Inc.; Chiaro Networks; Myricom; and Entropia. Industrial involvement is mutually beneficial. It helps CGrADS researchers reach the right people inside the companies to facilitate transferring CGrADS technologies and research results into their products. At the same time, it provides the companies with the opportunity to help shape the Center's research direction by bringing the problems that they encounter to the PIs' attention and by providing direct feedback on the research programs, through forums like the CGrADS Industrial Council.

7.2.4 Community of Computational Scientists

The aforementioned activities all target specific groups—our collaborators on individual applications; members of alliances like NCSA, NPACI, and LACSI; and individual companies. To interact with the computational science community in a broader sense, the Center will sponsor workshops that focus on critical problems in Grid computing. These workshops, typically held in conjunction with an already-established scientific meeting, should provide an environment for the open exchange of ideas within the community—to foster new ideas, to report new results, to identify critical problems, and to suggest approaches for solving them.

Each workshop will be led by one or more of the CGrADS PIs. To leverage CGrADS funding, we will seek joint sponsorship with other organizations, including the PACIs, LACSI, and related professional organizations such as SIAM and the ACM SIGs.

We have also reserved a small amount of funding to support and encourage CGrADS participants who are interested in organizing technical workshops on other topics relevant to CGrADS research. The Center for Research on Parallel Computation (CRPC) used this kind of support to encourage the early workshops on MPI and PVM; these efforts evolved beyond CRPC support and produced useful community-based standards and software. We hope to be able to claim similar successes for CGrADS.

Finally, the Center will avail itself of the classic academic methods of knowledge transfer—producing students, publications, and software. Students trained in CGrADS research programs will go forth to jobs and create ties between the PIs and their new organizations. The PIs will aggressively publish research results in national and international forums, and make presentations to a variety of audiences. The software systems that result from CGrADS research will embody the research results in a concrete form. The PIs have a long history of making such systems available in the community; they intend to continue that tradition.

7.3 Knowledge Transfer to Students

A natural synergy exists between some of the knowledge transfer activities and our efforts in Education and Human Resource (EHR) Development (see § 6). For example, our programs focusing on promoting graduate study among domestic undergraduate students will involve these students in CGrADS research activities, enabling the students to experience the knowledge discovery process first hand. These students will take this knowledge with them, and disseminate it during their careers. Similarly, our EHR programs that develop courseware, that train K–12 teachers, and that sponsor long-term graduate-student exchanges will involve both education and knowledge transfer components.

The PIs are actively involved in the education of students at the undergraduate and graduate level. Their work in CGrADS will have a direct impact on the training of these students, through research insights and problems brought into the classroom, through selection of thesis topics, and through the creation of an intellectual community focused on the problems of Grid-based computing. However, the distributed nature of the Center makes it important to have effective mechanisms for cross-site interaction—particularly for students. The Center will direct some of its knowledge transfer effort at facilitating cross-site interaction and increasing the impact of CGrADS research on education at the participating universities.

The Center will create specific opportunities for student involvement. These will include recruiting undergraduates to work on CGrADS research projects during the summer and the school year; a program to encourage long-term graduate student exchanges between universities; and the involvement of graduate students at the regular internal workshops.

One mechanism for spreading knowledge of Grid-based computing beyond the PIs' own students is the incorporation of research results and insights into new courses. Again, the PIs have a history of developing new courses, writing textbooks, and creating new degree programs. Based on this history, we expect to incorporate knowledge from CGrADS research activities into courses in Computer Science, Applied Mathematics, and multidisciplinary Computational Science/Engineering degree programs at the participating institutions.

7.4 Summary

Each of the PI's associated with this project has a successful history of transferring the knowledge and technology of their research activities to the broader community. CGrADS will build upon this history not only by leveraging the programs and the infrastructure these investigators already have in place, but by implementing new programs for effective knowledge transfer as well. We hope that the programs focused on domestic graduate and undergraduate students and the special workshop series will achieve a level of success and community visibility that make them signature successes for CGrADS.

8 Rationale for the Center Concept

To achieve the goal of making the Grid usable, we must bring together a team of experts from a number of related areas. This group, already collaborating on the GrADS project, is committed to the long-term, large-scale effort required to develop effective, usable tools for programming Grid-enabled applications. To achieve true success, however, will require an effort of the scale and duration of a Science and Technology Center for the following reasons:

- The project requires an effort involving leading scientists in a variety of areas in computer science and applied mathematics, working closely with application scientists and engineers. To carry out an effort of this scope requires the collaboration of scientists at many different universities and organizations, which in turn requires a coordinated, cross-institutional management structure that is found only in a center. The project is too broad for any single institution to sustain, particularly with support only through traditional single-investigator grants or even through medium-sized grants to research teams.
- Success will require the construction of large, complex software systems that include sophisticated compilers, libraries, programming environments, networking and communications infrastructure, and performance monitoring systems. Building such experimental software systems is beyond the scope of normal single-investigator or team projects because they require years of effort from many different investigators at different institutions to build and refine. Once constructed, however, these systems can serve as frameworks for many different research projects. Hence, a center is needed to build the appropriate software resources to jump-start a broader effort by the Grid-computing research community.
- Our early experience porting applications to the Grid has revealed a set of problems that must be solved before we can achieve the project's goal of making the Grid usable. Individually, these problems are challenging; solving them together will require coordination and knowledge-sharing between the research efforts, as well as shared interfaces and components in the software systems. As yet, we do not understand all the relationships among the component technologies that will be required to address these problems. For example, what are the proper roles of the compiler, library developer, programming environment, scheduling and service negotiation system, and the performance monitor in carrying out a Grid computation? Because of these uncertainties, the entire effort will need to be able to change research directions and seize new opportunities in response to lessons learned through experience. This kind of adaptivity can only be provided in a center with long-term funding and the flexibility to pursue new directions.
- For the Grid to become accessible to a larger community of scientists and engineers, an effort will be needed by a broad and diverse collection of computer science researchers. Not all of these researchers will be associated with CGrADS. However, as a Science and Technology Center, CGrADS can serve as a focal point for these efforts. The visibility and credibility that come from selection as a Science and Technology Center will play an important role in our ability to attract users for our early prototype systems, to influence emerging standards, and to educate the community about the potential for Grid-based computing.
- Our previous experience with the NSF Center for Research on Parallel Computation has shown us that a Science and Technology Center generates a high level of intellectual ferment at all the participating sites, creating an atmosphere that makes it possible to pursue new visions through coordinated research planning; to generate excitement in the community; and to pursue new educational programs needed to increase the awareness of technological issues and to increase participation of underrepresented groups. This intellectual ferment will draw many new researchers and graduate students into the community effort that will be needed to make Grid computing widely accessible.

For all these reasons a Science and Technology Center, with multiple-year support and the flexibility to pursue a long-term research vision through to its conclusion, is the ideal vehicle for pursuing the CGrADS vision of making Grid computing into a commodity for scientists and engineers.

9 Management

Our strategy for managing the CGrADS efforts will be modelled after the strategy employed by the Center for Research on Parallel Computation (CRPC)—a successful, geographically-dispersed, Science and Technology Center. The strategy is based on an ongoing, cyclic pattern of review, planning, and implementation. This approach, which was reviewed and endorsed by the National Academy of Public Administration, allowed the CRPC to both adapt to changes in personnel and institutions and adopt new research directions, while staying focused on its central mission.

Management Structure CGrADS will include include researchers from seven institutions: Rice University; University of California, San Diego; University of Chicago; University of Houston; University of Illinois, Urbana-Champaign; University of Southern California, Information Sciences Institute; and University of Tennessee, Knoxville. Rice University will be the lead institution, with Ken Kennedy serving as Director. He will be assisted by an Executive Director, Linda Torczon; a Director for Education and Human Resource Development, Richard Tapia; an Associate Director for Business, Danny Powell; and an Associate Director for External Relations. Three committees will provide further guidance: an Executive Committee (EC), an External Advisory Committee (EAC), and an Industrial Council (IC). Figure 3 depicts the reporting relationships between these individuals and committees; the individuals and committees have the following roles:

- The Director is responsible for oversight, planning, budgetary decisions, and reporting on CGrADS activities to the NSF. The director receives advice and assistance from the individuals and committees shown in Figure 3. The director serves as chair of the Executive Committee.
- The Executive Director is responsible for implementing decisions made and policies set by the Director; reporting CGrADS activities to the NSF through annual reports and site visits; and facilitating communication between the director and the individuals and committees that advise him.
- The Director for Education and Human Resource Development is responsible for designing and implementing CGrADS educational programs, particularly programs designed to encourage minorities and women to pursue careers in technical fields.
- The Associate Director for Business is responsible for financial aspects of the Center, including financial management, financial reports to the NSF, and the subcontracts between the lead institution and the other six participating institutions.
- The Associate Director for External Relations is responsible for producing and disseminating information about CGrADS research, education, and knowledge transfer activities through a CGrADS newsletter and a CGrADS Web site; arranging CGrADS workshops and committee meetings; and facilitating communication with external groups.

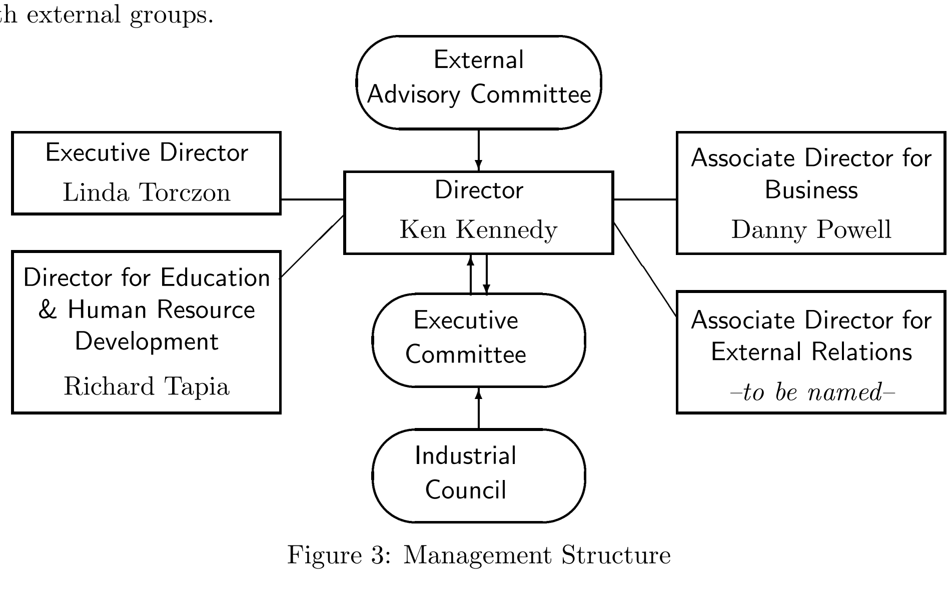

Figure 3: Management Structure

- The Executive Committee will convene regularly via meetings, teleconferences, and e-mail to plan research, education, outreach, and technology transfer programs; review progress; and approve major budgetary commitments. The EC will consist of Ken Kennedy (chair), Keith Cooper, John Mellor-Crummey, Richard Tapia, and Linda Torczon (Rice); Fran Berman and Andrew Chien (UCSD); Ian Foster (U. Chicago); Lennart Johnsson (UH); Dan Reed (UIUC); Carl Kesselman (USC-ISI); and Jack Dongarra and Rich Wolski (Tennessee).
- The External Advisory Committee will convene annually to provide the director with an external view of CGrADS progress, plans, and management, and to advise the director on future priorities. The EAC will produce a summary of their findings and advice. Additional meetings may be held at the discretion of the EAC. The EAC will consist of a diverse group of eight members, including members from academia, government, and industry. Members of the EAC will be chosen by the Executive Committee, with advice from both the participating institutions and the NSF. Criteria for selection will include scientific credentials, vision, and experience. Every effort will be made to ensure that there is a balance among memberships from universities, industry, and government labs. We will also make strong efforts to include members of underrepresented groups in science and engineering.
- The Industrial Council, composed of representatives from companies supporting core CGrADS activities, will convene annually to discuss current and future directions, promote technology transfer, explore potential joint projects, and provide the director with an industrial view of CGrADS progress and plans.

Each CGrADS site will have both a senior researcher and an administrative contact to assist in operational matters, including reporting and budgeting. Researchers and administrative personnel will communicate regularly through e-mail, the World Wide Web, teleconferences, and meetings.

The senior management team brings to the task a wealth of prior experience in managing large, collaborative, distributed research projects, including the CRPC. The PIs have prior experience working together as collaborators on a joint project, the Grid Applications Development Software project.

Review, Planning and Implementation The research and development activities of the Center will be organized into three types of activities. The *basic research* activity will draw upon our preliminary research and implementation experiences to develop principles and new research directions. The *infrastructure development* activity will develop the GrADS software base (GrADSoft), incorporating the results of our research, and providing a basis for experimental verification of our ideas and for technology transfer. The *application evaluation* activity will involve joint investigations with partners in the PACIs, and elsewhere, in which GrADSoft will be used to develop compelling Grid-enabled applications. The emerging MicroGrid and MacroGrid testbeds will be used to test, tune, and monitor these applications. These evaluations will produce critical feedback for the research efforts and provide a venue for technology transfer.

These three activities form an integrated and interconnected set of projects that can only succeed if they are closely coordinated and tightly coupled. As principles emerge from the prototype tools and applications, the CGrADS researchers will review them and feed the best ideas back into the research program. Thus we plan a series of software demonstrations that will exercise, successively, deeper and deeper ideas about ways to construct Grid-enabled applications. This kind of close coordination, rapid feedback, and cross-fertilization is needed if the project is to achieve its overall goal.

To coordinate activities among the various projects and sites, we will have three internal meetings per year. These meetings will combine an internal technical workshop with a planning session. They will serve as forums where we can review what we have accomplished, refine the overall design of the system based on what we have learned, and assess new ideas and technologies. At each of the internal meetings, most of the EC will be present. The Director and the EC will hold a separate planning session to address higher level technical issues, set goals, initiate and/or terminate projects, and deal with issues of policy or resource allocation. In addition to these meetings, the EC will communicate regularly via teleconferences and e-mail. These internal meetings embody the management strategy over short time periods: review, plan, and implement.

Over longer periods of time, the Director and the EC will receive external reviews from a number of sources. The EAC will meet annually to review all aspects of the Center, including research, education, outreach, management, and technology transfer. The IC will meet annually to discuss research progress and plans, to provide the EC with an industrial perspective on its research and knowledge transfer activities,

Name	Who Attends	How Often	Purpose
Internal Workshops	PIs, researchers, students	*3/year*	Present progress, discuss & refine plans; EC meets in an addition planning session
EAC	EAC & EC	*annual*	Review progress & plans for all center activities, (research, education, outreach, management, & tech. transfer)
IC	IC & EC	*annual*	Provide EC with industrial perspective on technical progress, plans, & technology transfer
NSF Site Visits	Site visitors, PIs & researchers	*periodic*	Review progress & plans for all center activities to provide advice, direction, & feedback
EC	EC	*3/year or more*	Review progress in all center activities, make plans, allocate resources, assess new directions & technologies

Figure 4: Summary of Meetings

and to discuss potential joint projects and technology transfer opportunities. The NSF will convene periodic site visits to review all CGrADS activities and to provide feedback, direction, and advice. These external reviews will play an important role in assessing the level of success of the various CGrADS activities. The Director and the EC will incorporate feedback from the external reviews into the internal review, plan, and implementation cycle.

Reporting and Information Dissemination Providing effective external communication is a critical part of building a successful Science and Technology Center. Many of the positive benefits associated with a center (see § 8) derive directly from external visibility for high-quality research programs. In addition to the normal academic venues (publication in journals and conferences, external talks, and software distribution), CGrADS will pursue a number of opportunities to raise its public profile.

- CGrADS will maintain a central Web site, produced and maintained by the Associate Director of External Relations and staff, that will include information on all aspects of the Center (research, education, outreach, and tech transfer).
- CGrADS will produce three hardcopy newsletters per year, written by the Associate Director for External Relations & staff with production help from NCSA, covering all Center activities. Newsletter content will also be available form the CGrADS Web site. We will seek out venues, like HPCWire for wider online distribution.
- CGrADS will host technical workshops for the scientific community on topics related to Center research. These workshops will bring CGrADS participants together with potential users, collaborators, and other researchers from outside the Center to discuss problems and progress in the area. (See § 7.)

CGrADS will also produce annual reports and renewal proposals as dictated by the NSF.

Involvement with External Groups The CGrADS meeting structure and information dissemination mechanisms serve as a starting point for involving other groups and researchers in CGrADS programs.

- CGrADS workshops, Web site, newsletters, and mailing lists are all mechanisms for reaching the outside community. The CGrADS External Relations staff will facilitate communication with external groups.
- Industrial collaborations will be managed by the researcher most closely involved with the industrial group. The IC provides a mechanism for industrial collaborators to feed their experience, evaluations, and suggestions into the CGrADS planning process.
- Collaborations with external researchers, in both academia and industry, will be handled by the CGrADS researcher with whom they work most closely. Active external collaborators will be invited to the regular internal workshops, where they can directly contribute to CGrADS growth, progress, and direction.

10 Lists of Project Personnel and Institutions

10 (a) Project Personnel

Academic Institutions

Aydt, Ruth, University of Illinois, Urbana-Champaign; College of Engineering; Computer Science Department; Senior Research Programmer

Angelo, Dave; University of Chicago; Division of Physical Sciences; Computer Science Department; Research Programmer.

Avery, Paul; University of Florida; Gainesville; Department of Physics; Collaborator.

Berman, Francine; University of California at San Diego; Jacobs School of Engineering; Computer Science & Engineering Department (CSE); Co-PI

Blackford, Susan, University of Tennessee at Knoxville; College of Arts and Sciences; Computer Science Department; Research Scientist.

Bourne, Philip E., University of California, San Diego; San Diego Supercomputer Center; Collaborator.

Chien, Andrew A., University of California at San Diego; Jacobs School of Engineering; Computer Science & Engineering Department (CSE); Co-PI

Cooper, Keith, Rice University, George R. Brown School of Engineering; Computer Science Department; Center for High Performance Software Research; Co-PI

Dail, Holly, University of California at San Diego, Jacobs School of Engineering; Computer Science Department; Research Programmer.

Dongarra, Jack, University of Tennessee at Knoxville; College of Arts and Sciences; Computer Science Department; Co-PI

Ellisman, Mark H., University of California, San Diego; Center for Research in Biological Structure; Depatment of Neurosciences; Collaborator.

Fowler, Robert J.; Rice University; George R. Brown School of Engineering; Computer Science Department; Center for High Performance Software Research; Research Scientist

Foster, Ian; University of Chicago; Division of Physical Sciences; Computer Science Department; Co-PI

Gullapalli, Sridhar, University of California at San Diego, Jacobs School of Engineering; Computer Science Department; Research Programmer.

Hanson, Richard; Rice University; George R. Brown School of Engineering; Computer Science Department; Center for High Performance Software Research; Research Scientist

Johnsson, S. Lennart; University of Houston; College of Natural Sciences and Mathematics; Computer Science Department; Co-PI

Kennedy, Ken; Rice University; George R. Brown School of Engineering; Computer Science Department; Center for High Performance Software Research; Project Director

Kesselman, Carl; University of Southern California, Information Sciences Institute; Co-PI

Lanius, Cynthia; Rice University; George R. Brown School of Engineering; Computational and Applied Mathematics Department; Outreach Project Manager.

Mazina, Mark; Rice University; George R. Brown School of Engineering; Computer Science Department; Research Programmer.

Mellor-Crummey, John; Rice University; George R. Brown School of Engineering; Computer Science Department; Center for High Performance Software Research; Co-PI

Mendes, Celso, University of Illinois at Urbana-Champaign; College of Engineering; Computer Science Department; Research Collaborator.

Obertelli, Graziano, University of California at San Diego, Jacobs School of Engineering; Computer Science Department; Research Programmer.

Reed, Daniel A.; University of Illinois, Urbana-Champaign; College of Engineering; Computer Science Department; Co-PI

Sievert, Otto, University of California at San Diego, Jacobs School of Engineering; Computer Science Department; Research Programmer.

Subramaniam, Shankar, University of California, San Diego; Departments of Bioengineering & Chemistry &

Biochemistry; Collaborator.

Swany, Martin, University of Tennessee, Knoxville; College of Arts and Sciences; Computer Science Department; Research Programmer.

Tapia, Richard A.; Rice University; George R. Brown School of Engineering; Computational and Applied Mathematics Department; Co-PI

Torczon, Linda; Rice University; George R. Brown School of Engineering; Computer Science Department; Co-PI.

Wolski, Richard; University of Tennessee, Knoxville; College of Arts and Sciences; Computer Science Department; Co-PI.

International Academic

Seidel, Ed, Max Planck Institute, Potsdam, Germany; Research Collaborator.

Industrial Participants

Knobe, Kathleen; Cambridge Research Laboratory, Compaq Computer Corporation; Collaborator

Wallach, Steven J.; Chiaro Networks; Richardson, Texas; Collaborator.

Seitz, Charles, L.; Myricom, Inc., Arcadia, California; Collaborator.

10 (b) Institutions

STC Academic Partners

Rice University

University of California, San Diego

University of Chicago

University of Houston

University of Illinois, Urbana-Champaign

University of Tennessee, Knoxville

University of Southern California, Information Sciences Institute

Other Academic Institutions

National Partnerships for Advanced Computational Infrastructure (NPACI), San Diego Supercomputer Center, University of California, San Diego

National Center for Supercomputing Applications (NCSA), University of Illinois, Urbana-Champaign

Federal Government (National Laboratories)

Argonne National Laboratory

Los Alamos National Laboratory

Industry

Compaq Computers

Chiaro Networks

Cray, Inc.

IBM T.J. Watson Research Center

Myricom, Inc.

Sun Microsystems, Inc.

International

Max Planck Institute, Potsdam, Germany

10 (c) Potential Reviewers with Conflicts of Interest

Allen, Fran, IBM-Watson, National Research Council report collaborations, Kennedy

Arnold, Dorian C., University of Tennessee, co-author, Kennedy

Balay, Satish, Argonne National Laboratorory, co-author, Kennedy

Balsara, Dinshaw, Univ of Illinois, at Urbana Champaign, research collaborator, Kennedy
Breshears, Clay, KDI, research collaboration, Kennedy
Briggs, Preston, Cray, Inc., co-author, Cooper, Torczon, Kennedy
Carle, Alan, Rice University, co-author, Kennedy
Chandy, Mani K., California Institute of Technology, co-author, Kennedy
Dawson, Clint N., The University of Texas at Austin, co-author, Kennedy
Dennis, John E., Rice University, co-author, Kennedy
Doeldel, Eusebius J., California Institute of Technology, co-author, Kennedy
Eijkhout, Victor, University of Tennessee, co-author, Kennedy
Fletcher, Katherine, Rice University, co-author, Cooper
Fox, Geoffrey, Florida State University, co-author, Kennedy
Gannon, Dennis; Indiana University; collaborator, Kennedy, Reed, Berman, Dongarra, Foster.
Gibson, Garth, Carnegie Mellon University, co-author, Reed
Gropp, William, Argonne National Laboratory, co-author, Kennedy
Harvey, Tim, Rice University, co-author, Cooper, Torczon
Hayder, Ehtesham, former Rice research scientist, Kennedy
He, Jingsong, Dell Computer, thesis advisor, Cooper
Henderson, Ron, California Institute of Technology, co-author, Kennedy
Karypis, George, University of Minnesota, co-author, Kennedy
Keller, Herb B., California Institute of Technology, co-author, Kennedy
Koelbel, Charles, National Science Foundation, co-author, Kennedy
Kumar, Vipin, University of Minnesota, co-author, Kennedy
Kurc, Tahsin M., University of Maryland, co-author, Kennedy
Lee, Wonsuck, The University of Texas, co-author, Kennedy
Li, Kai, Princeton University, co-author, Reed
Lu, John, LSI Logic, co-author, Cooper
Madhyastha, Tara, UC Santa Cruz, co-author, Reed
McInnes, Lois Curfman, Argonne National Laboratory, co-author, Kennedy
McIntosh, Nathaniel, Hewlett Packard, co-author, Cooper
McKinley, Kathryn, University of Massachusetts, co-author, Cooper, Torczon, Kennedy
Meiron, Dan, California Institute of Technology, co-author, Kennedy
Mirkovic, Dragan, University of Houston, research collaborator, Johnsson
Oldehoeft, Rod, Los Alamos National Laboratory, research collaborator, Kennedy
Parashar, Manish, Rutgers State University of New Jersey, co-author, Kennedy
Petitet, Antoine, University of Tennessee, research collaborator, Dongarra
Reynders, John, Mind Spring Corporation, research collaborator and co-author, Kennedy
Saltz, Joel, University of Maryland, co-author, Reed
Samtaney, Ravi, California Institute of Technology, co-author, Kennedy
Schielke, Philip, Texas Instruments, co-author, Cooper
Schloegel, Kirk, University of Minnesota, co-author, Kennedy
Simitci, Huseyin, Seagate, co-author, Reed
Simpson, L. Taylor, BOPS Computer, co-author, Cooper
Smarr, Larry, University of California at San Diego, research collaborator, Kennedy
Smirni, Evgenia, College of William and Mary, co-author, Reed
Smith, Barry F., Argonne National Laboratory, co-author, Kennedy
Smith, Burton, Cray Inc., research collaborator, Kennedy
Soni, Bharat, Mississippi State University, co-author, Kennedy
Sorensen, Dan, Rice University, research collaborator and co-author, Kennedy
Stevens, Rick, Argonne National Laboratory, research collaborator and co-author, Kennedy
Subramanian, Devika, Rice University, research collaborator, co-author, Cooper
Thakur, Rajeev, Argonne National Laboratory, co-author, Kennedy
Thompson, Joe, Mississippi State University, co-author, Kennedy
Thorp, John, Los Alamos National Laboratory, research collaborator, Kennedy
Towns, John, University of Illinois at Urbana Champaign, research collaborator, Kennedy

Tuecke, Steve, Argonne National Laboratory, technical collaborator, Kennedy,
Vetter, Jeffrey, Lawrence Livermore National Laboratory, co-author, Reed
Vick, Christopher, Sun Microsystems/JavaSoft, co-author, Cooper
Whalley, David, Florida State University, research collaborator and co-author, Mellor-Crummey
Wheeler, Mary F., University of Texas at Austin, research collaborator and co-author, Kennedy
White, Andy, Los Alamos National Laboratory, research collaborator and co-author, Kennedy
Wienskoski, Edmar, Motorola, thesis advisor, Cooper
Wu, Ahijun, Rice University, co-author, Kennedy

Federal Government (national laboratories)
Advanced Computing Laboratory, Los Alamos National Laboratory, research collaborators, Kennedy
DARPA/NSF, National Compiler Infrastructure Workshop, Chair, Kennedy
Lawrence Livermore National Laboratory, consultant, Kennedy
NASA Space Shuttle Program Lead Office, Visiting Scientist, Kennedy

Non-governmental Organizations
Computing Research Association, Member, Board of Directors, Kennedy
Center for the Analysis and Prediction of Storms, Member, External Advisory Board, Kennedy

Industry
Cray, Inc., Member, Board of Directors, numerous research collaborations, Kennedy
IBM T.J. Watson Research Center, Member, Advisory Committee, Deep Computing Institute, Kennedy .

11 Intellectual Property Rights

The research performed under the auspices of CGrADS will undoubtedly lead to the development of intellectual property that can be protected with copyrights or patents. The inter-institutional nature of CGrADS complicates many issues surrounding intellectual property. Each institution involved in CGrADS has its own policies regarding the ownership, disclosure, and licensing of intellectual property.

Inventions The policies of most of the institutions participating in CGrADS give ultimate ownership of inventions to the institution, with some compensation to the inventor. However, the details differ widely between the eight participating institutions. These complications prevent us from making strong blanket statements about rights in inventions. However, we have agreed to the following principles:

1. Any invention developed solely by employees of a participating institution shall be governed by the intellectual property policies of that institution.
2. Any invention developed collaboratively by personnel from multiple institutions shall be governed by a cooperative agreement between the developer's institutions. Such agreements will be negotiated as the need for them arises.
3. All the institutions involved in CGrADS agree, in advance, to defer such negotiations and to negotiate, in good faith and in a timely manner, when the need arises.

The institutions participating in CGrADS have a history of negotiating inter-institutional agreements for sharing technology and intellectual property.

Software The investigators believe strongly in the value of open source approaches as a means of achieving broad dissemination of the results of publicly-funded research. To the extent possible under institutional policies, we will work to achieve liberal policies for access to CGrADS-developed software. The Consortium for Open Grid Software may prove to be an effective body for disseminating such software.

12 Shared Experimental Facilities

The CGrADS budget does not include any support for the purchase of dedicated, shared facilities. Given the importance of empirical studies in actual and simulated Grid environments, however, CGrADS will establish both a distributed Grid testbed, called the *Macrogrid* and a set of Grid simulation tools, called the *MicroGrid*.

The MacroGrid will be assembled from resources provided by CGrADS institutions and other partners. These resources will be funded through cost-sharing on CGrADS and funds from external sources. It will be available for use by CGrADS researchers and the application partners associated with CGrADS. The Macrogrid resources are distributed among the CGrADS home institutions. The initial Macrogrid consists of the following hardware resources (listed by institution).

Institution	CPU Type	O.S. Type
Rice	IBM SP2 (24 nodes)	AIX
UCSD	Pentium Cluster (16 nodes)	Debian Linux
UofC	Pentium Cluster (16 nodes)	SUSE Linux
UH	IBM SP2 (64 nodes)	AIX
UH	Pentium III Cluster (8 nodes) + Xeon Cluster (3 nodes)	Linux
UIUC	Pentium Cluster (24 nodes)	Red Hat Linux
USC/ISI	SGI Origin 2000 (10 nodes)	IRIX
USC/ISI	SGI Origin 2000 (64 nodes)	IRIX
USC/ISI	Pentium III Cluster	Linux 6.2
UTK	Pentium Cluster (8 nodes)	Red Hat Linux
UTK	Pentium Cluster (16 nodes)	Debian Linux

Our expectation is that the Macrogrid testbed will grow, both in terms of computing capability and resource heterogeneity, as the project proceeds. However, because there is no funding in the CGrADS budget for this activity, we cannot provide any specific assurances as to the resources that will be available. In addition, the National Partnership for Advanced Computational Infrastructure (NPACI) and the National Computational Science Alliance (NCSA) will provide access to additional resources, as shown in their letters of support.

The MicroGrid is a collection of software that is easily deployed on a single PC or a cluster of PCs. It allows simulation of Grid-enabled applications in a closely controlled environment where the behavior of processing elements, networks, and storage resources can be monitored and studied. The MicroGrid is an important facility for CGrADS research, and, because it is software, an easily replicated resource. Initially, we will provide the MicroGrid tools to our application partners. As the project progresses, the capabilities of the MicroGrid Software will increase, and successive releases of the software will be made available.

13 Committed Funding by Source

	Year 1		Year 2		Year 3	
Committed Funding	*Cash*	*In-Kind*	*Cash*	*In-Kind*	*Cash*	*In-Kind*
NSF	3,804,082	0	3,810,285	0	3,887,235	0
Industry	0	0	0	0	0	0
Academic Institution	355,347	790,935	431,335	712,405	437,055	728,365
Other Government	0	0	0	0	0	0
Other	0	0	0	0	0	0
Total	4,159,429	790,935	4,241,620	712,405	4,324,290	728,365

	Year 4		Year 5	
Committed Funding	*Cash*	*In-Kind*	*Cash*	*In-Kind*
NSF	3,934,189	0	3,998,578	0
Industry	0	0	0	0
Academic Institution	439,094	738,758	605,479	591,538
Other Government	0	0	0	0
Other	0	0	0	0
Total	4,373,283	738,758	5,195,595	591,538

14 Institutional and Other Sector Support

The CGrADS project has a team of thirteen PIs from seven institutions. Each of these institutions has enthusiastically endorsed this project as an effort that will have a significant and positive impact on the progress of computational science, as well as on the education of their own students. Each institution also has a demonstrated history of supporting progress in the computational sciences, in improving the diversity of student participation, in increasing the participation of American students, and in providing space and facilities to enhance interdisciplinary collaboration in research and education. Each institution has agreed to cost share 30% of the budget amount requested by their institution in support of the project efforts.

Cost sharing is composed of institutional commitments in cash and in kind. Approximately 32% of the cost share will be provided as high-end computer equipment or as IT and networking support for the project's Microgrid and Macrogrid infrastructure. More than 28% of the cost share (approximately $1.67 million) will take the form of direct support for the education and human-resource development efforts of the Center. This includes graduate-student stipends, support for undergraduates, and funding for programs designed to increase the number of women and minorities in the computational sciences. Discretionary funds will account for approximately 7.5% of the cost share. Such funds are particularly valuable because they allow the Center to take advantage of new opportunities that arise during the course of the project. Other support comes in the form of extra travel funds, faculty salaries, administrative salaries, reduced overhead, and student tuition waivers. It is important to note that over 40% of the total cost share will be provided in cash. This percentage is significant; it provides a good indication of the institutions' strong commitment to CGrADS.

In addition to these direct cost-sharing contributions, each institution will also provide computing cycles on some of its leading-edge machines. (See § 12.) These cycles will be used to create a testbed for Center research activities and the prototype software systems that CGrADS will produce. Most of the PIs are involved in one or both of the NSF PACI programs (NCSA, based at the University of Illinois in Urbana-Champaign, and NPACI, based at the University of California in San Diego). Because the PACIs cannot directly support research, both PACI programs have shown a strong interest in CGrADS. As one measure of this interest, the PACIs will make computational resources available to the project, as indicated in their letters of commitment.

The CGrADS principal investigators are each nationally recognized in their own right. The PIs will continue to seek funding for projects that are complementary to the CGrADS effort, both on individually and as part of larger collaborative groups. These funds and related projects will leverage the efforts of the proposed STC project. All of the PIs are already involved in such efforts. These include:

- *The GRADS Project:* GrADS is funded by the NSF Next Generation Software program through 2002. All of the CGrADS PIs are members of GrADS.
- *The PACI Programs:* Every PI participates in one or both PACI programs. The PACIs provide a natural technology transfer venue for the project.
- *The Los Alamos Computer Science Institute (LACSI):* LACSI is funded by the DOE through Los Alamos National Laboratory. The following PIs are involved in LACSI: Ken Kennedy, Keith Cooper, Jack Dongarra, Lennart Johnsson, John Mellor-Crummey, Richard Tapia, Dan Reed, and Linda Torczon. The LACSI research program has components that complement the proposed CGrADS research.

As the GrADS software matures, we expect that additional collaborations will form around new problems exposed by our experience with Grid-aware applications.

References

[1] R. Aiken, M. Carey, B. Carpenter, I. Foster, C. Lynch, J. Mambretti, R. Moore, J. Strasnner, and B. Teitelbaum. Network Policy and Services: A Report of a Workshop on Middleware. IETF, RFC 2768, http://www.ietf.org/rfc/rfc2768.txt, 2000.

[2] H. Casanova, J. Dongarra, C. Johnson, and M. Miller. Application–Specific Tools. In I. Foster and C. Kesselman, editors, *The Grid: Blueprint for a New Computing Infrastructure*, pages 159–180. Morgan Kaufmann, 1998.

[3] CDDB, `http://www.cddb.com`.

[4] S. Chauveau and F. Bodin. Menhir: An environment for high performance Matlab. *Scientific Programming*, 7:303–312, 1999.

[5] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and S. Tuecke. The Data Grid: Towards and Architecture for the Distributed Management and Analysis of Large Scientific Data Sets. *Journal of Network and Computer Applications*, 2000.

[6] L. Childers, T. Disz, M. Hereld, R. Hudson, I. Judson, R. Olson, M. E. Papka, J. Paris, and R. Stevens. ActiveSpaces on the Grid: The Construction of Advanced Visualization and Interaction Environments. In M. Griebel, D. E. Keyes, R. M. Nieminen, D. Roose, and T. Schlick, editors, *Simulation and Visualization on the Grid: Lecture Notes in Computational Science and Engineering 13*, pages 64–80, Berlin, 2000. Springer-Verlag.

[7] K. Czajkowski, S. Fitzgerald, I. Foster, and C. Kesselman. Grid Information Services for Distributed Resource Sharing, 2001. Submitted for publication.

[8] K. Czajkowski, I. Foster, N. Karonis, C. Kesselman, S. Martin, W. Smith, and S. Tuecke. A Resource Management Architecture for Metacomputing Systems. In *Proceedings of the Fourth Workshop on Job Scheduling Strategies for Parallel Processing*, 1998.

[9] S. E. Czerwinski, B. Y. Zhao, T. D. Hodes, A. D. Joseph, and R. H. Katz. An Architecture for a Secure Service Discovery Service. In *Mobicom '99*. ACM Press, 1999.

[10] T. DeFanti, I. Foster, M. Papka, R. Stevens, and T. Kuhfuss. Overview of the I–WAY: Wide–Area Visual Supercomputing. *The International Journal of Supercomputer Applications and High Performance Computing*, 10(2):123–130, Summer/Fall 1996.

[11] L. DeRose and D. Padua. A MATLAB to Fortran 90 translator and its effectiveness. In *Proceedings of the 10th International Conference on Supercomputing*, May 1996.

[12] P. Dinda and D. O'Hallaron. An Evaluation of Linear Models for Host Load Prediction. In *Proceedings of the 8th IEEE Symposium on High-Performance Distributed Computing (HPDC)*. IEEE Press, 1999.

[13] A. Downey. Predicting Queue Times on Space-Sharing Parallel Computers. In *Proceedings of the International Parallel Processing Symposium*, 1997.

[14] Entropia, `http://www.entropia.com`.

[15] S. Fitzgerald, I. Foster, C. Kesselman, G. von Laszewski, W. Smith, and S. Tuecke. A Directory Service for Configuring High–Performance Distributed Computations. In *Proceedings of the Sixth IEEE Symposium on High–Performance Distributed Computing*, pages 365–375, Aug. 1997.

[16] I. Foster and C. Kesselman, editors. *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann, 1998.

[17] I. Foster and C. Kesselman. The Globus Toolkit. In I. Foster and C. Kesselman, editors, *The Grid: Blueprint for a New Computing Infrastructure*, pages 259–278. Morgan Kaufmann, 1998.

[18] I. Foster, C. Kesselman, C. Lee, R. Lindell, K. Nahrstedt, and A. Roy. A Distributed Resource Management Architecture that Supports Advance Reservations and Co-Allocation. In *Proceedings of the International Workshop on Quality of Service*, pages 27–36, 1999.

[19] I. Foster, C. Kesselman, and S. Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International Journal of Supercomputer Applications*, 2001.

[20] E. Gabriel, M. Resch, T. Beisel, and R. Keller. Distributed Computing in a Heterogenous Computing Environment. In *Proc. EuroPVMMPI'98*. 1998.

[21] D. Gannon, R. Bramley, M. Govindaraju, N. Mukhi, M. Yechuri, and B. Temko. A Componentized Services Architecture for Building Distributed Grid Applications. In *Proceedings of Ninth IEEE International Symposium on High Performance Distributed Computing*, Pittsburgh, Pennsylvania, Aug. 2000.

[22] A. S. Grimshaw, W. A. Wulf, and the Legion Team. The Legion Vision of a Worldwide Virtual Computer. *Communications of the ACM*, 40(1), 1997.

[23] Grid Physics Network (GriPhyN) Project, `http://www.griphyn.com`.

[24] E. Guttman, C. Perkins, J. Veizades, and M. Day. Service Location Protocol, Version 2. IETF RFC 2165, 1998.

[25] S. Guyer and C. Lin. An annotation language for optimizing software libraries. In *Proceedings of the Second Conference on Domain-Specific Languages*, Oct. 1999.

[26] B. Hahn. *Essential MATLAB for Scientists and Engineers*. Arnold, 1997.

[27] T. A. Howes and M. Smith. A Scalable, Deployable Directory Service Framework for the Internet. Technical Report 95-7, Center for Information Technology Integration, Univerity of Michigan, 1995.

[28] W. E. Johnston, D. Gannon, and B. Nitzberg. Grids as Production Computing Environments: The Engineering Aspects of NASA's Information Power Grid. In *Proceedings of the 8th IEEE Symposium on High-Performance Distributed Computing (HPDC)*. IEEE Computer Society Press, 1999.

[29] N. H. Kapadia, J. A. B. Fortes, and C. E. Brodley. Predictive Application-Performance Modeling in a Computational Grid Environment. In *Proceedings of the 8th IEEE Symposium on High-Performance Distributed Computing (HPDC)*, Aug. 1999.

[30] K. Kennedy, B. Broom, K. Cooper, J. Dongarra, R. Fowler, D. Gannon, L. Johnsson, J. Mellor-Crummey, and L. Torczon. Telescoping Languages: A Strategy for Automatic Generation of Scientific Problem-Solving Systems from Annotated Libraries. *Journal of Parallel and Distributed Computing*, 2001. To appear.

[31] Khoral Software. *Khoros Pro Version 2.2*, 1998.

[32] T. Kimura and H. Takemiya. Local Area Metacomputing for Multidisciplinary Problems: A Case Study for Fluid/Structure Coupled Simulation. In *Proc. Intl. Conf. on Supercomputing*, pages 145–156. 1998.

[33] M. Lutz. *Programming Python*. O'Reilly & Associates, Oct. 1996.

[34] P. Lyster, L. Bergman, P. Li, D. Stanfill, B. Crippe, R. Blom, C. Pardo, and D. Okaya. CASA Gigabit Supercomputing Network: CALCRUST Three–Dimensional Real–Time Multi–Dataset Rendering. In *Proceedings of Supercomputing '92*, Minneapolis, Minnesota, Nov. 1992. (Poster session.).

[35] V. Menon and K. Pingali. A case for source-level transformations in MATLAB. In *Proceedings of the Second Conference on Domain-Specific Languages*, pages 53–65, Oct. 1999.

[36] V. Menon and K. Pingali. High-level semantic optimization of numerical codes. In *Proceedings of the International Conference on Supercomputing 1999*, pages 434–443, 1999.

[37] Napster, `http://www.napster.com`.

[38] D. Narayanan, J. Flinn, and M. Satyanarayanan. Using History to Improve Mobile Application Adaptation. In *Proceedings of the Third Workshop on Mobile Computing Systems and Applications*, Dec. 2000.

[39] Parabon, `http://www.parabon.com`.

[40] D. Reed, C. Elford, T. Madhyastha, E. Smirni, and S. Lamm. The Next Frontier: Interactive and Closed Loop Performance Steering. In *Proceedings of the 1996 International Conference on Parallel Processing Workshop*, pages 20–31, Aug. 1996.

[41] D. Reed and R. L. Ribler. Performance Analysis and Visualization. In I. Foster and C. Kesselman, editors, *The Grid: Blueprint for a New Computing Infrastructure*, pages 367–394. Morgan Kaufmann, 1998.

[42] R. Ribler and D. Reed. The Autopilot Performance–Directed Adaptive Control System. In *Proceedings of 11th ACM International Conference on Supercomputing—Workshop on Performance Data Mining: Automated Diagnosis, Adaption and Optimization*, Vienna, Austria, July 1997.

[43] T. Sheehan, W. Shelton, T. Pratt, P. Papadopoulos, P. LoCascio, and T. Dunigan. Locally Self Consistent Multiple Scattering Method in a Geographically Distributed Linked MPP Environment. *Parallel Computing*, 24, 1998.

[44] S. Smale. Dynamics in General Equilibrium Theory. *American Economic Review*, 66(2):284–294, May 1976.

[45] W. Smith, I. Foster, and V. Taylor. Predicting Application Run Times Using Historical Information. In *The 4th Workshop on Job Scheduling Strategies for Parallel Processing*, 1998.

[46] R. Stevens, P. Woodward, T. DeFanti, and C. Catlett. From the I–WAY to the National Technology Grid. *Communications of the ACM*, 40(11):50–60, Nov. 1997.

[47] J. Subhlok, P. Lieu, and B. Lowekamp. Automatic Node Selection for High Performance Applications on Networks. In *Proceedings of the Seventh ACM SIGPLAN Symposium on the Principles and Practice of Parallel Programming (PPoPP'99)*, pages 163–172. ACM Press, 1999.

[48] T. Tannenbaum and M. Litzkow. The Condor Distributed Processing System. *Dr. Dobbs Journal*, Feb. 1995.

[49] M. van Steen, F. Hauck, P. Homburg, and A. Tanenbaum. Location Objects in Wide-area Systems. *IEEE Communications Magazine*, pages 104–109, 1998.

[50] J. Waldo. The Jini Architecture for Network-centric Computing. *Communications of the ACM*, 42(7):76–82, July 1999.

[51] C. A. Waldspurger, T. Hogg, B. A. Huberman, J. O. Kephart, and W. S. Stornetta. Spawn: A Distributed Computational Economy. *IEEE Trans. on Software Engineering*, 18(2):103–117, Feb. 1992.

[52] G. E. Weaver, K. S. McKinley, and C. C. Weems. Score: A compiler representation for heterogeneous systems. In *Proceedings of the 1996 Heterogeneous Computing Workshop*, Honolulu, Apr. 1996.

[53] R. Wolski. Dynamically Forecasting Network Performance to Support Dynamic Scheduling Using the Network Weather Service. In *Proceedings of the Sixth IEEE Symposium on High–Performance Distributed Computing*, Aug. 1997.

Ken Kennedy

Professional Preparation:

B.A., Rice University, 1967 (Mathematics, Summa Cum Laude)
M.S., New York University, 1969 (Mathematics)
Ph.D., New York University, 1971 (Computer Science)

Appointments:

1999-	Director, Center for High Performance Software Research, Rice University
1997-	Ann and John Doerr Professor, Dept. of Computer Science, Rice University
1997-	Distinguished Adjunct Professor, Computer Science Dept., University of Houston
1990-1992	Chair, Dept. of Computer Science, Rice University
1989-2000	Director, Center for Research on Parallel Computation, Rice University
1986-1992	Director, Computer and Information Technology Institute, Rice University
1985-1986	Visiting Professor, Dept. of Computer Science, Stanford University
1985-1997	Noah Harding Professor, Dept. of Computer Science, Rice University
1984-1988	Chair, Dept. of Computer Science, Rice University
1982-1984	Chair, Computer Science Program Committee, Rice University
1980-1984	Professor, Dept. of Mathematical Sciences, Rice University
1976-1980	Associate Professor, Dept. of Mathematical Sciences, Rice University
1971-1976	Assistant Professor, Dept. of Mathematical Sciences, Rice University

Synergistic Activities:

Dr. Kennedy's research interests include parallel computing, scientific programming environments, and optimization of compiled code. He has published over one hundred fifty technical articles and supervised thirty-four Ph.D. dissertations on programming support software for high-performance computer systems. He has supervised the construction of several software systems for programming parallel machines, including an automatic vectorizer for Fortran 77, an integrated scientific programming environment with support for parallel program construction and debugging, and a compiler infrastructure for Fortran 90 and High Performance Fortran. His current research focuses on programming tools for parallel computer systems and high-performance microprocessors, seeking to develop new strategies for supporting architecture-independent parallel programming, especially in science and engineering. He currently directs the GrADS Project, a collaborative eight-institution research effort started in 1999 with NSF support, which is focused on application development support for computational grids. He is also the director of the academic component of the Los Alamos Computer Science Institute, a collaboration with Los Alamos National Laboratory.

Professor Kennedy was the founding chair of the Rice Computer Science Department (1984) and Rice Computer and Information Technology Institute (1987). In 1989, he established the Center for Research on Parallel Computation (CRPC), a NSF Science and Technology Center, and directed it throughout its lifetime. He is a member of the National Academy of Engineering (1990) and a Fellow of the AAAS (1994), the ACM (1995) and the IEEE (1995). In recognition of his achievements in compilation for high performance computer systems, he received the 1995 W. W. McDowell Award, the highest research award of the IEEE Computer Society. In 1999, received the ACM SIGPLAN Programming Languages Achievement Award, the third time this award was given.

Professor Kennedy's service to the national community includes time as Co-chair (1997-99) of the President's Information Technology Advisory Committee (PITAC), a member of the Commission on Physical Sciences, Mathematics, and Applications (1995-97), a member of the Advisory Committee for the NSF CISE Directorate (1995-97), Chair of the High Performance Fortran Forum (1992-97), Chair of the NRC Workshop Series on High Performance Computing and Communications (1993-96). He is also on the Editorial Boards of three major journals. His service has been recognized through two national awards: the CRA Distinguished Service Award (1999) and the RCI Seymour Cray HPCC Industry Recognition Award (1999)

Related Publications:

1. "The ParaScope Parallel Programming Environment," (with K. D. Cooper, M. W. Hall, R. T. Hood, K. McKinley, J. Mellor-Crummey, L. Torczon, and S. K. Warren. *Proceedings of the IEEE* 81(2): 244-263, February 1993.
2. "Interprocedural Analysis and Optimization," (with K. D. Cooper, M. W. Hall, L. Torczon). The *Communications in Pure and Applied Mathematics* 48: 947-1003, 1996.
3. Compilers, Languages, and Libraries," In *The Grid: Blueprint for a New Computing Infrastructure*, (I. Foster and C. Kesselman, editors), Morgan Kaufmann Publishers, Inc., 181- 204, August 1998.
4. "Telescoping Languages: A Compiler Strategy for Implementation of High-level Domain- specific Programming Systems," *Proceedings of the International Parallel and Distributed Processing Symposium 2000*, Cancun, Mexico, May 2000.
5. *Advanced Compilation for High Performance Computing* (with R. Allen), Morgan Kaufmann, San Francisco, CA, *to appear September 2001.*

Other Significant Publications:

1. "Improving Register Allocation for Subscripted Variables," (with D. Callahan and S. Carr). *Proceedings of the ACM SIGPLAN '90 Conference on Programming Language Design and Implementation*, *SIGPLAN Notices* 25(6): 53-65, June 1990.
2. "Improving Cache Performance in Dynamic Applications through Data and Computation Reorganization at Run Time," (with C. Ding). *Proceedings of the ACM SIGPLAN '99 Conference on Programming Language Design and Implementation*, Atlanta, GA, May 1999.
3. "Improving Memory Hierarchy Performance for Irregular Applications," (with J. Mellor- Crummey and D. Whalley). *Proceedings of the 13th ACM International Conference on Supercomputing*, Rhodes, Greece, 425-433, June 1999.
4. "Greedy Weighted Fusion," *Proceedings of International Conference on Supercomputing 2000*, Santa Fe, NM (May 2000).
5. "The Cost of Being Object-oriented: A Preliminary Study," (with Z. Budimlic and J. Piper). *Scientific Programming* 7(2): 87-95, 1999.

Collaborators:

Current Collaborators: Vikram Adve (UIUC), Francine Berman (UCSD), Alok Choudhary (Northwestern), Jack Dongarra (Tennessee), Ian Foster (Argonne, Chicago), Geoffrey Fox (Syracuse), Dennis Gannon (Indiana), Lennart Johnsson (UH), Carl Kesselman (USC ISI), Chuck Koelbel (NSF), Dan Reed (UIUC), Joel Saltz (Maryland), David Whalley (FSU), Rich Wolski (Tennessee) plus numerous collaborators associated with the following projects: NSF CRPC, NSF NCSA (the Alliance), NSF NPACI, NSF GrADS, and the Los Alamos Computer Science Institute.

Thesis Advisees: Steve Carr (MTU), Chen Ding (Rochester), Ervan Darnell (Oracle), Gina Goff (DoD), Mary Hall (USC ISI), Paul Havlak (UH), Reinhard von Hanxleden (Daimer Benz AG), Ulrich Kremer (Rutgers), Lorie Liebrock (MTU), Nat McIntosh (HP) Kathryn McKinley (UMass), Doug Monk (Trinity Group), Ajay Sethi (Oracle), Jerry Roth (Gonzaga U), and Chau-Wen Tseng (Maryland). (Total over career: 34)

Thesis Advisors: Jacob T. Schwartz, New York University and John Cocke, IBM (retired).

Francine Berman

Professional Preparation:

B.A., UCLA., June, 1973 (Mathematics)
M.A., University of Washington, February, 1976 (Mathematics)
Ph.D., University of Washington, August, 1979 (Mathematics)

Appointments:

February, 2001–present	Director	San Diego Supercomputer Center
February, 2001–present	Director	National Partnership for Advanced Computational Infrastructure
July, 1993–present	Professor	U.C. San Diego
July, 1987–June, 1993	Associate Professor	U.C. San Diego
July, 1984–June, 1987	Assistant Professor	U.C. San Diego
August, 1979–July, 1984	Assistant Professor	Purdue University

Synergistic Activities:

Dr. Berman is a Fellow of the ACM and has served on numerous editorial boars, steering committees, and program and conference committees in the areas of parallel computing and Grid computing. She has been Program Chair of the Heterogeneous Computing Workshop and IPDPS Conference and Program Co-Chair of the High Performance Distribute Computing Conference. A member of the Computing Research Association's Committee on the Status of Women in Computer Science and Engineering (CRA-W) since its inception a decade ago, Dr. Berman has served as co-chair of CRA=W, is a member of its steering committee, and has served as General Chair and Program Chair of the 1999 Careers Workshop for Women.

Related Publications: (*available at* `http://apples.ucsd.edu`)

1. H. Casanova, A. Legrand, Z. Zaogordnov, and F. Berman, 'Heuristics for Scheduling Parameter Sweep Applications in Grid Environments," Proceedings of the 2000 Heterogeneous Computing Workshop.

2. H. Casanova, G. Obertelli, R. Wolski and F. Berman, "The AppLeS Parameter Sweep Template: User-Level Scheduling Middleware for the Grid," Best paper finalist, Proceedings of Supercomputing 2000.

3. Berman, F., "High-Performance Schedulers," book chapter for *The Grid: Blueprint for a New Computing Infrastructure*, edited by Ian Foster and Carl Kesselman, Morgan-Kaufmann, 1999.

4. Beck, M., Casanova, H., Dongarra, J., Moore, T., Planck, J., Berman, F. and Wolski, R., "Logistical Quality of Service in NetSolve," to appear in the Special Issue on Network-Based Parallel and Distributed Computing of *Computer Communications*.

5. S. Smallen, Walfredo Cirne, Jaime Frey, Francine Berman, Rich Wolski, Mei-Hui Su, Carl Kesselman, Steve Young, Mark Ellisman, "Combining Workstations and Supercomputers to Support Grid Applications: The Parallel Tomography Experience," Proceedings of the 2000 Heterogeneous Computing Workshop

Other Significant Publications:

1. W. Cirne and F. Berman, "A Model for Moldable Supercomputer Jobs," Proceedings of IPDPS 2001.

2. F. Berman, R. Wolski, and Z. Zaogordnov, "Application Scheduling on the Information Power Grid," [invited paper] *International Journal of High Performance Computing*, 14(3), Fall 2000.

3. M. Faerman, A. Su, R. Wolski, and F. Berman, "Adaptive Performance Prediction for Distributed Data-Intensive Applications," Proceedings of Supercomputing '99.

4. Berman, F., Wolski, R., Figueira, S., Schopf, J. and Shao, G., "Application-Level Scheduling on Distributed Heterogeneous Networks," Proceedings of Supercomputing '96.

5. Berman, F. and Wolski, R., "Scheduling from the Perspective of the Application," (*invited paper*), Proceedings of the 1996 High-Performance Distributed Computing Conference.

Collaborators:

Micah Beck	Silvia Figueira	Jim Planck	Atsuko Takefusa
Henri Casanova	Jaime Frey	Jennifer Schopf	Steve Young
Walfredo Cirne	Andrew Grimshaw	Gary Shao	Rich Wolski
Holly Dail	Aubin Jarry	Shava Smallen	Dmitrii Zagorodnov
Jack Dongarra	Carl Kesselman	Michelle Strout	
Mark Ellisman	Arnaud LeGrand	Alan Su	
Marcio Faerman	Terry Moore	Mei-Hui Su	

Graduate and Postdoctoral Advisors:
Ph.D. Dissertation advisors: Robert Ritchie (HP) and Michael Fischer (Yale University)

Thesis Advisees (last 5 years):
Leesa Hicks (Creedence Software Engineering), Silvia Figueira (Santa Clara University), Jennifer Schopf (Northwestern University), Walfredo Cirne (Universidade Federal da Paraiba), Gary Shao (will postdoc at UCSD)

Postdoctoral Advisees:
Rich Wolski (University of Tennessee), Henri Casanova (UC San Diego)

Jack J. Dongarra

Education:

Ph.D., University of New Mexico, 1980 (Applied Mathematics)
M.S., Illinois Institute of Technology, 1973 (Computer Science)
B.S., Chicago State University, 1972 (Mathematics)

Positions:

University Distinguished Professor, Department of Computer Science, University of Tennessee (1990-Present)
Distinguished Scientist, Mathematical Sciences, Oak Ridge National Laboratory (1990-2000)
Senior Scientist, Argonne National Laboratory (1980-1990)

Relevant Publications:

1. "A Portable Programming Interface for Performance Evaluation on Modern Processors", S. Browne, J. Dongarra, N. Garner, G. Ho, and P. Mucci, *International Journal of High Performance Computing Applications*, Volume 14, Number 3, Fall 2000, pp 189—204.
2. "Automatically Tuned Linear Algebra Software", Clint Whaley and Jack Dongarra, SC98 Conference, Orlando, FL, November, 1998.
3. "Numerical Linear Algebra Algorithms and Software", Jack Dongarra and Victor Eijkhout, *Journal of Computation and Applied Mathematics*. Volume 123, 2000, pp 489-514.
4. "Algorithmic Issues on Heterogeneous Computing Platforms", Pierre Boulet, Jack Dongarra, Fabrice Rastello, Yves Robert, and Frederic Vivien, *Parallel Processing Letters*, Volume 9, Number 2, pp 197-213, 1999.
5. "Programming Tools and Environments, J. Saltz, A. Sussman", S. Graham, J. Demmel, S. Baden, and J. Dongarra, *Communications of the ACM*, November 1998, Vol. 41, No. 11, pp 64-73, ISSN 0001-0782.

Significant Publications:

1. *MPI – The Complete Reference*, Volume 1, The MPI-1 Core, Second Edition, Marc Snir, Steve Otto, Steven Huss-Lederman, David Walker, Jack Dongarra, MIT Press, September 1998, ISDN 0-262-69215-5.
2. *Numerical Linear Algebra for High-Performance Computers*, Jack Dongarra, Iain Duff, Danny Sorensen, and Henk van der Vorst, SIAM Publication, November 1998.
3. "Using Agent-based Software for Scientific Computing in the NetSolve System", Henri Casanova and Jack Dongarra, *Parallel Computing*, Vol. 24, No. 12-13, November, 1998, pp 1777-1790, ISSN 0167-8191.
4. *ScaLAPACK Users' Guide*, L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley SIAM Publications, Philadelphia, 1997.
5. *LAPACK Users' Guide – Third Edition*, E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, SIAM Publication, Philadelphia, 1999.

Synergistic Activities:

Dongarra specializes in numerical algorithms in linear algebra, parallel computing, use of advanced-computer architectures, programming methodology, and tools for parallel computers. His research includes the development, testing, and documentation of high quality mathematical software. He has contributed to the design and implementation of Open Source software packages EISPACK, LINPACK, the BLAS, LAPACK, ScaLAPACK, Netlib, PVM, MPI, NetSolve, ATLAS, and PAPI all of which have become professional and industrial standards. He has published approximately 200 articles, papers, reports, and technical memoranda, and he is coauthor of several books. He was named a fellow of the AAAS in 1995, of the IEEE in 1999, and of the ACM in 2001, and elected to the National Academy of Engineering in 2001. The NECI Scientific Literature Digital Library lists Dongarra as the fifth most cited author in computer science.

In addition to the position in Tennessee, Dongarra has an Adjunct Professorship in Computer Science at Rice University. He has been a visiting professor or visiting scientist at universities and laboratories throughout the United States and Europe including LANL, IBM T.J. Watson Research Center, the Center for Supercomputer Research and Development at the University of Illinois, Urbana-Champaign, AERE Harwell Laboratory in England, ETH in Zurich, Ecole Normale Superieure de Lyon in France, the Danish Technical University in Lyngby, the Swiss Scientific Computing Center in Lugano, the Australian National University in Canberra, the Royal Institute of Technology in Stockholm, the Computational Lab at Oxford University, and the EPFL in Lausanne.

Selected Professional Activities:

Editor-in-Chief: Netlib (since 1987), International Journal of High-Performance Supercomputer Applications (since 1992), and SIAM Series on Software, Environments, and Tools for Scientific Computing (since 1995)
Editorial Boards: Applied Numerical Mathematics (since 1994), Electronic Transactions on Numerical Analysis (since 1993), International Journal of High Speed Computing (since 1994), Journal of Distributed and Parallel Computing (since 1988), Journal of Numerical Linear Algebra with Applications (since 1994), Journal of Supercomputing (since 1987), Numerical Linear Algebra with Applications (since 1994), Numerical Algorithms (since1994), Parallel Computing (since 1987), and Parallel Processing Letters (since 1993). Chair (1993-95), Message Passing Interface Forum.

Students: Richard Barrett, Henri Casanova, Stan Green, Youngbae Kim, Lorie Liebrock, Robert Manchek, Phil Mucci, James Payne, Antoine Petitet, Rick Phillips, Keith Seymour, Majed Sidani, Paul McMahan, Rick Phillips, Delphy Nypaver, Clint Whaley, Brian Larose, Steve Moulton, Mitchell Duitz, Nathan Garner, George Ho, Ganapathy Raman.

Postgraduate-Scholar Sponsor: Henri Casanova, Jaeyoung Choi, Andy Cleary, Victor Eijkhout, Peter Newton, Antoine Petitet, Roldan Pozo, Erich Strohmaier, Francoise Tisseur, Robert van de Geijn.

Collaborators: Zhaojun Bai (U. of California, Davis), Fran Berman (UCSD), Jim Demmel (UC Berkeley), Iain Duff (Rutherford Lab), Al Geist (ORNL), Sven Hammarling (NAG Ltd.), Ken Kennedy (Rice U), Bo Kågstrom (U. of Umeå), Roldan Pozo (NIST), Yves Robert (ENS, Lyon), Charles Romine (ORNL), Danny Sorensen (Rice), Vaidy Sunderam (Emory), Henk van der Vorst (U. of Utrecht), David Walker (University of Wales).

Ph.D. Thesis Advisor: Cleve Moler, University of New Mexico
M.S. Thesis Advisor: Brian Smith, Illinois Institute of Technology

Ian T. Foster

Professional Preparation:

University of Canterbury, New Zealand	Computer Science	BS (Hons I), 1977-1979
Imperial College, London, England	Computer Science	PhD, 1986-1988
Argonne National Laboratory (ANL)	Computer Science	Postdoc, 1989-1990

Appointments:

1996-present	Associate Professor, Computer Science, UC
1998-present	Senior Scientist, Mathematics and Computer Science Division (MCS), ANL
1999-present	Executive Committee, Computation Institute, UC
1999-present	Associate Director, MCS, ANL
1997-present	Executive Committee, ASCI Flash Center, UC
1999-present	Enabling Technologies Steering Committee, National Computational Science Alliance
1992-1998	Scientist, MCS, ANL
1990-1992	Assistant Scientist, MCS, ANL

Synergistic Activities:

- *Pedagogical:* Development of widely used texts: *Strand: New Concepts in Parallel Programming* (Prentice Hall, 1990), *Designing and Building Parallel Programs* (www.mcs.anl.gov/dbpp, Addison-Wesley, 1995), and *The Grid: Blueprint for a Future Computing Infrastructure* (www.mkp.com/grid, Morgan Kaufmann, 1999); also, teaching of numerous related tutorials.
- *Research tools:* Development of software tools and systems that have seen extensive use in research and teaching, including: Program Composition Notation compiler and runtime system, Parallel Spectral Transform Shallow Water Model, Nexus communication library (www.mcs.anl.gov/nexus), Globus distributed computing toolkit (www.globus.org).
- *Service:* Including: numerous program committees and review committees; chair of numerous technical workshops and conferences, including IEEE HPDC '98, IEEE Frontiers '99; Software Architect, 1995 I-WAY, ACM/IEEE SC '95; SC 'XY steering committee, 1999-present; Editorial Board, IEEE Trans. on Parallel and Distributed Systems, 1997-present.
- *Leadership:* Convenor of the Grid Forum, an international organization focused on defining standards and best practices in "Grid" computing (www.gridforum.org).
- *Awards and Honors:* British Computer Society Award for Technical Innovation, 1989; Best paper award, IEEE/ACM Supercomputing '95 Conference, 1995; GII Next Generation Award, 1997.

Related Publications:

1. I. Foster, C. Kesselman, "Globus: A Metacomputing Infrastructure Toolkit," Intl J. Supercomputer Applications, 11(2):115-118, 1997.
2. I. Foster, C. Kesselman, C. Lee, R. Lindell, K. Nahrstedt, A. Roy, "A Distributed Resource Management Architecture that Supports Advance Reservations and Co-Allocation," Proc. Intl. Workshop on Quality of Service, 27-36, 1999.
3. I. Foster, C. Kesselman, G. Tsudik, S. Tuecke, "A Security Architecture for Computational Grids," ACM Conference on Computers and Security, 83-91, 1998.
4. I. Foster, J. Insley, C. Kesselman, G. von Laszewski, M. Thiebaux, "Distance Visualization: Data Exploration on the Grid," IEEE Computer, December 1999.
5. I. Foster, J. Geisler, C. Kesselman, S. Tuecke, "Managing Multiple Communication Methods in High-Performance Networked Computing Systems," J. Parallel and Distributed Computing, 40:35- 48, 1997.

Other Significant Publications:

1. K. Czajkowski, I. Foster, C. Kesselman, "Resource Co-Allocation in Computational Grids," Proc. 8th IEEE Symp. on High Performance Distributed Computing, IEEE, 1999.

2. I. Foster and N. Karonis, "A Grid-Enabled MPI: Message Passing in Heterogeneous Distributed Computing Systems," Proc. SC'98, 1998 (CD ROM).

3. I. Foster, J. Geisler, W. Nickless, W. Smith, S. Tuecke, "Software Infrastructure for the I-WAY Metacomputing Experiment," Concurrency: Practice and Experience, 10(7):567–581, 1998.

4. I. Foster, P. Worley, "Parallel Algorithms for the Spectral Transform Method," SIAM J. Scientific Computing, 18(3), 1997.

5. I. Foster, "Compositional Parallel Programming Languages," ACM Trans. Prog. Lang. Syst., 18(4):454–476, 1996.

Collaborators:

Current Collaborators: David Abramson (Monash); Deb Agarwal (LBNL); Richard Alkire (UIUC); Bruce Allen (Wisc- Milwaukee); John Anderson (ILM); Paul Avery (Florida); Ruth Aydt (UIUC); Ray Bair (PNNL); Steve Barnard (NASA Ames); Fran Berman (UCSD); Robert Biswas (NASA Ames); Maxine Brown (UIC); Randy Butler (NCSA); Charlie Catlett (ANL/Chicago); David Ceperley (NCSA); Mani Chandy (Caltech); Ann Chervenak (USC); Andrew Chien (UCSD); Alok Choudhary (NWU); Karl Czajkowski (USC); H. Dachsel (PNNL); Terrence Disz (ANL); Jack Dongarra (Tennessee); John Drake (ORNL); Steve Fitzgerald (USC); Dennis Gannon (Indiana); Jonathan Geisler (NWU); Bill Gropp (ANL); Steve Hammond (NCAR); Robert Harrison (PNNL); Bill Hibbard (Wisc-Madison); Bob Hollebeck (Pennsylvania); Rob Jacob (Chicago); Chris Johnson (Utah); Andrew Johnson (UIC); Lennart Johnsson (Houston); Nancy Johnston (LBNL); Nick Karonis (NIU); Ricky Kendall (Ames); Ken Kennedy (Rice); Steven Kent (Chicago); Carl Kesselman (USC); Dave Kohr (SGI); Rakesh Krishnaiyer (Intel); Tim Kuhfuss (ANL); Stephen Lau (LBNL); Craig Lee (Aerospace); Jason Leigh (UIC); Kai Li (Princeton); Bob Lindell (USC); Miron Livny (Wisc-Madison); Stu Loken (LBNL); Bob Lucas (LBNL); Rusty Lusk (ANL/Chicago); Andrea Malagoli (Chicago); Joe Mambretti (Northwestern); Rick McMullen (Indiana); Michael McRobbie (Indiana); Reagan Moore (SDSC); Richard Mount (SLAC); Klara Nahrstedt (UIUC); Cliff Neuman (USC); Harvey Newman (Caltech); Jarek Nieplocha (PNNL); Jason Novotny (NCSA); Mike Papka (ANL/Chicago); Manish Parashar (Rutgers); Larry Price (ANL); Dan Reed (UIUC); John Reynders (LANL); Bob Rosner (Chicago); Alain Roy (Chicago); Subhash Saini (NASA Ames); Chuck Salisbury (ANL); Volker Sander (Juelich); Ed Seidel (MPI); Larry Smarr (NCSA); John Shalf (NCSA); Arie Shoshani (LBNL); Rok Sosic (Active Tools); Paul Stelling (Aerospace); Rick Stevens (ANL/Chicago); Wai-Mo Suen (WUStL); Valerie Taylor (NWU); George Thiruvathukal (Loyola); Brian Tierney (LBNL); Michael Tobis (Wisc-Madison); Brian Toonen (ANL); Gene Tsudik (Irvine); Steven Tuecke (ANL); Dean Williams (LLNL); Rich Wolski (Tennessee); Patrick Worley (ORNL); Rob Van der Wijngaart (NASA Ames); Gregor von Laszewski (ANL); Maurice Yarrow (NASA Ames); Lou Zechter (NASA Ames).

Advisees and Postgraduate Scholars: (2 Ph.D., 6 Postdoc) Rakesh Krishnaiyer (Intel), Ravi Nanjundiah (Indian Institute of Science), Juan Restrepo (U. Arizona), Michael Tobis (U. Wisconsin), Warren Smith (NASA Ames), George Thiruvathukal (Loyola), Gregor von Laszewski (Argonne), Ming Xu (Platform Computing).

Thesis Advisor: Keith Clark (Imperial College, London)

Daniel A. Reed

Professional Preparation:

University of Missouri at Rolla	Computer Science	B.S. (summa cum laude), 1978
Purdue University	Computer Science	M.S., 1980
Purdue University	Computer Science	Ph.D., 1983

Appointments:

Director, National Computational Science Alliance, 2000-
Director, National Center for Supercomputing Applications, 2000-
Head of the Department of Computer Science, University of Illinois at Urbana-Champaign, 1997-2001
Assistant, Associate, and Professor, Department of Computer Science, 1984-
Senior Research Scientist, National Center for Supercomputing Applications (NCSA), 1996-
Assistant Professor, Department of Computer Science, University of North Carolina, 1983-1984

Biographical Sketch:

Dr. Reed is currently Professor and Head of the Department of Computer Science at the University of Illinois at Urbana-Champaign. In addition, he holds a joint appointment as Director of the National Computational Science Alliance and as director of the National Center for Supercomputing Applications (NCSA). He is also a member of the Caltech Facility for Simulating the Dynamic Response of Materials.

He is the author of research papers and monographs on algorithms, architectures, and performance evaluation techniques for high-performance computing and virtual environments. He has been a principal in the multi-agency (NSF, DARPA, DOE, and NASA) national Scalable I/O Initiative (SIO) and a member of a collaborative, NSF-funded Grand Challenge group with Caltech to explore the input/output performance of scientific codes using the Pablo instrumentation software. This work led to the forthcoming book Scalable Input/Output: Achieving System Balance, to be published by MIT Press.

With Los Alamos National Laboratory, he organized a series of workshops on performance analysis techniques for scalable parallel systems. The workshop goals have been to understand the technical, economic, and political issues that affect development and deployment of software tools on scalable parallel systems. The latest of these workshops led to the 1996 book Debugging and Performance Tuning for Parallel Computing Systems, which has catalyzed national discussion of HPCC software testing and evaluation.

Professor Reed has served on the editorial board of IEEE Transactions on Software Engineering and IEEE Transactions on Parallel and Distributed Systems. He currently serves on the boards of Concurrency Practice and Experience and Performance Evaluation and Modeling for Computer Systems. He is a past treasurer of the ACM SIGMETRICS special interest group on performance measurement and analysis and a past member of the NSF CISE Advisory Committee.

Synergistic Activities:

- Member: NSF CISE Advisory Committee, 1997-2000
- Board of Directors, Computing Research Association, 1998-present
- Illinois VentureTECH Advisory Committee, appointed by Governor Ryan, 2000-
- Current Editorial Boards: Concurrency: Practice and Experience, Performance Evaluation and Modeling for Computer Systems
- Program chair, Workshop on I/O for Parallel and Distributed Systems, 1998
- Treasurer, ACM SIGMETRICS, SIG on Performance Measurement & Analysis, 1995-1997

Related Publications:

1. D. A. Reed, R. C. Giles, and C. E. Catlett, "Distributed Data and Immersive Collaboration," *Communications of the ACM*, Vol. 40, No. 11, pp. 38-49, November 1997.

2. D. A. Reed and R. L. Ribler, "Performance Analysis and Visualization," in *The Grid: Blueprint for a New Computing Infrastructure*, Ian Foster and Carl Kesselman (eds), Morgan- Kaufmann, August 1998.

3. D. A. Reed, C. L. Elford, T. M. Madhyastha, E. Smirni, and S. L. Lamm, "The Next Frontier: Closed Loop and Interactive Performance Steering," *Proceedings of the 1996 ICPP Workshop on Challenges for Parallel Processing*, August, 1996, pp. 20-31.

4. "Debugging and Performance Tuning for Parallel Computing Systems," (with M. L. Simmons, A. H. Hayes, and J. J. Brown), *IEEE Computer Society Press*, 1996.

5. D. A. Reed, D. A. Padua, I. T. Foster, D. B. Gannon, and B. P. Miller, "Delphi: An Integrated, Language-Directed Performance Prediction, Measurement, and Analysis Environment," *Frontiers '99: The 9th Symposium on the Frontiers of Massively Parallel Computation*, Feb. 1999.

Other Significant Publications:

1. C. L. Mendes and D. A. Reed, "Integrated Compilation and Scalability Analysis for Parallel Systems," *Parallel Architectures and Compilation Techniques (PACT '98)*, Oct. 1998.

2. D. A. Reed, R. A. Aydt, L. DeRose, C. L. Mendes, R. L. Ribler, E. Shaffer, H. Simitci, J. Vetter, D. R. Wells, S. Whitmore, and Y. Zhang, "Performance Analysis of Parallel Systems: Approaches and Open Problems," *Japan JSPP*, pp. 239–256, June 1998.

3. T. M. Madhyastha and D. A. Reed, "Input/Output Access Pattern Classification Using Hidden Markov Models," *Proceedings of the Workshop on Input/Output in Parallel and Distributed Systems (IOPADS)*, Nov. 1997.

4. T. M. Madhyastha and D. A. Reed, "Intelligent, Adaptive File System Policy Selection," *Proceedings of Frontiers '96*, Oct. 1996.

5. V. S. Adve, J. Mellor-Crummey, M. Anderson, K. Kennedy, J-C. Wang, and D. A. Reed, "An Integrated Compilation and Performance Analysis Environment for Data Parallel Programs," *Supercomputing '95*, Dec. 1995.

Collaborators:

Current Collaborators: R. Aydt (UIUC), A. Chien (UCSD), A. Choudhary (Northwestern), F. Berman (UCSD), J. Dongarra (Tennessee), D. Gannon (Indiana) G. Gibson (CMU), K. Kennedy (Rice), T. Lenard (Caltech), K. Li (Princeton), V. McKoy (Caltech), J. Mellor-Crummey (Rice), P. Messina (Caltech), B. Miller (Wisconsin), L. Peterson (Princeton), J. Pool (Caltech), J. Saltz (Maryland), M. Simmons (SDSC), R. Stevens (ANL), R. Wolski (Tennessee) (List of hundreds of PACI and ASCI collaborators omitted for brevity.)

Graduate and Postdoctoral Advisees: P. E. Crandall (LANL), B. K. Totty (Inktomi), C. L. Mendes (UIUC), T. M. Madhyastha (UCSC), C. L. Elford (Intel), T. Kwan (McKinsey), R. Ribler (Lynchburg College), H. Simitci (Seagate), H. Simitci (Seagate), E. Smirni (William and Mary), L. Tavera (Inktomi), J.-C. Wang (IBM), L. DeRose (IBM), J. Vetter (LLNL)

Thesis Advisor: Herbert D. Schwetman, Purdue University (now a private consultant)

Ruth A. Aydt

Professional Preparation:

B.A., Knox College, 1980 (Mathematics)
M.C.S., University of Illinois at Urbana-Champaign, 1992 (Computer Science)

Appointments:

1996–	Senior Research Programmer, Pablo Group, Department of Computer Science, University of Illinois at Urbana-Champaign
1991–1996	Research Programmer, Pablo Group, Department of Computer Science, University of Illinois at Urbana-Champaign
1988–1991	Research Programmer, Tapestry Project, Department of Computer Science, University of Illinois at Urbana-Champaign
1983–1988	Research Programmer, Computer Research Laboratory, Department of Computer Science, University of Illinois at Urbana-Champaign
1980–1983	Programmer/Analyst, Uniq Computer Corporation, Batavia, IL

Synergistic Activities:

Ms. Aydt's research interests include performance analysis techniques for parallel and distributing computing, scalable input/output, and computational steering. She is especially interested in designing and building elegant software components that can be easily combined to form flexible and powerful performance tools for a variety of computational platforms, and in working with graduate students to help them successfully contribute to large projects in a team setting. She is currently working on the Autopilot project that provides a scalable and extensible infrastructure for real-time adaptive resource control in distributed computing environments. In addition, she contributes to the management of the adaptive filesystem, I/O characterization, and performance instrumentation and visualization efforts within the Pablo group.

Ms. Aydt is currently participating in the NSF-funded GrADS Project, which is focused on application development support for computational grids, in the Los Alamos Computer Science Institute, and in the Global Grid Forum, where she is on the steering committee and active in the performance working group. As a member of the Pablo group, Ms. Aydt has been involved in the multi-agency (NSF, DARPA, DOE and NASA) national Scalable I/O Initiative (SIO), a collaborative NSF-funded Grand Challenge group with Caltech to explore input/output performance of scientific codes using the Pablo instrumentation software, the DOE/ASCI Illinois Center for Simulation of Advanced Rockets, the distributed storage and collaboration-enabling technologies team for the PACI National Computational Science Alliance, and the Center for Research on Parallel Computation (CRPC).

Related Publications: (*see* `http://www-pablo.cs.uiuc.edu/Publications/publications.htm`)

1. L. DeRose, M. Pantano, R. Aydt, E. Shaffer, B. Schaeffer, S. Whitmore, and D. Reed, "An Approach to Immersive Performance Visualization of Parallel and Wide-Area Distributed Applications," Proceedings of the Eighth IEEE International Symposium on High Performance Distributed Computing, pp. 247–254, August 1999.

2. D. A. Reed, R .A. Aydt, L. DeRose, C. L. Mendes, R. L. Ribler, E. Shaffer, H. Simitci, J. Vetter, D. R. Wells, S. Whitmore, and Y. Zhang, "Performance Analysis of Parallel Systems: Approaches and Open Problems," Joint Symposium on Parallel Processing (JSPP), pp. 239–256, June 1998.

3. E. Smirni, R. A. Aydt, A. A. Chien, and D. A. Reed, "I/O Requirements of Scientific Applications: An Evolutionary View," Proceedings of the Fifth IEEE International Symposium on High Performance Distributed Computing, pp. 49–59, August 1996.

4. D. A. Reed, R. A. Aydt, R. J. Noe, P. C. Roth, K. A. Shields, B. Schwartz, and L. F. Tavera, "Scalable Performance Analysis: The Pablo Performance Analysis Environment," Proceedings of the Scalable Parallel Libraries Conference, pp.104–113, October, 1993.

5. D. Reed, J. Arendt, R. Aydt, T. Birkett, D. Jensen, T. Madhystha, B. Nazief, T. Nelson, R. Olson, and B. Totty, "Scalable Performance Environments for Parallel Systems," Proceedings of the Sixth Distributed Memory Computing Conference, pp. 562–569, May 1991.

Collaborators:

Current Collaborators: John Ambrosiano (LANL), Randy Bramley (Indiana), Francine Berman (UCSD), Andrew Chien (UCSD), Luiz DeRose (IBM), Jack Dongarra (Tennessee), Chris Elford (Intel), Ian Foster (Argonne, Chicago), Dennis Gannon (Indiana), Dan Gunter(LBL), Lennart Johnsson (Houston), Carl Kesselman (USC ISI), Darcy Quesnel (ANL), Tara Madhyastha (UCSC), John Mellor-Crummey (Rice), Bart Miller (Wisconsin), Dave Padua (UIUC), Randy Ribler (Lynchburg College), Huseyin Simitci (Seagate), Evgenia Smirni (William and Mary), Warren Smith (NASA), Valerie Taylor (NWU), Brian Tierney(LBL), Jeff Vetter (LLNL), Rich Wolski (Tennessee).

Andrew A. Chien

SAIC Chair Professor, Dept CSE
University of California, San Diego
9500 Gilman Drive, Dept 0114
La Jolla, CA 92093-0114

Email: achien@cs.ucsd.edu
Phone: 858-822-2458
Fax: 858-822-2459
http://www-csag.ucsd.edu

Education:
Ph.D., Massachusetts Institute of Technology, 1990 (Computer Science)
S.M., Massachusetts Institute of Technology, 1986 (Computer Science)
S.B., Massachusetts Institute of Technology, 1984 (Electrical Engineering)

Current Positions:
1998-present SAIC Chair Professor, Dept of Computer Science & Eng, UCSD
Chief Technology Officer, Entropia Inc.
Adjunct Professor, Dept of Computer Science, University of Illinois

Awards and Honors:
National Science Foundation Young Investigator Award 1994; C. W. Gear Outstanding Junior Faculty Award, 1995; Senior Xerox Award for Excellence in Research, 1996

Research Interests:
Parallel computer architecture, object-oriented programming, language implementation (compilers and runtime systems), distributed objects, operating systems, networks, and reconfigurable computing.

Biographical Sketch:
Andrew Chien has authored over seventy publications in the areas of compilers, system software, networks and processor architecture of high-performance systems. Over the past fifteen years, Andrew Chien has been involved in numerous parallel computing architecture and software. As a graduate student at MIT, he participated in Arvind's tagged-token Dataflow Architecture project, and then Dally's J-Machine project, an early low-overhead message passing machine. As a faculty member at the University of Illinois, Andrew pursued a wealth of architecture, networking, and language implementation projects. Andrew is well known for his work on high speed cluster communication -- the Fast Messages (FM) and High Performance Virtual Machines (HPVM) systems which are the basis for a wealth of clustering research and NCSA's Windows NT Clustering efforts. These systems have also been widely disseminated around the world to over 500 academic, national laboratory, and commercial environments.

Andrew's early work on concurrent object-oriented programming systems began at MIT and continued at Illinois, providing key intellectual input to the pC++ project and the parallel software efforts of the Japanese Real World Computing Project (RWC) which led to the HPC++ standard. His Concert project developed a range of compiler and runtime techniques for fine-grained concurrent object-oriented languages and foreshadows aggressive type, cloning, and data structure optimizations which have yet to appear in commercial compilers for languages such as C++ and Java.

Selected Professional Activities
Editorial and Program Committees:
Associate Editor, IEEE Transactions on Parallel and Distributed Computing
Guest Editor, IEEE Computer Special Issue on Network Interfaces, 1998.
Program Chair, ACM Symposium on the Principles and Practice of Parallel Programming, 1999.
Program Vice Chairman, IEEE International Parallel Processing Symposium, 1995.
Member, numerous program committees in compilers, object-oriented systems, computer architecture, etc.

Invited/Keynote/Panel Talks:
SIAM Parallel Processing 2001, USENIX Windows 2000, Cluster 2000, SC'1999, 1999 SIGPLAN PPoPP Symposium, HiPC '98, ADASS '98, NATO Workshop on HPC 1998, HPCA 1998, DOE Scalable Clusters Workshop 1997, ONR PCRCW 1997, IPPS 1997, SIAM Parallel Processing, 1997, Frontiers 1996, IWPC++ 1996, IBM CASCON 1995, IPPS 1995, Frontiers 1995, ONR PCRCW 1994, Hot Interconnects 1993.

Five Closely Related Publications:
H. Song, X. Liu, D. Jakobsen, X. Zhang, K. Taura, and A. Chien, The MicroGrid: A Scientific Tool for Modeling Computational Grids, finalist for Best Paper Award, SC'2000, Dallas, Texas, November 2000.
J. Dolby and Andrew A. Chien, Automatic Inlining Optimization and its Evaluation}, SIGPLAN Symposium on Programming Language Design and Implementation, Vancouver, British Columbia, June 2000.
B. Ganguly and Andrew A. Chien, High-Level Parallel Programming of An Adaptive Mesh Application Using the Illinois Concert System, in International Symposium on Computing in Object-Oriented Parallel Environments, Sante Fe, New Mexico, December 8-11, 1998.
V. Karamcheti and Andrew A. Chien, A Hierarchical Load-Balancing Framework for Dynamic Multithreaded Computations, in SC '98: High Performance Networking and Computing Conference, November 1998, Orlando, Florida.
V. Karamcheti, J. Plevyak, and A. A. Chien, Runtime Mechanisms for Efficient Dynamic Multithreading, in the Journal of Parallel and Distributed Computing, Volume 37, pgs 21-40.

Five Other Significant Publications:
M. Lauria, S. Pakin, and Andrew A. Chien, Efficient Layering for High Speed Communication: Fast Messages 2.x, in Proceedings of the Seventh High Performance Distributed Computing Conference (HPDC'7), Chicago, Illinois, July, 1998.
J. Dolby and Andrew A. Chien, An Evaluation of Object Inline Allocation Techniques, in the Proceedings of the 1998 Object-oriented Programming, Systems, Languages, and Architectures Conference, (OOPSLA '98), October 1998, Vancouver, British Columbia, Canada.
Scott Pakin, Vijay Karamcheti, and Andrew A. Chien, Fast Messages: Efficient, Portable Communication for Workstation Clusters and Massively-Parallel Processors, in IEEE Concurrency, 1997.
Mario Lauria and Andrew A. Chien, MPI-FM: A High Performance MPI on Workstation Clusters, Journal of Parallel and Distributed Computing, Volume 40, Number 1, January 1997.
Andrew A. Chien, Julian Dolby, Bishwaroop Ganguly, Vijay Karamcheti, and Xingbin Zhang, Evaluating High Level Parallel Programming Support for Irregular Applications in ICC++, in International Scientific Computing in Object-oriented Parallel Environments Conference (ISCOPE), Marina del Rey, December 1997, Springer-Verlag LNCS.

Collaborators:
Other PI's on this proposal, Co-authors of the ``Grids'' book.
Vijay Karamcheti (NYU), Dennis Gannon (Indiana), Reagan Moore (SDSC/UCSD), David Culler (UC Berkeley), Rajesh Gupta (UCI), Alex Nicolau (UCI), Nikil Dutt (UCI), Bishwaroop Ganguly (MIT Lincoln Laboratories), William Weihl (Compaq Research), Matt Buchanan (Compaq Computer), Jae Kim (Hal Computer).
Jane Liu, Wen-mei Hwu, Klara Nahrstedt, David Padua, and Weng Chew (Illinois)

Graduate and Post Doctoral Advisors:
Ph.D. Thesis Advisor: Professor William J. Dally (MIT => Stanford)
S.M. Thesis Advisor: Arvind (Massachusetts Institute of Technology)

Thesis Advisor and Postgraduate-Scholar Sponsor:
Andrew has supervised 15 Master's students, 6 Ph.D. students, and 4 postdoctoral fellows. These alumni are employed variously as faculty at research universities, corporate research laboratories, and startup companies. He currently supervises a group with ten graduate students.

Keith D. Cooper

Professional Preparation:

B.S., Rice University, 1978 (Electrical Engineering)
M.A., Rice University, 1982 (Mathematical Sciences)
Ph.D., Rice University, 1983 (Mathematical Sciences)

Appointments:

Professor, Department of Computer Science and Department of Electrical and Computer Engineering, Rice University, July 2000
Associate Professor, Department of Computer Science, Rice University, June 1990 to present
Research Scientist, Department of Computer Science, Rice University, 1983 to June 1990

Synergistic Activities:

Dr. Cooper's research group develops new compiler techniques and transfers them to both industry and academia. Methods developed in the group appear in compilers from Bops, Compaq, Ericsson, Hewlett-Packard, Ibm, Intel, Silicon Graphics, Sun Microsystems, and Texas Instruments. Most of his Ph.D.s accept positions in industry; he has placed students at Cray (Tera, head of the compiler group), Sun (head of the Hotspot compiler group), TI, Hewlett-Packard, LSI Logic, Dell, and Bops (manager of compiler development). The research compiler tools developed in the group have been used in both industry and academia, including studies at Sequent Computer, DEC (now Compaq), Michigan Technological University, University of Delaware, and Clemson. The implementations have been used in classes at CalTech and the University of Massachusetts.

The second major focus of Dr. Cooper's group has been to improve the exchange of information between the compiler construction community and other groups-bringing outsiders into the community and applying code optimization methods in new domains. With Linda Torczon, he has written a textbook on introductory compiler construction; it will appear in print in March 2002 from Morgan Kaufmann. With Devika Subramanian and Linda Torczon, he has worked to apply techniques from artificial intelligence to scheduling and allocation. With John Bennett and Linda Torczon, he worked on ways to apply classic compiler optimizations to Vhdl-based circuits.

Dr. Cooper has been involved in outreach to K–12 teachers. He has lectured several times in Richard Tapia's GirlTech/Mcsa program—a summer institute that trains master teachers in technology while sensitizing them to the unique issues that arise in educating and retaining women and minorities. He has served as a *pro bono* consultant to several schools in the Houston area on classroom technology, on the role of technology in the curriculum, and on the low-level issues that arise in wiring schools.

Dr. Cooper has been active in his professional community. He served as Program Chair for the Acm Sigplan 98 Conference on Programming Language Design and Implementation and as Tutorials Chair for the same conference in 1993. He was a member of the Technical Steering Committee for the NSF-funded Center for Research on Parallel Computation from 1991 to 2000, and is currently a member of the Academic Planning and Coordination Committee for the Los Alamos Computer Science Institute. He served on the Policy Board of the Concurrent Supercomputing Consortium from 1991 to 1996

Dr. Cooper was on the design team for Duncan Hall–the new Computational Engineering Building at Rice University. It houses four departments, several research centers, and a research institute. The building was designed to encourage collaboration–both within disciplines and across disciplines. It has become the hub of activity in the School of Engineering and the space used for most of the University's programs that reach out to minorities, to women, and to educators from elementary and secondary schools.

Related Publications: (*see* `http://www.cs.rice.edu/~keith/Promotion`)

1. "Operator Strength Reduction," (with L.T. Simpson and C. Vick), ACM Transactions on Programming Languages and Systems (TOPLAS), to appear, 2001.

2. "Optimizing for Reduced Code Space Using Genetic Algorithms," (with P. Schielke and D. Subramanian), Proceedings of the 1999 ACM SIGPLAN Workshop on Languages, Compilers, and Tools for Embedded Systems (LCTES), May 1999, Atlanta, GA, USA,.

3. "Enhanced Code Compression for Embedded RISC Processors," (with N. McIntosh), Proceedings of the ACM SIGPLAN 99 Conference on Programming Language Design and Implementation, Atlanta, GA, USA, May 1999, pages 139-149.

4. "Compiler-controlled Memory," (with T.J. Harvey), Proceedings of the Eighth International Conference on Architectural Support for Programming Languages and Systems (ASPLOS), San Jose, CA, USA, October 1998, pages 2-11.

5. "Practical Improvements to the Construction and Destruction of Static Single Assignment Form," (with P. Briggs, T.J. Harvey, and L.T. Simpson), SoftwarePractice and Experience, 28(8), July 1998, pages 859-881.

Other Significant Publications and Patents:

1. "Non-local Instruction Scheduling with Limited Code Growth," (with P.J. Schielke), Proceedings of the 1998 ACM SIGPLAN Workshop on Languages, Compilers, and Tools for Embedded Systems (LCTES), June 1998, Montreal, CA, pages 193-207.

2. "Live Range Splitting in a Graph Coloring Register Allocator," (with L.T. Simpson), Proceedings of the 1998 International Conference on Compiler Construction (CC 99), March/April 1998, Lisbon, PT., pages 174-187.

3. "Combining Analyses, Combining Optimizations," (with C. Click), ACM Transactions on Programming Languages and Systems (TOPLAS), 17(2), March 1995, pages 181-196.

4. "Improvements to Graph Coloring Register Allocation," (with P. Briggs and L. Torczon), ACM Transactions on Programming Languages and Systems (TOPLAS), 16 (1994), pp. 428-455.

5. "Digital Computer Register Allocation and Code Spilling Using Interference Graph Coloring," (with P. Briggs, K. Kennedy, and L. Torczon), Patent Number: 5.249.295.

Collaborators:

Current Collaborators: John Bennett, Bradley Broom, Rob Fowler, Tim Harvey, John Mellor-Crummey, Ken Kennedy, Jan Sjodin, Devika Subramanian, and Linda Torczon (all at Rice), Dennis Gannon (Indiana U.), Fran Berman and Andrew Chien (UCSD), Carl Kesselman, (ISI,USC), Lennart Johnsson (U. of Houston), Dan Reed and Ruth Aydt (U. of Illinois, Urbana/Champagne), Jack Dongarra and Rich Wolski (U. of Tennessee), John Reynders (SUN), Rod Oldehoeft (and others from the Advanced Computing Lab at the Los Alamos National Laboratory), and all of the PIs on this proposal.

Recent Co-authors: (excluding those listed for other reasons): Kathryn McKinley (U. Mass.), Mary Hall (USC ISI), John Mellor-Crummey (Rice),

Thesis Advisees: Preston Briggs (Tera), Cliff Click (SUN Microsystems), L. Taylor Simpson (Intel), Nathaniel McIntosh (joint with K. Kennedy, Hewlett-Packard), John Lu (LSI Logic), Edmar Wienskoski (Motorola), Philip Schielke (Texas Instruments), Karim Esseghir (MS 93), Chris Vick (MS 94, SUN Microsystems), Jingsong He (MS 00, Dell), Li Xu, Tim Harvey, Todd Waterman,

Thesis Advisor: Ken Kennedy (Rice University)

Richard Hanson

Professional Preparation:

B.S., Oregon State University, 1960, (Mathematics)
M.S., Oregon State University, 1962 (Mathematics)
Ph.D., University of Wisconsin, 1965, (Mathematics)

Appointments:

1999- Research Scientist, Rice University Center for High Performance Software Research

1989-1999 Senior Principal Scientist, IMSL/Visual Numerics, Inc., Houston, TX

1987-1989 Principal Scientist, Applied Dynamics International, Ann Arbor, MI. Also Adjunct Professor of Mathematics, University of Michigan, 1988-1989.

1976-1987 Member of the Technical Staff, Sandia National Laboratory, Albuquerque, NM

1972-1976 Associate Professor and Professor, Depts of Pure and Applied Mathematics, Computer Science, Washington State University

1966-1972 Staff Scientist, Jet Propulsion Lab., California Institute of Technology

1964-1966 Assistant Professor, Mathematics Dept, University of Southern California

Synergistic Activities: Throughout his career in academia, government laboratories, and industry Dr. Hanson has pursued the common goal of advancing the state-of-the-art in numerical analysis and mathematical software to solve significant problems in applied mathematics, utilizing the full potential of available computing hardware.

At Rice University, Center for High Performance Software Research, he was the University Lead for the DoD/PET Modernization program. A significant accomplishment coming from this work was the writing of an interface to POSIX threads, using standard Fortran 95. This work makes it possible for Fortran programmers to have easy access to thread technology without writing codes in two languages. This capability will often result in major performance improvements and can make possible new approaches to problem solving.

At IMSL (now VNI) he assumed major project responsibility, including technical supervision, project management, and involvement in the full cycle of software development. Early IMSL projects included notable improvements in the IMSL real, symmetric, dense eigenvalue-eigenvector code, from which (given the right environment) an order of magnitude speed-up is obtained compared with LAPACK.

Some of the most professionally challenging work of his career began in 1992, with design and implementation of a new object-oriented interface for mathematical software libraries made possible by ISO Fortran 90. This consists of defined operators and generic functions in standard Fortran 90, and also interfaces with the IMSL Fortran 77 libraries. The use of innovative Fortran 90 features in these codes led to their being used extensively by hardware vendors as rigorous tests of new Fortran 90 compilers under development. This IMSL MP Library is widely available with DEC Fortran (Professional Version), as well as on many other computers.

His most significant work at VNI was to design and implement MPI-enhanced subroutines and functions for scientific applications on distributed network computers. The IMSL Distributed Network Fortran Library (DNFL) was developed on arrays of workstations, and is so far available for the IBM SP-2, Hitachi SR2201, and Cray T3E. His public service includes serving as editor for Collected Algorithms from ACM and Associate Editor for ACM Transactions on Mathematical Software.

Related Publications: (*See also* `http://home.att.net/~hanson.kool/pubs.html#paperlist`)

1. "A Fortran Interface to POSIX Threads," (with Clay Breshears and Henry Gabb), submitted to ACM-Trans. Math. Software, (July, 2000).

2. "Evaluating American-Style Option Prices Based on Constrained Least Squares," from Iterative Methods in Scientific Computation (IV), edited by David Kincaid and Anne Elster, IMACS, Dept. of Computer Science, Rutgers Univ., New Brunswick, NJ 08903, (1999).

3. Solving Least Squares Problems, with Charles L. Lawson, SIAM Publication, Classics in Applied Mathematics, No. 15, (1995).

Collaborators:

Current Collaborators: Clay Breshears and Henry Gabb, Kuck and Associates, Inc, Champaign-Urbana, IL; Jack Dongarra, University of Tennessee. All are co-authors of papers.

Thesis Advisor: Wolfgang R. Wasow, University of Wisconsin.

S. Lennart Johnsson

Professional Preparation:

Tekniska Lroverket Vsteras, Sweden, Electrical Engineering, Ingenjr (B.S.), 1963
Chalmers Institute of Technology, Gteborg, Sweden, Engineering Physics, Civilingenjr (M.S.), 1969
Chalmers Institute of Technology, Gteborg, Sweden, Control Engineering, Tekn. Lic (Ph.D.), 1971

Appointments:

Director, Texas Learning and Computation Center, University of Houston, 1999 -
Member, Executive Committee, W.M. Keck Center for Computational Biology, 1999 -
Member, Executive Committee, the Los Alamos Computer Science Institute, 1999 -
Chair, The Swedish National Allocations Committee for High Performance Computation, The Swedish Council for Planning and Coordination of Research, Stockholm, Sweden, 1999 -
Chair, The External Advisory Board, the National High Performance Computation and Visualization Center, PDC, Stockholm, Sweden, 1999 -
University of Houston Internet2 representative, 1997 -
Chair, Department of Computer Science, University of Houston, 1996 - 1999
Chair, Scientific Board, the National High Performance Computation and Visualization Center, PDC, Stockholm, Sweden, 1996 - 1999
Hugh Roy and Lillie Cranz Cullen Distinguished Professor of Computer Science, Mathematics and Electrical and Computer Engineering, University of Houston, 1995 -
Adjunct Professor, Department of Computer Science, Rice University, 1995 -
Visiting Professor, Department of Numerical Analysis and Computing Sciences, the Royal Institute of Technology, Stockholm, Sweden, 1995 -
Gordon McKay Professor of the Practice of Computer Science, Harvard University, 1990 - 1996
Director, Computational Sciences, Thinking Machines Corp., Cambridge, MA, 1987 - 1995
Associate Professor of Computer Science, and of Electrical Engineering, Yale University, 1983 - 1990
Senior Research Associate, California Institute of Technology, Pasadena, CA, 1979 - 1983
Manager, Systems Engineering, Central Research-&-Development, ASEA AB (now ABB), 1974 - 1980
Systems Engineer, Central Research-&-Development, ASEA AB (now ABB), 1970 - 1974
Postdoctoral Fellow, Systems Science Department, UCLA, Los Angeles, CA, 1970 - 1971

Synergistic Activities: At ASEA AB (now ABB), Dr. Johnsson implemented one of the first commercial-strength sparse-matrix packages, and led the development of systems for real-time supervision, control, and optimization of electric utility network operations, and for industrial process control. Within five years after entering the market for control centers for electric utilities, ASEA AB was a world leader with a revenue of about $50M (1987 dollars) for this product area and over 200 people engaged.

In 1982 at Caltech Dr Johnsson in collaboration with Dr Fornberg of the Applied Mathematics Dept. introduced one of the first courses in the country on large-scale scientific and engineering computation on scalable parallel architectures. Revisions of this course was later introduced by Dr Johnsson at Yale University (1983) and Harvard University (1990). At Yale University, both faculty and graduate students attended the course the first time it was taught. It was one of very few courses related to parallel architectures, their programming and use. At the University of Houston Dr Johnsson has also introduced a course on Advanced Networking addressing issues n the design and use of high-performance networks.

Some of the results of Dr Johnsson's research on network routing influenced the definition of the primitives in the MPI standard, and were adopted by vendors such as Intel and IBM in implementing the standard, and heavily influenced the Connection Machine Run-Time System.

At Thinking Machines Corp., Dr. Johnsson lead the design, development, and maintenance of the Connection Machine Scientific Software Library (CMSSL) and part of the Connection Machine Run-Time System (CMRTS). The CMSSL included several novel features, such as algorithm selection at run-time, and multiple-instance functionality for consistency with languages with array syntax.

Jointly with Rice University and Baylor College of Medicine Dr Johnsson established the Texas GigaPoP and was responsible for the first MPI applications for Globus demonstrated at SC97. He also lead the effort at two of five institutions performing an interactive, collaborative VR demonstration at Alliance '98 resulting in permanent Nordunet connectivity to the Abilene and vBNS networks and significantly increased Nordunet transatlantic capacity (from 34 Mbps to three OC-3 connections and doubled to six OC-3 connections this year). Johnsson has also actively contributed to establishing the European Grid Forum.

Related Publications:

1. Ken Kennedy, Bradley Broom, Keith Cooper, Jack Dongarra, Rob Fowler, Dennis Gannon, Lennart Johnsson, John Mellor-Crummy and Linda Torczon, "Telescoping Languages," submitted to the Journal of Parallel and Distributed Computing.
2. Y. Charlie Hu, G. Jin, L. Johnsson, D. Kehagias and N. Shalaby, "HPFBench: A High Performance Fortran Benchmark Suite," ACM Transaction on Mathematical Software, 26(1), pp. , 2000.
3. Michael Feig, Matin Abdullah, Lennart Johnsson and Montgomery Pettitt, "Large Scale Data Repository: Design of a Molecular Dynamics Trajectory Database," Future Generation Computer Systems, Elsevier, North- Holland, 16(1), pp. 101 - 110, 1999.
4. Kapil K. Mathur and Lennart Johnsson, "High Performance, Scalable Scientific Software Libraries," in Portability and Performance in Parallel Processing, John Wiley & Sons, pp. 159-208, 1994.
5. CMSSL, The Connection Machine Scientific Software Library, Thinking Machines Corp., 1986 - 1995.

Other Significant Publications:

1. Yu Charlie Hu, Shang-Hua Teng and Lennart Johnsson, "High Performance Fortran for Highly Irregular Problems," in Proceedings of the 6th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, ACM Press, pp 13-24, 1997.
2. Yu Hu and Lennart Johnsson, "Implementing N-body algorithms efficiently in data parallel languages," Journal of Scientific Programming, 5(4), pp. 337-364, 1996.
3. Zdenek Johan, Kapil K. Mathur, Lennart Johnsson and Thomas J.R. Hughes, "Scalability of Finite Element Applications on Distributed-Memory Parallel Computers," Computer Methods in Applied Mechanics and Engineering, 119(1-2), pp. 61-72, November 1994.
4. Lennart Johnsson, "Data Partitioning for Load-Balance and Communication Bandwidth Preservation," Proceedings of The Second International Conference on Massively Parallel Processing using Optical Interconnections, IEEE Computer Soc. Press, pp. 214-219, 1995.
5. Ted Nesson and S. Lennart Johnsson, "ROMM Routing on Mesh and Torus Networks," Proc. of the 7th Annual ACM Symp. on Parallel Algorithms and Architectures, ACM Press, 1995, pp. 275-287.

Collaborators:

Current Collaborators: Ralph Brickner (LANL/Darwin), Jean-Philippe Brunet (SUN), Barbara Chapman (UH), Wah Chiu (Baylor), Bjrn Engquist (UCLA), Michael Feig (Scripps), Ian Foster (ANL), Roland Glowinski (UH), Thomas Hughes (Stanford), Zdenek Johan (ONERA), Ken Kennedy (Rice), Carl Kesselman (USC/ISI), David Kramer (Oracle), Yuri Kuznetsov (UH), Bowen Loftin (UH), Kapil Mathur (D.E. Shaw), Montgomery Pettitt (UH), Ridgway Scott (Univ of Chicago), Jaspal Subhlok (UH), Paul Swarztrauber (NCAR), and Shang-Hua Teng (UIUC).

Recent Advisees and Postdoctoral Scholars: Ching-Tien Ho (IBM Almaden), Yu Charlie Hu (Rice), Olle Larsson (ANL), Ted Nesson (VideoGuide), Abhiram Ranade (formerly UC Berkeley), Nadia Shalaby, and Manish K Singh (Lucent). Dr Johnsson is currently supervising 10 graduate students and two Research Associates.

Thesis Advisor: Birger Qvarnstrm

Carl Kesselman

Information Sciences Institute	+1 310 488-9338
University of Southern California	+1 310 823-6714
4676 Admiralty Way, Suite 1001	carl@isi.edu
Marina del Rey, CA 90292-6695	http://www.isi.edu/~carl

Professional Preparation

State University of New York at Buffalo	Electrical Engineering	B.S. 1978-1982
University of Southern California	Electrical Engineering	M.S. 1982-1985
University of California at Los Angeles	Computer Science	Ph.D. 1985-1991

Appointments

2000-present	Senior Project Leader, Information Sciences Institute, University of Southern California
1996-2000	Project Leader, Information Sciences Institute, University of Southern California
1997-present	Research Associate Professor, Computer Science, University of Southern California
1996-present	Visiting Associate, Computer Science, California Institute of Technology
1993-1996	Member of the Beckman Institute, California Institute of Technology
1991-1993	Senior Research Fellow, Computer Science, California Institute of Technology
1990	Assistant Scientist, Mathematics and Computer Science Division, Argonne National Laboratory
1989	Visiting Scientist, Swedish Institute of Computer Science
1982-1991	Member of the Technical Staff, The Aerospace Corporation

Related Publications

[1] The Anatomy of the Grid: Enabling Scalable Virtual Organizations," Ian Foster, Carl Kesselman, Steve Tuecke, **International Journal of Supercomputer Applications**, to appear.

[2] The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets**.** A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, S. Tuecke, *proc. NetStore '99,* http://www.globus.org/documentation/incoming/datagrid.pdf.

[3] A Distributed Resource Management Architecture that Supports Advance Reservations and Co-Allocation. I. Foster, C. Kesselman, C. Lee, R. Lindell, K. Nahrstedt, A. Roy. *(Intl Workshop on Quality of Service, 1999).* ftp://ftp.globus.org/pub/globus/papers/hbm.pdfb2

[4] A Resource Management Architecture for Metacomputing Systems. K. Czajkowski, I. Foster, N. Karonis, C. Kesselman, S. Martin, W. Smith, S. Tuecke. *Proc. IPPS/SPDP '98 Workshop on Job Scheduling Strategies for Parallel Processing,*1998

[5] A Directory Service for Configuring High-Performance Distributed Computations. S. Fitzgerald, I. Foster, C. Kesselman, G. von Laszewski, W. Smith, S. Tuecke. *Proc. 6th IEEE Symp. on High-Performance Distributed Computing,* pg. 365-375, 1997, ftp://ftp.globus.org/pub/globus/papers/hpdc97-mds.pdf

Other Significant Publications

[1] I. Foster, J. Insley, C. Kesselman, G. von Laszewski, M. Thiebaux, "Distance Visualization: Data Exploration on the Grid," IEEE Computer, December 1999.

[2] The Globus Project: A Status Report. I. Foster, C. Kesselman, *Proc. IPPS/SPDP '98 Heterogeneous Computing Workshop*, pg. 4-18, 1998.

[3] A Security Architecture for Computational Grids. I. Foster, C. Kesselman, G. Tsudik, S. Tuecke, *Proc. 5th ACM Conference on Computer and Communications Security Conference*, pg. 83-92, 1998. ftp://ftp.globus.org/pub/globus/papers/security.pdf.

[4] Real-Time Analysis, Visualization, and Steering of Microtomography Experiments at Photon Sources, Gregor von Laszewski, Mei-Hui Su, Joseph A. Insley, Ian Foster, John Bresnahan, Carl Kesselman, Marcus Thiebaux, Mark L. Rivers, Steve Wang, Brian Tieman, Ian McNulty, Ninth SIAM Conference on Parallel Processing for Scientific Computing, Apr. 1999, http://wwwfp.globus.org/documentation/incoming/siamCmt99.pdf.

[5] Implementing Distributed Synthetic Forces Simulations in Metacomputing Environments. S. Brunett, D. Davis, T. Gottshalk, P. Messina, C. Kesselman, *Proceedings of the Heterogeneous Computing Workshop*. Mar. 1998, ftp://ftp.globus.org/pub/globus/papers/sf-express.pdf

Synergistic Activities

- Pedagogical: Co-editor of widely used text: *The Grid: Blueprint for a Future Computing Infrastructure* (Morgan Kaufmann, 1999: www.mkp.com/grid); also, teaching of numerous related tutorials.
- Research tools: Developed numerous software systems that have seen extensive use in the research community including the Compositional C++ programming environment, the Nexus communication library and the Globus Grid toolkit.item1;item2;item3.
- Service: Program committee: SC Program Committee (1994,1997,2000), the International Symposium on Parallel Processing (1996), the International Symposium on Concurrent Object Oriented Programming (1997,1999), High-Performance Distributed Computing Conference (1992,1993,1998,1999,2000). Program Co-chair High Performance Distributed Computing Conference in 1999..
- Awards and Honors: 1998 Global Information Infrastructure Award.

Collaborators

Bruce Allen (U. Wisc.), Joseph Bannister (USC), Chaitanya Baru (UCSD), Francine Berman (UCSD), Joseph Bester (ANL.), John Bresnahan (ANL), Sharon Brunett (CIT), Randy Butler (NCSA), Julian Bunn (CIT) Henri Casanova (UCSD), Charles Catlett, (Argonne Natl. Lab.), Ann Chervenek (USC), Andrew Chien (UCSD), K. Mani Chandy (CIT.), Keith Cooper (Rice U.), Karl Czajkowski (USC), Dan Davis (USC), Tom DeFanti (U. of Ill.),Jack Dongarra (U. of Tenn.), Mark Ellisman (UCSD), Steven Fitzgerald (USC), Ian Foster (Argonne Natl. Lab.), Geoffrey Fox (Syracuse U.), Michael Franklin (Berkeley), Dennis Gannon (Indiana U.), Robert Gardner (Indiana), Jonathan Geisler (Northwestern U.), Thomas Gottschalk (CIT),Andrew Grimshaw (UVA), Bill Gropp (ANL), Roch Guerin (IBM), Robert Hollebeek (Penn), Jeffrey Hollingsworth (U. of Maryland), John Huth (Harvard), Joseph Insley (ANL), Chris Johnson (U. of Utah), Lennart Johnson (U. of Houston), William Johnston (NASA), Tom Jordan (USC), Nickolas Karonis (Northern Ill. U.), Ken Kennedy (Rice U.), Stephen Kent (Chicago), Albert Lazzarini (CIT), Craig Lee (The Aerospace Corp.), Jason Leigh (U. of Ill.), Robert Lindell (USC), Miron Livny (U. of Wisc.), Meloney Loots (NCSA), Richard Marciano (UCSD), Stuart Martin (Argonne Natl. Lab.), Keith Marzullo, (UCSD), Ian McNulty (ANL), Paul Messina (CIT.), Bart Miller (U. of Wisc.), Reagan Moore (UCSD), Richard Mount (SLAC), Klara Nahrstedt (U. of Ill.), Harvey Newman (Caltech), Larry Peterson (Princeton U.), Lawrence Price (ANL), Thomas Prince (Caltech), Arcot Rajasekar (UCSD), Daniel Reed (U. of Ill.), Mark L. Rivers (U. of Chicago), Alain Roy (U. of Chicago), Joel Saltz (U. of Maryland), Henning Schulzrinne (Columbia U.), Steven Schwab (Network Associates), Arie Shoshani (LBNL), Warren Smith (NASA), Marc Snir (IBM), James Stepanek (The Aerospace Corp.), Rick Stevens (ANL), Mei Su (USC), Alexander Szalay (Johns Hopkins), Marcus Thiebaux (USC), Brian Tieman (ANL), Gene Tsudik (UCI), Steven Tuecke (Argonne Natl. Lab.), Joseph Romano (Texas Brownsville),Michael Wan (UCSD), Steve Wang (Argonne Natl. Lab.), Mary Wheeler (U. of Texas), Roy Williams (Caltech), Richard Wolski (U. of Tenn.), Gregor von Laszewski (Argonne Natl. Lab.)

Graduate Advisor: Milos Ercegovac, University of California at Los Angeles

Thesis Advisor & Postgraduate-Scholar Sponsor:

Soonwook Hwang, University of Southern California

John Mellor-Crummey

Professional Preparation:

B.S.E., Princeton University, 1984 (Electrical Engineering and Computer Science, Magna Cum Laude)
M.S., University of Rochester, 1986 (Computer Science)
Ph.D., University of Rochester, 1989 (Computer Science)

Appointments:

1998 - Senior Faculty Fellow, Department of Computer Science, Rice University
1992 - 1998 Faculty Fellow, Department of Computer Science, Rice University
1989 - 1992 Research Associate, Department of Computer Science, Rice University

Synergistic Activities: John Mellor-Crummey is a Senior Faculty Fellow in the Department of Computer Science, a research faculty appointment commensurate with the rank of Associate Professor. In 1996, he was named as a member of the technical steering committee of the CRPC. In 1999, Rice University and Los Alamos National Laboratory jointly established the Los Alamos Computer Science Institute (LACSI). Since the founding of LACSI, Mellor-Crummey has served as a member of its Executive Committee and has led the Institute's research effort in compilation, systems and performance evaluation.

Mellor-Crummey's research has focused on topics in parallel computing. His early research focused on debugging and performance analysis of parallel programs. Techniques he developed for replaying executions of parallel programs for the purpose of debugging are today used in environments such as IBM s DejaVu debugger for Java. He is most widely known for multi-processor synchronization algorithms he developed with Michael Scott (University of Rochester), which are widely used in practice. A significant focus of his research since 1993 has been the development of data-parallel compiler technology to support parallel scientific computing.

Mellor-Crummey is currently an investigator in an ASCI Level-2 academic partnership in which he is leading a research effort to develop techniques to improve scalar performance of ASCI scientific codes. As part of that effort, he is leading development of HPCView, a performance analysis tool that is currently being used by application scientists at LANL and Sandia National Laboratories.

In the early 1990s, Dr. Mellor-Crummey worked with the Center for Analysis and Prediction of Storms on automatic parallelization of a research code for mesoscale modeling of severe storms and with a team from Rice and the University of Houston on the parallelization of CHARMM, a widely used computational chemistry program developed at Harvard.

He is currently a member of program committees for the upcoming Los Alamos Institute Computer Science Symposium (Santa Fe, NM), International Symposium on Distributed Computing 2001 (Lisbon, Portugal), and Parallel Computing 2001 (Naples, Italy). He was previously a member of the program committees for Supercomputing (2000), International Conference on Parallel Processing (1999), International Parallel Computing Conference (1999,1997), International Conference on Supercomputing (1997, 1996), and the Symposium on Parallel and Distributed Tools (1996). He was a member of the technical infrastructure working group at the Workshop on Software Tools for High Performance Computing Systems (1996), the software tools working group at the Pasadena Workshop on System Software and Tools for High Performance Computing Environments (1992), and a NSF Review Panel for Experimental Software Systems (1997).

Related Publications:

1. Vikram S. Adve and John M. Mellor-Crummey. "Using Integer Sets for Data-Parallel Program Analysis and Optimization," Proceedings of the ACM SIGPLAN '98 Conference on Programming Language Design and Implementation, Montreal, CA, (June 1998).

2. Vikram S. Adve, Guohua Jin, John M. Mellor-Crummey and Qing Yi. "High Performance Fortran Compilation Techniques for Parallelizing Scientific Codes," Proceedings of Supercomputing 98: High Performance Networking and Computing, Orlando, FL, (November 1998).

3. Bo Lu and John M. Mellor-Crummey. “Compiler Optimization of Implicit Reductions for Distributed-Memory Multiprocessors,” Proceedings of the 12th International Parallel Processing Symposium, Orlando, FL, (May 1998).

4. John M. Mellor-Crummey and Vikram S. Adve. “Simplifying Control Flow in Compiler-Generated Parallel Code,” International Journal of Parallel Programming, 26(5), 1998. Special Issue with selected papers from the 10th International Workshop on Languages and Compilers for Parallel Computing.

5. K.D. Cooper, M.W. Hall, R.T. Hood, K. Kennedy, K.S. McKinley, J.M. Mellor-Crummey, L. Torczon, and S.K. Warren. “The ParaScope Parallel Programming Environment,” Proceedings of the IEEE, 81(2):244–263, February 1993.

Other Significant Publications:

1. John M. Mellor-Crummey and Michael Scott. “Algorithms for Scalable Synchronization on Shared-Memory Multiprocessors,” ACM Transactions on Computer Systems, 9(1):21–65, February 1991.

2. Thomas J. LeBlanc and John M. Mellor-Crummey, “Debugging Parallel Programs with Instant Replay,” IEEE Transactions on Computers, C-36(4):471–482, April 1987.

3. John M. Mellor-Crummey and Thomas J. LeBlanc. “A Software Instruction Counter,” Proceedings of the Third International Conference on Architectural Support for Programming Languages and Operating Systems, (April 1989), 78–86.

4. John M. Mellor-Crummey, David Whalley and Ken Kennedy. “Improving Memory Hierarchy Performance for Irregular Applications,” Proceedings of the International Conference on Supercomputing, Rhodes, Greece ,(June 1999), 425–433. (Selected to appear in a future special issue of the International Journal of Parallel Programming.)

5. John M. Mellor-Crummey. “On-the-fly Detection of Data Races for Programs with Nested Fork-Join Parallelism,” Proceedings of Supercomputing ’91, Albuquerque, NM, (November 1991), 24–33.

Collaborators:

Current Collaborators: Vikram Adve (UIUC), Ralph Brickner, Alan Cox (Rice), Robert Fowler (Rice), Guohua Jin (Rice), Ken Kennedy (Rice), Gerald Roth (Gonzaga U.), David Whalley (Florida State), Qing Yi (Rice), Ewing Lusk (ANL), Barbara Chapman (U. of Houston), Jarek Nieplocha (PNNL), Katherine Yelick (UC Berkeley), Tom Sterling (Caltech/JPL), Bill Gropp (ANL), Ricky Kendall (Ames Laboratory), Robert Lucas (LBNL), Guang Gao (U. of Delaware), D. Panda (OSU), and Marianne Winslett (UIUC). As part of a multi-institutional effort, I interact with Dennis Gannon (Indiana U.), Fran Berman and Andrew Chien (UCSD), Carl Kesselman, (USC ISI), Lennart Johnsson (U. of Houston), Dan Reed and Ruth Aydt (UIUC), Jack Dongarra and Rich Wolski (U. of Tennessee).

Thesis Advisees: Bo Lu (National Semiconductor), Collin McCurdy (U. of Wisconsin), Kai Zhang (Compaq), Daniel Chavarria-Miranda

Thesis Advisor: Thomas LeBlanc (University of Rochester)

Richard Tapia

Professional Preparation:

University of California, Los Angeles, California, B.A. (Mathematics), 1961.
University of California, Los Angeles, California, M.A. (Mathematics), 1966.
University of California, Los Angeles, California, Ph.D. (Mathematics), 1967.

Appointments:

Adjunct Professor, University of Houston, 2000-present

Director, Center for Excellence and Equity in Education, Rice University 1999-present

Cluster Leader, Alliances for Graduate Education in the Professoriate (AGEP), Rice University, 1998-present

Noah Harding Professor of Computational and Applied Mathematics, Rice University, 1991-present

Associate Director of Graduate Studies, Office of Graduate Studies, Rice University, 1989-present

Director of Education and Outreach Programs, Center for Research on Parallel Computation Rice University, 1989-2000

Professor of Mathematical Sciences, Rice University, 1976-present

Lecturer, Department of Community Medicine, Baylor College of Medicine, 1986-1988

Adjunct Professor, Texas Institute of Rehabilitation and Research, Baylor College of Medicine, 1978-1983

Chair, Department of Mathematical Sciences, Rice University, 1978-1983

Visiting Associate Professor of Operations Research, Stanford University, 1976-1977

Associate Professor of Mathematical Sciences, Rice University, 1972-1976

Assistant Professor of Mathematical Sciences, Rice University, 1970-1972

Assistant Professor, Mathematics Research Center, University of Wisconsin–Madison, 1968-1970

Instructor in Mathematics, University of California-Los Angeles, 1967-1968

Synergistic Activities:

- Developed and directed Spend a Summer with a Scientist, a SMET graduate studies recruitment and retention program for underrepresented minorities and women, 1989. Institutionalized at Rice in 1998 and exported to the University of Wisconsin–Madison.
- Developed and directed K–12 Teacher Professional Development Programs: The Mathematical and Computational Sciences Awareness Workshop (1989) and GirlTECH (1995) that address the underrepresentation of women and underrepresented minorities in computational sciences.
- Chair Advisory Committee for Houston Independent School District's NSF Urban Systemic Initiative. (1998-present)
- Co-developing high school curriculum materials: Introduction to Linear Algebra, (with Tamara Carter and Ann Papakonstantinou), (2000), `http://ceee.rice.edu/Books/LA/`.
- Computational Science: Tools for a Changing World, (with Cassandra McZeal and Cynthia Lanius), `http://ceee.rice.edu/Books/CS/`.

Related Publications:

1. Underrepresented Minority Achievement and Course Taking—The Kindergarten-Graduate Continuum, 2000 NISE Forum: Diversity and Equity Issues in Mathematics and Science Education, `http://ceee.rice.edu/Books/DV/continuum/`.

2. Assessing and Evaluating the Evaluation Tool—The Standardized Test, 1998 NISE Forum: Assessment and the Promotion of Change, `http://www.caam.rice.edu/~rat/nise.html`.

3. Computing an Exact Solution in Interior-Point Methods for Linear Programming, Contemporary Mathematics, 252 (1999), 9-29 (with P. Williams and A. El-Bakry).

4. The Solution of the Metric STRESS and SSTRESS Problems in Multidimensional Scaling Using Newton's Method, Computational Statistics, 13 (3) (1998), 369-396 (with A. Kearsley and M. Trosset).

5. Perturbation Lemma for Newton's Method with Application to the SQP Newton Method, Journal of Optimization Theory and Applications, 97 (1) (1998), (with D. Cores).

Other Significant Publications:

1. On Effectively Computing the Analytic Center of the Solution Set by Primal-Dual Interior-Point Methods, SIAM Journal on Optimization, 8 (1) (1998), (with M. Gonzalez-Lima and F. Potra).

2. A Robust Choice of the Lagrange Multiplier in the SQP Newton Method, Investigacin Operativa, 7 (1,2) (1997), (with D. Cores).

3. On the Convergence of the Mizuno-Todd-Ye Algorithm to the Analytic Center of the Solution Set, SIAM Journal on Optimization, 7 (1) (1997), 47-65 (with C. Gonzaga).

4. On the Quadratic Convergence of the Simplified Mizuno-Todd-Ye Algorithm for Linear Programming, SIAM Journal on Optimization, 7 (1) (1997), 66-85 (with C. Gonzaga).

Collaborators:

Current Collaborators: Baine Alexander, University of Wisconsin–Madison, Miguel Argaez, Debora Cores, Amr El-Bakry, Maria Gonzalez-Lima, Terry Millar, University of Wisconsin–Madison, Zeferino Parada, George Phillips, Michael Trosset, , and Yin Zhang, Rice University

Thesis Advisees: Miguel Argaez, Richard H. Byrd, Maria Rosa Celis, Jershan Chiang, Debora Cores, Edward Dean, Mahmoud El-Alem, Amr El-Bakry, Mohammedi El-Hallabi, Rodrigo Fontecilla, Naresh Garg, Mark Gockenbach, Maria Gonzalez-Lima, Victor M. Guerra, Kathie L. Hiebert, Anthony Kearsley, Shou-Bai Li, Hector J. Martinez, Mary Anne McCarthy, Gilbert Mayor de Montricher, Jorge Nocedal, Zeferino Parada, Teresa Parks, Marcos Raydan, Catherine Samuelsen, David W. Scott, Leticia Velazquez, M. Cristina Villalobos, Donald Williams, Pamela Williams, Cassandra McZeal, and Diane Jamrog.

Graduate and Postdoctoral Advisors: Magnus Hestenes, David A. Sanchez, and Charles B. Tompkins

Linda Torczon

Professional Preparation:

B.S., Rice University, 1980 (Chemical Engineering, Magna cum Laude)
M.S., Rice University, 1984 (Computer Science)
Ph.D., Rice University, 1985 (Computer Science)

Appointments:

Research Scientist, Department of Computer Science, Rice University, 1985 to present
Executive Director, CRPC, Rice University, 1990 to 2000.

Synergistic Activities:

Linda Torczon's research interests include code generation, interprocedural data-flow analysis and optimization, and programming environments. In the code generation realm, she published a set of improvements to graph coloring register allocation. She is also one of the key implementors of an optimizing compiler for Fortran. In the area of interprocedural analysis and optimization, she developed techniques for interprocedural constant propagation and recompilation analysis. She also completed a study on the effectiveness of several interprocedural constant propagation techniques and collaborated on a study of the effectiveness of inline substitution. In the programming environment arena, she was one of the driving forces behind the ParaScope programming environment project at Rice. She was a principal architect of the framework for whole program analysis in the ParaScope programming environment. Techniques that she developed are widely used in industrial and research compilers.

From 1990 to 2000, Dr. Torczon served as executive director of the Center for Research on Parallel Computation (CRPC), a National Science Foundation Science and Technology Center. In this capacity, she coordinated extensive research efforts, education and outreach programs, and technology transfer activities (http://www.crpc.rice.edu/CRPC). She is currently editing the *CRPC Handbook of Parallel Computation*, intended as a resource for computer science and application researchers, as well as for computational science and parallel computing education and training.

Dr. Torczon has been involved in activities intended to increase the number of women and underrepresented minorities entering mathematics and science related fields, particularly the field of computational science and engineering. With Ken Kennedy, she has initiated several CRPC outreach activities, including the CRPC GirlTECH program and the Girl Games effort. She has made presentations to students participating in Expanding Your Horizons, The Galveston Bay Project, and Girl Games programs that encourage middle-school girls to pursue technical careers. She has made presentations to K–12 teachers as part of GirlTECH and other Rice University programs aimed at improving mathematical and computational skills among K–12 teachers. She served on the Shared Decision Making Team of The Rice School/La Escuela Rice, a Houston Independent School District K–8 laboratory school. Finally, as tutorial chair for the ACM SIGPLAN Conference on Programming Language Design and Implementation, she directed an NSF-funded tutorial program that brought faculty members from undergraduate institutions, particularly women's colleges and institutions with large minority enrollments, to the conference and tutorials.

With Keith Cooper, Linda Torczon is co-authoring *Engineering a Compiler*, which is intended as a textbook for senior-level courses on compiler construction and as a resource for compiler implementors.

Dr. Torczon has served on the Program Committee for ACM SIGPLAN PLDI Conference (1994 & 2000), on the Tutorial Committee of Supercomputing 97, as Tutorial Chair of the ACM SIGPLAN PLDI Conference (1997), as Treasurer of the ACM SIGPLAN PLDI Conference (1996), on the NSF Postdoctoral Research Associates Program Panel for the ASC Division (January 1995), and on the NSF Postdoctoral Panel for the New Technologies Program of the ASC Division (February 1994).

Related Publications:

1. "How to Build an Interference Graph," (with K. D. Cooper and T. J. Harvey), Software-Practice and Experience, April 1998.
2. "Interprocedural Analysis and Optimization," (with K. D. Cooper, M. W. Hall, and K. Kennedy), Communications on Pure and Applied Mathematics, 48, 1996, pages 947–1003.
3. "Improvements to Graph Coloring Register Allocation," (with P. Briggs and K. D. Cooper), ACM Transactions on Programming Languages and Systems (TOPLAS), 16 (1994), pp. 428–455.
4. "Interprocedural Constant Propagation: A Study of Jump Function Implementations," (with D. Grove), Proceedings of the ACM SIGPLAN 93 Conference on Programming Language Design and Implementation (PLDI), SIGPLAN Notices, 28 (1993), pp. 90–99.
5. "Interprocedural Optimization: Eliminating Unnecessary Recompilation," (with M. Burke), ACM Transactions on Programming Languages and Systems (TOPLAS), 15 (1993), pp. 367–399.

Other Significant Publications and Patents:

1. "Compilers," (with K.D. Cooper and K. Kennedy), article in *Encyclopedia of Physical Science and Technology*, 3^{rd} Edition, Academic Press, *to appear, 2001.*
2. "An Efficient Representation for Sparse Sets," (with P. Briggs), ACM Letters on Programming Languages and Systems, 2 (1993), pp. 59–69.
3. "The ParaScope Parallel Programming Environment," (with K. D. Cooper, M. W. Hall, R. T. Hood, K. Kennedy, K. S. McKinley, J. Mellor-Crummey, and S. K. Warren), Proceedings of the IEEE, 81(2), February 1993, pp. 244–263.
4. "Rematerialization," (with P. Briggs and K. D. Cooper), Proceedings of the ACM SIGPLAN 92 Conference on Programming Language Design and Implementation, SIGPLAN Notices, 27(7), July 1992, pp. 311–321.
5. "Digital Computer Register Allocation and Code Spilling Using Interference Graph Coloring," (with P. Briggs, K. D. Cooper, and K. Kennedy), Patent Number: 5.249.295.

Collaborators:

Current Collaborators: As executive director of both the CRPC and the Los Alamos Computer Science Institute (LACSI), I have interacted with researchers at Argonne National Laboratory (ANL), Caltech, Los Alamos National Laboratory (LANL), Rice University, Syracuse University, University of Houston, University of Illinois at Urbana-Champaign, University of Tennessee, and the University of Texas. For the sake of brevity, I have not exhaustively listed members of the CRPC or LACSI efforts. Other collaborators include: John Bennett, Bradley Broom, Keith Cooper, Rob Fowler, Tim Harvey, Ken Kennedy, John Mellor-Crummey, and Devika Subramanian (Rice); Dennis Gannon (Indiana University); Fran Berman and Andrew Chien (UCSD); Carl Kesselman (ISI, USC); Jack Dongarra and Rich Wolski (University of Tennessee); John Reynders (SUN), Kathryn McKinley (University of Massachusetts), and all of the authors and editors of the *CRPC Handbook of Parallel Computation*.

Thesis Advisee: Daniel Grove

Thesis Advisor: Ken Kennedy (Rice University)

Richard Wolski

Professional Preparation:

California Polytechnic University, San Luis Obispo, Computer Science, B.S., 1986
University of California at Davis/Livermore Campus, Computer Science, M.S., 1989
University of California at Davis/Livermore Campus, Computer Science, Ph.D., 1994
University of California at San Diego, Distributed Computing, 1994-1999

Appointments:

Assistant Professor, University of Tennessee, 1999-present

SDSC Junior Fellow, University of Tennessee, 1994-present

Research Professor, University of California, San Diego, 1995-1999

Postdoctoral Researcher, University of California, San Diego, 1994-1995

Relevant Publications:

1. Wolski, R., Spring, N. and Hayes, J., The Network Weather Service: a Distributed resource Performance Forecasting Service for Metacomputing, *Journal of Future Generation Computing Systems*, Volume 15, Numbers 5-6, pp. 757-768, October, 1999. (available from `http://www.cs.utk.edu/~rich/publications/nws-arch.ps.gz`)

2. Wolski, R., Brevik, J., Krintz, C., Obertelli, G., Spring, N., and Su, A., Running EveryWare on the Computational Grid, *Proceedings of SC99*, November, 1999. (available from `http://www.cs.utk.edu/~rich/publications/ev-sc99.ps`)

3. Wolski, R., Spring, N. and Hayes, J., Predicting the CPU Availability of Time-shared Unix Systems on the Computational Grid, *Proceedings of 8th High-performance Distributed Computing Systems Conference*, August, 1999. (available from `http://www.cs.utk.edu/~rich/publications/nws-cpu.ps.gz`)

4. Wolski, R., Spring, N., Peterson, C., Implementing a Performance Forecasting System for Metacomputing: The Network Weather Service, Proceedings of SC97, November, 1997. (available from `http://www.cs.utk.edu/~rich/publications/nws-impl.ps.gz`)

5. Spring, N. Wolski, R., Application Level Scheduling of Gene Sequence Comparison on Metacomputers, Proceedings of ACM 1988 International Conference on Supercomputing, July, 1998. (available from `http://www.cs.utk.edu/~rich/publications/apples-ics.ps.gz`)

Other Significant Publications:

1. Berman F. Wolski, R., Figueira, S., Schopf, J., and Shao, G., Application-Level Scheduling on Distributed Heterogeneous Networks, *Proceedings of Supercomputing 1996*, November, 1996. (available from `http://www.cs.utk.edu/~rich/publications/apples.ps.gz`)

2. Wolski, R., Static Scheduling of Hierarchical Program Graphs, *Journal of Parallel Processing Letters*, Volume 5, Number 4, pp. 611-622, December, 1995. (available from `http://www.cs.utk.edu/~rich/publications/sched.ps.gz`)

3. Wolski, R., Feo, J.T., Program Partitioning for NUMA Multiprocessor Computer Systems, *The Journal of Parallel and Distributed Computing*, Volume 19, No. 3, pp. 203-218, November, 1993.

4. Griswold, W., Wolski, R. Baden, S., Fink, S., Kohn, S., Programming Language Requirements for the Next Millennium, *ACM Computing Surveys*, Volume 28, No. 4, pp. 194-196, December, 1996.

5. Gorda, B. and Wolski, R., Time Sharing Massively Parallel Machines, *Proceedings of the International Conference on Parallel Processing*, August, 1995.

Synergistic Activities :

1. **Partner in the National Partnership for Advanced Computational Infrastructure (NPACI)**. As a member of this partnership, Dr. Wolski is working to make the research-grade software artifacts he produces available to the community at large. The technology transfer he leads benefits both the computer and computational science communities.

2. **Leader, Network Weather Service Project**. Dr. Wolski's research in dynamic network performance forecasting has lead to an on-line and supported permanent installation of the Network Weather Service – a facility for predicting nation-wide resource load in near real-time.

3. **Co-Leader, Application-level Scheduling Project**. Dr. Wolski (with his collaborator Professor Francine Berman) co-leads the **App**lication **Le**vel **S**cheduling (AppLeS) project which is building dynamic, adaptive scheduling tools for the Computational Grid.

Collaborators:

Micah Beck (Univ. Tennessee), Francine Berman (UC San Diego), Jack Dongarra (Univ. Tennessee, ORNL) John Feo (Tera Computer Corp.), Ian Foster (Argonne National Lab) Carl Kesselman (USC/ISI), Reagan Moore (SDSC), James Plank (Univ. Tennessee), Margaret Simmons (NPACI), Andrew Grimshaw (Univ. Virginia)

Thesis Advisor: John Feo (now at Tera Computer Corp.)

Postdoctoral Sponsor: Francine Berman (UC San Diego)

Recent Graduate Student and Post-Doctoral Advisees: James Calloway (UT), Walfredo Cirne (UCSD), Marcio Faerman (UCSD), Gary Shao (UCSD), Alan Su (UCSD), Martin Swany (UT)

total advised: 6

Contact Information:

Department of Computer Science
University of Tennessee
1122 Volunteer Blvd.
Knoxville, TN 37996-3450

(865) 974-4394
(865) 974-4404 *FAX*
`rich@cs.utk.edu`
`www.cs.utk.edu/~rich/`

FACILITIES, EQUIPMENT & OTHER RESOURCES

FACILITIES: Identify the facilities to be used at each performance site listed and, as appropriate, indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed and at sites for field studies. USE additional pages as necessary.

Laboratory: **Not applicable.**

Clinical: **Not applicable.**

Animal: **Not applicable.**

Computer: **Each of the departments of the project PIs maintain computing infrastructures of the highest caliber. Each location maintains and operates an extensive shared computing facility available to all members of that institution's department. This infrastructure includes very high bandwidth network connections, laboratories of**

Office: **Each institution maintains printers, copiers, fax machines, and other normal office equipment - all readily available to the project researchers.**

Other: **At Rice University, all project researchers are housed in Duncan Hall, a recently constructed building for the support of Computational and Computer Engineering. It specifically houses educational and research personnel in the fields of computational and applied mathematics, computer science, electrical and computer engineering, information technology, and statistics - fostering a higher level of interaction and**

MAJOR EQUIPMENT: List the most important items available for this project and, as appropriate identifying the location and pertinent capabilities of each.

The currently proposed, distributed Grid testbed, called the Macrogrid, is composed of:

From the University of Southern California's Information Services Institute (1) a 10 processor SGI Origin/2000/Onyx 2 system with 3 Infinite Reality 2 graphics pipelines, running the IRIX operating system, (2) an SGI Origin 2000 (64 nodes) and a tape robot operating under Unitree , and (3) a Pentium III Cluster (128 dual processor nodes)

OTHER RESOURCES: Provide any information describing the other resources available for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available for the project. Include an explanation of any consortium/contractual arrangements with other organizations.

Another important facility is the MicroGrid software which enables Grid researchers and application scientists to simulate the dynamics of both Grids and Grid applications. The MicroGrid can be easily be deployed on a single or cluster of PC's, and allows simulation and study of a variety of Grid processing, network, and storage resources. We will provide CGrADS application partners access to the MicroGrid tools in software release as well as a network accessible testbed. As the project progresses, the capabilities of the MicroGrid software will increase, and successive

Continuation Page:

COMPUTER FACILITIES (continued):

machines for use by students and visitors, departmentally shared compute servers and file servers, along with printers, PCs and/or Macs, and backup facilities.

Each of the project institutions has participated in the very-high-speed Backbone Network Service (vBNS), an experimental network launched in 1995 that connects national supercomputing centers and universities across the country to collaborate and share powerful computing and information resources. Most of the institutions are now also connected to other very high speed networks (e.g. Abilene, NTON, Cairn, etc).

For high-end visualization, most of the PIs have direct access to local ImmersaDesks, typically driven by SGI Onyx 2 graphics engines. In addition, we have a number of desk-top machines with high-end graphics cards for less demanding visualization tasks.

OTHER FACILITIES (continued):

collaboration between different computational disciplines. The building contains numerous instructional and computational facilities .

Likewise, centralized facilities at UIUC (through NCSA), UCSD (through SDSC/NPACI), USC (through ISI), UH (through TLCC), and Tennessee (through the new Innovation Computing Laboratory's Center for Information Technology Research) , provides not only an enormous set of computing and networking infrastructure and equipment resources, they also provide an environment, similar to Duncan Hall, where researchers with overlapping computational interests can more readily share experiences.

In addition, resource access will be provided by the National Partnership for Advanced Computational Infrastructure (NPACI) and the National Computational Science Alliance (NCSA) to additional resources, per the included letters of support.

MAJOR EQUIPMENT (continued):

running Linux 6.2 -- there are plans to at least double the size of this last cluster by the time the STC award is made.

From Rice University, an IBM SP2 (24 nodes) running AIX.

From the University of Tennessee at Knoxville, (1) a Pentium Cluster (8 nodes) running Red Hat Linux and (2) a Pentium Cluster (16 nodes) running Debian Linux.

From the University of California at San Diego, a Pentium Cluster (16 nodes) running Debian Linux.

Continuation Page:

MAJOR EQUIPMENT (continued):

From the University of Illinois at Urbana-Champaign, a - Pentium Cluster (24 nodes) running Red Hat Linux.

From the University of Chicago, a Pentium Cluster (16 nodes) running SUSE Linux.

This mix of high end computer facilities will certainly change over time, but we fully expect this change will represent growth, both in terms of computing capability and resource heterogeneity. Other major equipment available for the project, but not yet proposed as being part of the Macrogrid are a 2.5 Terabyte, 16 node storage cluster at USC/ISI, and (from the University of Houston) a 4-processor HP V-Class SMP, and an inhomogeneous cluster with eight dual Pentium III 500 MHz processors with 256 MB main memory, and three four-processor Xeon 550MHz systems with 512 MB memory.

OTHER RESOURCES (continued):

releases of the software will be made available.

RICE

JORDAN KONISKY
Vice Provost
for Research and Graduate Studies

April 2, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts:

Rice University is pleased to submit this research proposal to the National Science Foundation (NSF), entitled "Center for Grid Application Development Software (CGrADS)", under the direction of Professor Ken Kennedy at Rice. Partnering with Rice University are the University of Tennessee at Knoxville, the University of Illinois at Urbana-Champaign, the University of California at San Diego, the University of Southern California Information Sciences Institute, the University of Chicago, and the University of Houston.

We are fully supportive of the proposed center, and will provide 30% of the amount funded from NSF as cost share for the life of the project. This support will be in the form of graduate student stipends and tuitions, cash funding to support the educational, research and infrastructure goals of the Center, and academic year salary for faculty to participate in the project beyond what is supported by the NSF budget. The sum of $19,434,367, including facilities and administrative costs, is requested from the NSF for the period of August 2002 through August 2007.

We are confident the proposed CGrADS effort has all the components to be a successful Science and Technology Center. It pulls together a truly impressive list of collaborators, with proven leadership skills, to work on a problem that represents the next large technological improvement needed in computing environments. We look forward to working with the NSF on this effort.

Sincerely,

Jordan Konisky

RICE UNIVERSITY · OFFICE OF VICE PROVOST FOR RESEARCH AND GRADUATE STUDIES–MS 13
P.O. BOX 1892 · HOUSTON, TEXAS 77251-1892 · 6100 MAIN STREET · HOUSTON, TEXAS 77005-1892
PHONE: 713-348-4002 · FAX: 713-348-5425 · E-MAIL: konisky@rice.edu

UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO SANTA BARBARA · SANTA CRUZ

OFFICE OF GRADUATE STUDIES AND RESEARCH
(619) 534-3555
FAX: (619) 534-3868

9500 GILMAN DRIVE
LA JOLLA, CALIFORNIA 92093-0003

March 23, 2001

Dr. Ken Kennedy, Director
Center for High Performance Software Research
Rice University
6100 Main Street (MS-41)
Houston, TX 77005

Dear Dr. Kennedy:

On behalf of the University of California, San Diego, I am pleased to offer cost sharing support for the National Science Foundation Science and Technology Center for Grid Application and Development Software. The total cash match from UCSD will be approximately $156,000 annually for the first five years from the following sources, should NSF provide funding at the level requested:

- $30,000 Department of Computer Science and Engineering
- $30,000 Fellowship support (from the Jacobs School of Engineering)
- $35,000 San Diego Fellowship graduate support (from the Office of Graduate Studies and Research)
- $10,000 Francine Berman, Computer Science and Engineering
- $10,000 Andrew Chien, Computer Science and Engineering
- $50,000 Equipment support (from the Cal(IT)^2 Institute)

Should funding be made at a reduced level, the Jacobs School of Engineering graduate student Fellowship support and the equipment support listed above will be reduced pro rata. Additional cost sharing support will be made at comparable levels for the lifetime of the award (up to ten years) should the Science and Technology Center be funded by NSF after year five.

Please let me know if you need additional information, and best wishes for success.

Sincerely,

Richard Attiyeh
Vice Chancellor for Research

THE UNIVERSITY OF CHICAGO
THE DIVISION OF THE PHYSICAL SCIENCES
5747 SOUTH ELLIS AVENUE
CHICAGO, ILLINOIS 60637

Office of the Dean
David W. Oxtoby, Dean

(773) 702-7950
FAX (773) 702-7915

March 19, 2001

Dr. Nathaniel Pitts
National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts:

The University of Chicago is pleased to participate with Rice University and the partner institutions (the University of Tennessee at Knoxville, the University of Illinois at Urbana-Champaign, the University of California-San Diego, the University of Southern California Information Sciences Institute, and the University of Houston) in submitting this proposal to the National Science Foundation for a new Science and Technology Center. At Chicago, the principal investigator will be Dr. Ian Foster. Our proposed effort is currently budgeted at $1,391,998 over the five-year life of the project.

I strongly support this proposal and expect it to have a significant, positive impact both on computer science itself and on the broader set of scientific and engineering disciplines that use computational techniques. At Chicago, I foresee strong collaborative links developing, with benefits to all involved, between the proposed center and other academic departments, as well as with researchers at Argonne National Laboratory, operated by the University on behalf of the Department of Energy. In addition, I am excited by the potential for new approaches to computer science education that may be enabled by harnessing the diverse talents and physical resources within this geographically distributed center.

The University of Chicago agrees with the underlying goals of the NSF STC program and will provide support over the life of the project, with commitments that match the goals of our institution to those of the NSF in general and to this proposed project in particular. We intend to encourage participation by both undergraduates and graduates by providing a total of $417,600 towards graduate student tuition, P.I. time and effort, and seed funding for special project requirements.

While some of these commitments represent new efforts at The University of Chicago, it is important to note that we have a demonstrated history of supporting progress in the computational sciences, in improving the diversity of student participation, in increasing the participation of American students, and in providing space and facilities to enhance interdisciplinary collaboration in research and education. Work on computational science has been enhanced in several academic departments by a very successful University of Chicago/Argonne National Laboratory collaborative grants program. Interdisciplinary work is encouraged through interdisciplinary research institutes and degree-granting "Committees"; a Committee on Computational Science is currently being formed.

We are confident that this project will be a successful Science and Technology Center. The problem to be addressed is of major importance to both computer science and science and engineering in general. The collaborators are all outstanding individuals with proven leadership skills. We look forward to working together as a partner in this important effort.

Yours sincerely,

David W. Oxtoby

hcp

cc: Stuart Kurtz
Gilda Reyes

Institutional Endorsement:

Mary Ellen Sheridan, Ph.D.
Assoc. VP for Research

UNIVERSITY OF HOUSTON SYSTEM
UNIVERSITY OF HOUSTON

Office of the Vice Chancellor for Research and Intellectual Property Management
UH System

Office of the Vice President for Research
University of Houston

March 30, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia, 22230

Dear Dr. Pitts:

The University of Houston is pleased to participate with Rice University and the other partner sites at the University of California at San Diego, the University of Chicago, the University of Illinois at Urbana-Champaign, the University of Southern California Information Sciences Institute, and the University of Tennessee at Knoxville, in submitting this proposal to the National Science Foundation in support of a new Science and Technology Center. The proposed total project costs for the University of Houston's part of the Center, with Lennart Johnsson as Principal Investigator, are $1,763,840 for the first five years. These costs are budgeted as $1,356,800 from NSF and $407,040 from the University of Houston.

The University of Houston computational and information related technologies and their use in science and engineering as well as other disciplines are an area of highest priority. The Center proposal addresses some of the core technologies in our vision of the future. We view the proven team of collaborators in this proposal as having the strong leadership and achievement on challenging research issues to successfully address the issues posed in the proposal. Several of the team members have had a significant impact on both academic research and on industry and we are convinced that the team will continue to strongly contribute to and influence the field.

We are in agreement with the underlying goals of the NSF STC program and will provide support over the ten-year life of the project with cost sharing commitments that match the goals of both the STC program and our institution. The support ($407,040), equivalent to 30% of the funds requested from NSF, will be in the form of a project coordinator along with systems support (totaling $212,266), faculty release time (totaling $44,774), and equipment (totaling $150,000) in the form of workstations, PCs and network servers as needed. The level of commitment of time and effort of the project coordinator and systems support are necessary to ensure the smooth operation of the program both technically and programmatically. Included in this cost sharing is the maintaining of high-speed network access to national and international networks.

We are confident that this project has all the ingredients for being a successful Science and Technology Center. It pulls together a truly impressive list of collaborators, with proven leadership skills, to work on a problem that will help leverage progress in many fields of scientific endeavor. We look forward to working together as a partner in this important effort.

Sincerely,

Arthur C. Vailas, Ph.D.
Vice Chancellor for Research and Intellectual Property Management
Vice President for Research

E. Cullen Building ■ Room 316 ■ Houston, Texas 77204-2163 ■ (713) 743-9104 Fax: (713) 743-9577

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research
Fourth Floor Swanlund Building
601 East John Street
Champaign, IL 61820-5711

March 26, 2001

William Marsh Rice University
6100 Main Street – MS 16
Houston, Texas 77005-1892

National Science Foundation
Science and Technology Center Program
4201 Wilson Blvd.
Arlington, VA 22230

Subject: NSF Science and Technology Center Preproposal with Rice University (University of Illinois Principal Investigator: Professor Daniel A. Reed)

To Whom It May Concern,

The University of Illinois at Urbana-Champaign is pleased to participate in a National Science Foundation preproposal with Rice University for the Center for Grid Application Development Software (CGrADS).

The research to be proposed involves advanced networking technologies that will soon make it possible to use the global information infrastructure in a qualitatively different way -- as a computational resource as well as an information resource. The goal of the Center for Grid Application Development Software is to simplify distributed, heterogeneous computing in the same way that the World Wide Web simplified information sharing via the Internet. The Center will explore the scientific and technical problems that must be solved to make Grid application development and performance tuning for real applications an every day practice.

In support of this exciting new research center, the University of Illinois commits to providing cost sharing in the amount $508,440, an amount equal to 30% of the proposed $1,694,800 to be subcontracted to Illinois by Rice University. This cost sharing will be provided in the form of both actual funds and in-kind contributions. Further, we intend to cost share at the required rate for years six through ten, if funded.

Sincerely,

Tony G. Waldrop
Vice Chancellor for Research

cc: Daniel A. Reed

USC
UNIVERSITY
OF SOUTHERN
CALIFORNIA

March 28, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Blvd.
Arlington, VA 22230

Department of
Contracts and Grants

Dear Dr. Pitts,

The University of Southern California-Information Sciences Institute is pleased to participate with Rice University, and the participating partner sites at the University of Tennessee at Knoxville, the University of Illinois at Urbana-Champaign, Indiana University, the University of California at San Diego, the University of Chicago, and the University of Houston in submitting this proposal to the National Science Foundation in support of a new Science and Technology Center. The University of Southern California-Information Sciences Institute's principal investigator will be Carl Kesselman. Our proposed effort is currently budgeted at $1,374,400 for the first five years of the project.

We are fully supportive of this proposal and believe it can have a significant and positive impact on the future of high performance, distributed computing, progress of science, quality of investigators, quality of leadership and the collaborative nature of the effort shall have valuable effect on the research produced by our institution.

The University of Southern California-Information Sciences Institute is in agreement with the underlying goals of the NSF STC program and will provide support over the life of the project with commitments that match the goals of our institution as well. The University's support will consist of allocating significant usage of USC/ISI's Virtual Reality facility for this project. This facility will be an important piece of the distributed MacroGrid for the CGrADS project. Additional support over the life of the project will be provided through the use of USC's High Performance Computing Center and networking capabilities.

While some of these commitments represent new efforts at the University of Southern California-Information Sciences Institute, it is important to note that we have demonstrated history of supporting progress in the computational sciences, in improving the diversity of student participation, in increasing the participation of American student and in providing space and facilities to accomplished researchers.

We are confident that this project has all the ingredients for being a successful Science and Technology center. It pulls together a truly impressive list of collaborators, with proven leadership skills, to work on a problem that will help leverage progress in may fields of scientific endeavor. We look forward to working together as a partner in this important effort.

Sincerely,

Lloyd Armstrong

Lloyd Armstrong, Jr.
Provost & Senior Vice President
For Academic Affairs

University of
Southern California
Los Angeles,
California 90089-1147
Tel: 213 740 7762
Fax: 213 740 6070
web page:
www.usc.edu/dept/
contracts/

Enclosures
cc: USC School of Engineering
Margie Schroeder, ISI Contract Manager
File

THE UNIVERSITY OF TENNESSEE

Office of Research
404 Andy Holt Tower
Knoxville, Tennessee 37996-0140
PHONE: (865) 974-3466
FAX: (865) 974-2805
URL:http://research.utk.edu/ora/

March 22, 2001

Director, OIA
National Science Foundation
4201 Wilson Blvd.
Arlington, VA 22230

Dear Director:

On behalf of The University of Tennessee, I am pleased to affirm the University's commitment of cost sharing on the Center for Grid Application Development Software proposal. The team of researchers is very impressive, including Jack Dongarra, who, in addition to Richard Wolski, will represent our University.

The cost sharing commitment is outlined as follows:

- Dr. Dongarra's salary, at a rate of 10%, including benefits and F&A costs
- A $15,000 commitment from the Computer Science Department to cover administrative costs
- Travel and communication funds obtained from non-federal grants, including F&A costs
- Reduced Facilities and Administrative Costs, to allow a higher percentage of funds for research activities.

The items mentioned above will account for at least, but no more than 30% of The University of Tennessee's requested budget total. We further intend to support the Center proposal by pledging these amounts for the full ten years of the proposal. The University of Tennessee realizes that the success of many projects relies on substantial gifts and commitments from various sources, and we are pleased to be able to offer our support.

Sincerely,

E. Christine Cox

E. Christine Cox
Director of Grants and Contracts

Compaq Computer Corporation
One Cambridge Center
Cambridge, MA 02142

Telephone 617-551-7600
Facsimile 617-551-7650

Cambridge Research Laboratory

Rich Zippel
Director

March 28, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts:

Compaq's Cambridge Research Lab (CRL) is very interested in the success of the CGrADS project.

We are convinced this work is critically important. The transition successful completion of this project could bring about is analogous to transition from having a small electric generator in each home or office to having centralized power source. Just as centralized power was key to the productivity, reliability, affordability and accessibility of electricity, the ability to harness the "centralized computational source" available on the web is key to the productivity, reliability, affordability and accessibility of computing resources.

From Compaq's perspective our we hope is that the early stages of this project will result in a better understanding of the very real and very current problem of optimizing computation in a varying environment. In many of today's systems the availability of compute power, storage capacity and bandwidth is undergoing constant change. One simple reason is that the systems are not necessarily dedicated to a single job.

Compaq and Cambridge Research Laboratory (CRL) have both benefited from our on-going close relationship with the partners in the CGrADS project, especially with Rice. We have a good record of recruiting both summer interns and graduates from Rice. We have long-standing technical ties. A member of the CRL research staff, Kathleen Knobe, returned recently from short term visiting research scientist position at Rice. We have donated equipment and financially supported research by both faculty members and students. Compaq hopes to maintain and strengthen our relationship with Rice. Toward this end we recently held a joint Compaq/Rice symposium.

CRL is very pleased to endorse this effort. We understand that the project is to last 10 years. Our current hope is that we could participate throughout the life of the project and we expect to continue our involvement after the NSF funding has ended. We are able to commit to supporting Kathleen Knobe for 4-6 weeks for each of the first 5 years of this effort subject to periodic evaluation of the success of the project and its continued value to Compaq.

COMPAQ

Kathleen Knobe is the best choice from our lab to participate in this project. She received her PhD from MIT in 1997. She has been working in the field of parallel systems, specifically compilers for such systems for almost twenty years. She is on the Executive Committee for Sigplan, the ACM special interest group in programming languages. In addition, she is very enthusiastic about the project.

We feel that based on both the extremely high caliber of the participants (they are the absolute top in the area of high performance parallel and distributed computation) and their track record in previous efforts, this project has a very good chance of success and we are very pleased to be able to contribute.

Please feel free to call me if you have any questions regarding this project.

Sincerely,

Richard Zippel
Director
Cambridge Research Laboratory

/br

CHIARO

March 31, 2001

Dr. Nathaniel Pitts
Director, Office of Integrative Activities
National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

Chiaro Networks
9 West Renner Parkway
Richardson, Texas 75080
Phone 972.644.4111
Fax 972.644.4101
www.chiaro.com

Dear Dr. Pitts:

Computer networks have evolved greatly over the last 10 years. From an initial view of just transferring data back and forth, today's networks should really be viewed as an extension of a single computer system to part of the symbiotic and heterogeneous architecture.

Several technologies have been the major enablers of this trend. Perhaps the most dominant is the availability of the cost effective and very high bandwidth transmission systems. Today's technologies permit transmission systems to carry data as fast if not faster, coast to coast, than previous systems provided such data transfer occurred within an integrated computer system. This last point is important—*within an integrated computer system.* Long haul fiber optic systems provided integrated hardware architecture. What is now needed is integrated software architecture to efficiently manage and provide access to geographically distributed data and applications.

If we look back to the 1980's, hardware supercomputer architectures were developed by many organizations. There were vector, very long word (VLIW), and massively parallel. But the main differentiation turned out to be ease of use, generality, and the ability to optimize user cycles as well as processor cycles. In contemporary systems, we need to apply the same efforts to large, distributed heterogeneous computing.

We need to simplify the use in the same way the World Wide Web simplified information sharing over the Internet. This requires research in many areas. These areas include:

• Software architectures that facilitate information flow among applications, libraries, compilers linkers, and runtime systems.

• Software technologies that support development and execution of performance-efficient grid applications. Among these technologies are: scheduling, resource discovery, and communications.

• Mathematical and data structure libraries for applications including numerical methods for control of accuracy and tolerant of the network latency.

• Languages, compilers, and environments that support the creation and maintenance of applications.

• New science and engineering applications that can take advantage of the new technologies as they are developed.

The objectives of the CGrADS proposal meet and exceed the objectives outlined above. In evaluating this proposal, the quality of the principal investigators (PI) must also be assessed. The accomplishments of the participating PI's would fill a book. The title of this book would be: *The Pioneers of the New Information Age in the 1980's and 1990's*. It is clear that with the establishment and full support of the CGrADS center, the title of a new book can become: *The Pioneers of Distributed Infrastructure; Hardware and Software Information Age*

Chiaro Networks will collaborate with CGrADS in many ways. One way will be assistance in understanding and utilizing the underlying networks and their topologies. Chiaro will also provide technical support and evaluation in areas such as: protocol development, security, and performance measurements.

Sincerely,

Steven J. Wallach
swallach@chiaro.com
V.P. Technology
Chiaro Networks
Richardson, Texas

Member PITAC
Member National Academy of Engineering

UNIVERSITY OF FLORIDA

Department of Physics

P.O. Box 118440
Gainesville, Florida 32611-8440
(352) 392-9264
Fax: (352) 392-8863

March 31, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts:

As Director of the GriPhyN Project, I wholeheartedly support the proposal "Center for Grid Application Development Software (CGrADS)," being submitted to the NSF Science and Technology Center (STC) program by Rice University. It is vitally important that the US seize the opportunity now to make fundamental and broad-based investments in Grid research, development and infrastructure. This university-based Center, with its long-term emphasis on developing Grid software technologies and training, will provide a key ingredient in this investment.

As you may know, GriPhyN is already deeply involved in the development of Grid infrastructure for four frontier experiments in high-energy physics, gravitational wave research and full-sky digital astronomy. These experiments will produce prodigious volumes of experimental and simulated data, with data sizes reaching 500 TB in 2001 and 100 PB by 2010. The collections, which will be nationally and even globally distributed for technical and political reasons, must be archived, processed, and analyzed by collaborations consisting of thousands of researchers at laboratories, universities and small colleges and institutes spread across the world. To meet these challenges, GriPhyN has embarked on an R&D program to develop Petascale Virtual-Data Grids (PVDGs) and plan to implement such Grids by the middle of the decade. We are also partnering with the US Particle Physics Data Grid Project and the EU DataGrid Project, as well as national efforts in the UK, France and Italy, in creating the international Virtual-Data Grid Laboratory (iVDGL), which aims to build a common global infrastructure that will support data intensive sciences.

GriPhyN is very interested in partnering with CGrADS since our projects have a lot to offer one another. Our application experiments can provide the CGrADS software research program with "real-world" environments and can offer its research teams valuable feedback and advice. Our computer science teams are also researching virtual data, job planning and scheduling, performance monitoring, and execution environments, areas which should complement and extend the CGrADS R&D progam. GriPhyN's Virtual Data Toolkit, which we will export to other scientific disciplines, would also benefit greatly from collaboration and joint development with this Center. Finally, the iVDGL provides a common infrastructure and sufficient size to conduct the experiments at the scale necessary to exercise Grid tools to their fullest extent.

Equal Opportunity/Affirmative Action Institution

The CGrADS proposal is extremely timely and exciting, and I am pleased to offer the commitment to work with Dr. Ken Kennedy and his colleagues in developing Grid tools and technology for the benefit of both our endeavors.

Sincerely,

Paul Avery

Paul Avery
Professor, Physics
Director, GriPhyN Project
University of Florida

UNIVERSITY OF CALIFORNIA, SAN DIEGO

UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

CENTER for RESEARCH in BIOLOGICAL STRUCTURE

Department of NEUROSCIENCES
Room 1000 Basic Science Building
LA JOLLA, California, 92093-0608
TEL: (858) 534-2251 FAX: (858) 534-7497
EMAIL: mhellisman@ucsd.edu

NATIONAL CENTER for MICROSCOPY and IMAGING RESEARCH at San Diego

A NIH Supported Resource Center for High Voltage Electron Microscopy and Image Analysis • established and sponsored by the National Center for Research Resources

March 19, 2001

Professor Ken Kennedy
Department of Computer Science
Rice University

Dear Ken,

As a neuroscientist who has worked extensively with Fran Berman, Carl Kesselman, Rich Wolski, Jack Dongarra and others in the emerging grid community, I wholeheartedly endorse your timely effort to create a science and technology center for "GrADS" (Grid Application Development System). I am eager to continue to work with the CGrADS P.I.s and will be happy to contribute challenging test-bed projects requiring extensions of tools and concepts for both computational and data grids. I understand that the project's main goal is to prototype a dynamic grid-aware programming environment. I believe that the development of a program and execution environment that supports adaptive and performance-oriented computing is critical to the wide-spread use and effectiveness of the grid. The grid is crucial for achieving critical advancements of ongoing projects at the National Center for Microscopy and Imaging Research (NCMIR) as well as the broader multiscale goals of the Center for Research in Biological Structure (CRBS), an organized research unit at UCSD whose purpose it is to integrate research activities in biological structure and function from atoms to organisms largely through enabling technologies from computer science. Significant recent progress has been made by collaborating groups at NCMIR, CRBS, SDSC and NPACI partners in addressing technological challenges of multiscale biological systems, like the brain. These grid-enabled advances in biological data representation and understanding have led the National Center for Research Resources of the NIH to begin construction of an expanded infrastructure for computational and data grid project to build a Biomedical Imaging Research Network (BIRN). CRBS and SDSC will likely serve as the central coordinating site for this new NIH initiative and one can expect substantive interaction between BIRN and GrADS should your program also receive funding.

Another very attractive feature of your program is the strong linkage to some of the clear successes of the NSF's PACI program that can be expected to provide important leverage for the GrADS Science and Technology Center projects. By coordination with the two PACI's your group will have even more widespread impact on the community. The consortium of grid experts gathered in this GrADS project (which include yourself, Andrew Chien, Ken Kennedy, Dan Reed, Ian Foster, Carl Kesselman, Jack Dongarra, Rich Wolski, and others) are among the brightest and most forward looking I've met in this community. There is no doubt that with such a stellar collection of coworkers you will accomplish the task of developing a program and executing the environment envisioned for the grid with unparalleled professionalism and creativity.

Best wishes on your proposal effort. I look forward to an exciting collaboration.

Sincerely,

Mark H. Ellisman
Professor of Neurosciences and Bioengineering
Director, National Center for Microscopy and Imaging Research
Director, Center for Research in Biological Structure

UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

Shankar Subramaniam
PROFESSOR
DEPARTMENTS OF BIOENGINEERING &
CHEMISTRY AND BIOCHEMISTRY &
DIRECTOR, BIOINFORMATICS GRADUATE PROGRAM &
SENIOR FELLOW, SAN DIEGO SUPERCOMPUTER CENTER
9500 GILMAN DRIVE
LA JOLLA, CA 92093-0427

TELEPHONE: (858) -822-0986
Fax: (858) 534-8380
E-MAIL: shankar@ucsd.edu

March 12, 2001

Professor Fran Berman
Director, San Diego Supercomputer Center
University of California at San Diego
La Jolla, CA 92093

Dear Fran,

I am writing to enthusiastically support your GraDS project and proposal, as well as express gratitude at your undertaking this most important project for computational and information sciences.

The post-genome era in biology and medicine warrants extensive interdisciplinary and integrative research efforts between computational and information sciences and biology and your proposed project will be a paradigm for such endeavors. As you know, my laboratory is engaged in one of the largest post-genome projects involving multiple institutions and multiple investigators. This project titled Alliance for Cellular Signaling is developing very large quantities of cellular signaling data that is of paramount importance to provide a quantitative understanding of the way our cells respond to stimuli. The complex nature of this project warrants extensive information and data communication, data modeling and intensive computational analysis. We are in desperate need of an environment such as the one you are proposing and we would like to be an early application.

I am looking forward to a fruitful collaboration.

Best regards.

Sincerely yours,

Shankar

Shankar Subramaniam

UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO SANTA BARBARA · SANTA CRUZ

NATIONAL PARTNERSHIP FOR ADVANCED COMPUTATIONAL INFRASTRUCTURE
SAN DIEGO SUPERCOMPUTER CENTER
(858) 534-5000

9500 GILMAN DRIVE, MC 0505
LA JOLLA, CALIFORNIA 92093-0505

Dr. Ken Kennedy
c/o Prof. Fran Berman
Director, San Diego Supercomputer Center
University of California San Diego
9500 Gilman Drive
La Jolla CA 92093-0505

March 17, 2001

Dear Dr. Kennedy:

The program development and execution system to be developed within the GrADS Science and Technology Center promises to be a major contribution to biologists. I would be very interested in being involved as an application developer for the GrADS system. We are presently inundated with biological data at the molecular level. While the human genome has captured the attention of the public the lesser know fact is that a new genome of a lesser species is being released every 1-2 weeks. At the same time data on gene expression, protein structure, protein-protein interactions, complex biological pathways etc. are being made public at exponential rates. Our approach to processing such data is to establish pipelines of existing software components developed by us and others, rigorously benchmark and test those pipelines, and then process large amounts of data in this way. Genome annotation and protein models are two results we are currently seeking through pipelining. The latter will provide targets in drug design experiments.

These pipeline codes are amenable to Grid computing and would be greatly improved by developing them in an environment which promotes their performance. Given the expertise of the GrADS Science and Technology Center proposal PIs, I would be eager to collaborate with this group to target biological codes we plan to develop to perform well in Grid environments. This effort should complement funded efforts and provide us an expanded platform on which to run our codes. I wish you luck with the proposal.

Sincerely,

Philip E. Bourne PhD
Professor of Pharmacology
Project Leader, The Protein Data Bank
Senior Principal Scientist, San Diego Supercomputer Center

Myricom

Myricom, Inc.
325 N. Santa Anita Ave.
Arcadia, CA 91006

626-821-5555
Fax: 626-821-5316
http://www.myri.com

29 March 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts:

Myricom, Inc., enthusiastically endorses the CGrADS project as an effort that will have a significant and positive impact on the progress of computer and computational science. The research objectives of the CGrADS will help in the development of standards for distributed and cluster-based computing. The development of such standards are critical for the field.

The project leaders are an impressive team of world-class researchers whose long-term commitments to distributed computing are legend. The success of the CGrADS project seems to me to be almost a foregone conclusion.

Myricom has benefited from on-going relationships with several of the partners in the CGrADS project. We were pleased to be asked to join the effort on this proposed NSF Science and Technology Center. We are already contributing people, equipment, and funding in the pursuit of distributed-computing research with colleagues at the University of Tennessee, and expect this relationship to expand and to continue for the duration of the project and beyond.

Sincerely,

Charles L. Seitz, Ph.D.
President & CEO

Los Alamos National Laboratory

Computer & Computational Sciences Division
P.O. Box 1663, Mail Stop B297
Los Alamos, New Mexico 87545
(505) 665-4700 / FAX: (505) 665-0120

Date: April 2, 2001

National Science Foundation
Director, Office of Integrative Activities
4201 Wilson Boulevard
Arlington, Virginia 22230

Dear Dr. Pitts,

We are very excited by the prospect of NSF support for the Center for Grid Application Development Software (CGrADS). CGrADS is a natural, ambitious and critical scientific enterprise that complements our collaborative work with Rice University through the Los Alamos Computer Science Institute. The computational grid is quite likely to be the general computing environment of the future. The obvious challenges - resource negotiation, scheduling, communication, languages, compilers, PSEs - are all very interesting. However, the most intriguing are the fundamental intellectual and research questions posed by this center, specifically the notion of a performance economy and the development of performance and resource forecasts. The results of this work are certainly important from an operational viewpoint, but this new environment may also change the way that applications are constructed forever.

To accomplish this daunting task, CGrADS has assembled a very impressive team. Ian Foster and Carl Kesselman provide expertise and links into the Grid work (e.g. Globus) that provides the foundation for CGrADS. Ken Kennedy is among the nation's leaders in languages and compilers. Jack Dongarra and Lennart Johnsson are world leaders in the construction of mathematical libraries for scientific computation. Now is the time to focus the attention of the computer science community on the issues of grid computation.

Why is this work important to the country? The two highest barriers standing in the way of general adoption of parallel computation are lack of (1) computing resources dedicated to a specific application and (2) powerful domain-specific problem solving environments. CGrADS will provide mechanisms for addressing both of these issues. In our work in crisis forecasting (e.g. wildfire, epidemiology, electrical power), these two issues have seriously affected our ability to transfer technology to parties charged with crisis management and mitigation. The existence of the CGrADS execution architecture that can be accessed by a domain PSE and that can then deliver distributed, heterogeneous computing resources not wholly-owned by the domain science organization is exactly what the doctor ordered. This is an essential capability in establishing predictive modeling and simulation as a regular part of the decision-making process.

The Center for Research on Parallel Computation (CRPC) is the perfect model for the management, education and outreach programs in CGrADS. CRPC's successful distributed management structure can also effectively integrate the CGrADS activities. Also, Richard Tapia has developed nationally celebrated programs for underrepresented groups and K-12 education. With Richard directing CGrADS's education and human resources activities they are certain to be innovative and effective.

Not surprisingly, CGrADS is of significant interest to Los Alamos above and beyond its specific technical goals. We are working to establish an Open Collaboration Network (OCN) -- in addition to our production green, yellow, and red partitions -- which will enable more effective collaboration with the external computer science and technology community. The OCN will be the interface at Los Alamos with the CGrADS effort. Also, it is clear that the computer room of the future will be composed of several generations of hardware and software that must effectively work together in solving scientific computing problems. One by-product of CGrADS will be an effective heterogeneous computing environment that will allow more effective utilization of existing computing resources.

In conclusion, we believe that CGrADS brings together the right team at the right time to focus on constructing a usable global computational resource.

Sincerely,

Andrew B. White, Jr.

Andrew B. White, Jr.
Director, Computer and Computational Sciences Division (acting)
Los Alamos National Laboratory

Cy. File