Proposal for an Interdisciplinary Laboratory for Scientific Computing

Scientific computing enhances our understanding of such diverse topics as the cosmos, genomics, the atmosphere and oceans, biodiversity, aerodynamics, cognitive science and linguistics, forest fires, and the weather. The increased use of computational methods in the sciences presents us with three challenges: the need for powerful hardware, efficient software, and the expertise in numerical methods. We wish to partner with Sun Microsystems, Inc. to create an Interdisciplinary Laboratory for Scientific Computing (ILSC) at UMass-Dartmouth which will meet these challenges by providing high quality computational resources and software and creating an interdisciplinary environment for scientists from universities and industry and computational mathematicians to collaborate on scientific problems requiring computational tools, and a resource for educating and training the computational scientists of tomorrow.

1 Scientific Computing

The last few decades have seen a dramatic increase in the speed and data handling capability of high performance computers, and the development of novel algorithms, which together have transformed scientific investigation. Scientific computation has joined experimentation and theory, to become central to scientific development and discovery. Large scale computational algorithms are utilized in virtually every scientific field. In biology, examples include the determination of protein structures from data given by nuclear magnetic resonance, x-ray crystallography, or electron microscopy, parallel computation of protein folding models and biomolecular dynamics simulations, and simulation of biological processes. Chemists use computing techniques extensively in spectroscopy, molecular modeling of organic and inorganic molecules, quantum chemical calculations, molecular dynamics and mechanics, and drug discovery. In physics, astrophysical and oceanographic models are simulated using some of the most advanced and complex numerical algorithms. Weather prediction is based on the numerical simulation of weather models. Engineers (mechanical, electrical, and computer) are by far the most sophisticated users of scientific computing, and applications such as computational fluid dynamics and image processing require extensive computational time and memory, and yield high level information for a wide variety of applications.

The overlap between computational research fields is fertile ground for scientific advancement. For example, in the field of computational linguistics, the intersection of computer science (in particular, artificial intelligence) and the cognitive sciences yields complex theories about linguistic knowledge that is required for understanding and generating language, as well as software for allowing computers to communicate with people. Amazingly, computational linguistics also has an impact on the protein folding problem in biochemistry; there has recently been developed a grammatical theory for problems in protein folding. These complex interactions between seemingly disparate fields such as mathematics, computer science, linguistics, bioinformatics, and biophysics highlight the need for interdisciplinary collaborations in the computational sciences.

The increased importance of computing in the sciences poses new challenges, as well. The first challenge is the availability and efficient use of computational facilities. Scientists need powerful high performance parallel computers to perform numerical simulations, and these computers need to be efficiently configured and connected. The second challenge is in providing the appropriate software for the problem. There are many commercial as well as free software packages available for a large variety of problems. It is necessary to identify the appropriate software packages and make them available for use by science and engineering faculty and their graduate students. The third challenge is in providing the experts in numerical methods who can direct and advise scientists on the use and limitations of numerical methods and innovate and develop new numerical methods tailored to a given scientific problem.

2 Goals and Vision

The UMass Interdisciplinary Laboratory for Scientific Computing (ILSC) will meet the challenges of the era of scientific computing by providing scientists with the hardware and software needed for numerical simulations, as well as enabling collaborations between the scientists and the numerical analysts. These collaborations will foster research activities that develop and utilize novel computational algorithms to aid scientific discovery.

1. **Hardware:** The ILSC will house a cluster of parallel workstations, which will be available for numerical simulations of various types of physical and engineering problems.

   Powerful workstations have become affordable in recent years, but are significantly more expensive than a desktop personal computer. These machines cannot be used efficiently if they are concurrently used for email, and internet and word-processing applications. It is essential that the machines be solely dedicated to running numerical simulations. Furthermore, parallel computing is needed for simulations which have large scale computation, data, and storage needs. For this, we require a parallel cluster of computers with large data capabilities.

   We are asking for the following equipment for the ILSC:

Quantity	Description
32	SUN Fire X4100:
	2 Opteron 254 CPUs, 8GB, 2x73GB SAS, DVD, 2PSU. A64-NPB22H8GCB7-IP
2	same as above, for communicating with storage
1	same as above, management node
1	PCI dual FC 2 Gb host Bus adapter SG-XPC12FC-QF2
2	SUN StorEdge 3511 12X500GB 7200 TA3511MO1A0A6000
35	Slide rail kit for SUN Fire X8029A
35	Topspin Interconnect 1235A
1	keyboard/monitor/mouse shelf unit (1RU) 1511A
1	SunRack 100-42 MPS 60 Amps SRS-1042-60TPL
1	Solaris 10
1	N1 grid engine 6 license for max of 50 CPUs
35	Sun N1 system manager license for server with up to 4 sockets

2. **Software** The ILSC will provide reliable software for scientific computing.

   The affiliates of ILSC will identify efficient software, which will be installed and tested on the workstations in the lab. The ILSC has affiliates who have the expertise to advise users on each package's use and its limitations. This will prevent the computer and human time loss which results from using an application that is ill-suited to a problem. There are many software packages available for scientific computing, a few examples include:
   HPF: The High Performance Fortran Forum (HPFF), a coalition of industry, academic and laboratory representatives, works to define a set of extensions to Fortran 90 known collectively as High Performance Fortran (HPF). HPF extensions provide access to high-performance architecture features while maintaining portability across platforms.
   PETSc is a suite of data structures and routines for the scalable (parallel) solution of scientific applications modeled by partial differential equations. This software is freely available.
   FLUENT is a finite volume based, commercially available computational fluid dynamics software.
   PseudoPack (2000) is a freely available support package for spectral methods codes for solution of partial differential equations, and provides subroutines for performing basic operations such as

generation of proper collocation points, differentiation and filtering matrices.
USEMe is an unstructured spectral element method code, which can be obtained from the author at no cost. It is designed for simulations of electromagnetics, fluid dynamics, and plasma dynamics. It can be implemented in parallel or serial.
GAUSSIAN is software for molecular modeling and chemical reaction dynamics, developed by Gaussian Inc.
MATLAB
MAPLE
MOLGEN is an efficient and portable tool for molecular structure elucidation for chemical research.
Molpro is a complete system of ab initio programs for molecular electronic structure calculations.
MSI software by Molecular Simulations Inc. for homology modeling and dynamics software (this software is extremely expensive and would not be purchased until the need and funding arises).
NAMD Molecular Dynamics Software is a program which runs molecular dynamics simulations. This program is designed to run on parallel machines. It is available free of charge for non-commercial use by individuals, academic or research institutions with completion of an online registration form.
Praat speech analysis software includes speech analysis (including spectral analysis, pitch analysis, formant analysis, intensity analysis, cochleagram, and excitation patterns) and speech synthesis. This software is freely available.
PredictProtein is a free service for sequence analysis, and structure prediction.
PSI is a free package of ab initio quantum chemistry software developed at the center for computational chemistry at the University of Georgia.
VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting. VMD supports computers running MacOS-X, Unix, or Windows, is distributed free of charge, and includes source code.
X-PLOR and **CNSsolve** are programs for structure determination from data obtained by x-ray crystallography and nuclear magnetic resonance. They are free to academic institutions.

Additionally, there is software necessary for visualization, and there will be more software added to the list as the need arises. There are problems for none of the available software is sufficient, and our affiliates are committed to developing softwares as part of their research. This software will then be released to the public through the lab web site.

Although some of the software must be purchased, a lot of it is freely available. However, unlike the software we use for everyday applications on a personal computer, this software is not self extracting and installing. Software must be downloaded and installed. Proper installation and maintenance is essential to assure that programs run properly in parallel. The technician will be able to support unix based computers outside of the lab, which are being used for research by our affiliates.

3. **Human Expertise: Collaboration and Innovation:** The ILSC hardware and software resources will be significant, but the most valuable aspect of the lab is the human capital, the ILSC affiliates.

   The ILSC's affiliates are scientists and applied mathematicians. We typically classify the scientists as experts in the physical problem, and the applied mathematicians as experts in the numerical methods. However, the ILSC affiliates have significant experience in both aspects of numerical simulation. The ICSC will bring together scientists in many fields as well as mathematicians, and the impact on each field will be greater than the sum of its parts.

   On its most basic level, the ILSC will bring together experts in physical problems and experts in numerical methods to identify numerical methodologies for the simulation of a given model. They will determine whether there are software packages which can be used for a given simulation, whether additional software is available for parts of the code which need to be written, or whether they must develop new code. This type of collaborative environment will prevent inappropriate choices of software for a given problem or choices of parameters or options that are not optimal. High performance computing requires an understanding of the underlying numerical methods as well as the parallel structure of the method. The ILSC will make the expertise of its affiliates available to a wide variety of scientists. There are many scientists who have, through training as well as years of experience and numerical experimentation, developed significant expertise in the use and function of numerical methods, and these scientists will share their expertise with other scientists as ILSC affiliates.

These technical collaborations will provide a springboard for more innovative collaborations. Once a physical problem is identified, the scientists involved may realize that it is necessary to tailor a numerical method for simulation of the problem, or perhaps create a new numerical method for the problem. This will lead to a collaboration which is innovative in both the physical and mathematical aspects. The opportunity for such collaborations is important to scientists and applied mathematicians, and the ILSC will provide the setting for truly innovative collaborative research.

The lab will be headed by Steven J. Leon, Sigal Gottlieb, and Jae-Hun Jung. The affiliates of the ILSC will be numerical analysts, engineers, physical scientists, computer scientists, and biological scientists. The affiliates will serve on a board of directors, which will guide the lab's development and direction. A graduate student will be funded as a technician and coordinator. Her/his responsibilities will include management of the workstations, downloading and testing of software, and setting up a website for the lab.

3 Affiliates of the ILSC

Although the need for computational resources is critical for the existence of the ILSC, its excellence will depend on its affiliates. Our affiliates represent a broad range of research interests which require high performance computing, including mathematics, physics, biology, chemistry, engineering, marine sciences, and computer information science. Many of our affiliates perform research which could benefit from high performance resources, such as parallel clusters for computing and visualization. In the past, the scope of our research and the choice of research interests has been limited by the absence of these computational resources on campus; the existence of the ILSC will encourage and stimulate more research which relies on high performance computing.

The following is a list of UMass-Dartmouth scientists who will be the initial affiliates of the ILSC. This list will grow as collaborations are formed and as we extend the lab to include the entire UMass system.

Ramprasad Balasubramanian *(Computer and Information Sciences)*: Computer Vision, Image Processing, Artificial Intelligence and HCI.

James Bisagni *(Marine Sciences)*: Global Ecosystems Dynamics.

Donald Boerth *(Chemistry/Biochemistry)* Molecular modeling, Molecular orbital self-consistent field calculations, Density functional theory, Modeling pesticide interactions with DNA molecules, Protein and nucleotide modeling for drug and pesticide design.

Changsheng Chen *(Marine Sciences)*: Ecosystems Dynamics Modeling.

Chi Hau Chen *(Electrical and Computer Engineering)*: Image Processing, Machine Vision.

Alex Fowler *(Mechanical Engineering)*: Fluid Flow in Porous Media, Heat Transfer, Thermodynamics, Bioengineering.

David Goodson *(Chemistry)*: Quantum Chemistry Methods Development, Quantum Molecular Dynamics.

Sigal Gottlieb *(Mathematics)*: Numerical Analysis, Parallel Computing, Shock Wave Calculation using WENO and Spectral methods, SSP Runge-Kutta methods, Gibbs phenomenon.
Jong-Ping Hsu *(Physics)*: Theoretical Physics: symmetry principles and gauge field theories.
Jae-Hun Jung *(Mathematics)*: Numerical Analysis, Parallel Computing, Spectral Methods, Gibbs phenomenon removal, Computational Fluid Dynamics.
Dayalan Kasilingam *(Electrical and Computer Engineering)*: Remote Sensing, Applied Electromagnetics, Wireless Communications, Adaptive Signal Processing.
Gaurav Khanna *(Physics)*: Theoretical and Computational Astrophysics, Black Hole Astrophysics, Gravitational Waves, Quantum Gravity, High-Performance Computing, Control and Dynamical System Theory.
Raymond Laoulache *(Mechanical Engineering)*: Thermodynamics, Multiphase Flow, Control Systems, Fluid Mechanics, Laser Doppler Anemometry, Parallel Computing.
Steven J. Leon *(Mathematics)*: Numerical Analysis, Inverse Problems, Matrix Computations.
Steven C. Nardone *(Electrical and Computer Engineering)*: Systems Theory, Control and Estimation Theory, Fuzzy Systems, Applications to Target Tracking.
Li Shen *(Computer Information Science)*: Image processing and computer vision, Pattern analysis and machine intelligence, Data mining and bioinformatics, Geometric modeling and graphics.
Eli Stahl *(Biology)*: Molecular Population Genetics, Population Structure, Conservation Genetics, Molecular Evolution, Disease Resistance.
Timothy C. Su *(Chemistry and Biochemistry)*: Ion Chemistry, Mass Spectrometry, Computational Chemistry, Drug Discovery.
Amit Tandon *(Physics)*: Physical Oceanography, Fluid Dynamics, Computational Physics, Motion Analysis and Visualization.
Marguerite Zarillo *(Physics)*: Transportation Engineering, traffic simulation, modeling and computing.

For more information about the ILSC affiliates, please look at the temporary ILSC web site: http://www.umassd.edu/cas/mathematics/ilsc

4 Plan for Future Funding

The ILSC will apply for federal grants to fund workshops and short courses in the computational sciences, which will be run by the affiliates of the lab. These workshops will raise the visibility of the ILSC, and of high performance computing and computational science, and will provide a rich basis for collaborations.

The lab will also benefit from research grants made to affiliates of the ILSC, to support research at the ILSC. These grants will pay for additional computational resources as they become necessary, for the funding of a technician, and for new and upgraded software.

5 Research Partnership with Sun Microsystems, Inc.

The ILSC will provide a springboard for collaborations within UMass Dartmouth, throughout the UMass system, and with scientists from Sun Miscrosystems. The availability of high performance computing resources will stimulate UMass scientists to work with Sun Microsystems scientists on problems of mutual interest, and will produce interest in new problems which importance to Sun researchers. The contact with UMass faculty and students will provide opportunities for Sun researchers to collaborate on research and thesis projects of mutual interest. The academic-industrial partnership will be beneficial to all the researchers, as well as for the students trained in such a partnership.

6 Benefits of a Scientific Computing Laboratory

The ILSC will have both educational and scientific impact on UMass Dartmouth and to the entire UMass system. It will generate scientific collaborations and educational opportunities on campus and throughout the scientific computing community. It will encourage and stimulate more research in computational science by raising the visibility of this field and providing the high performance computing resources necessary. It will also impact on the international academic community by providing research software developed by its affiliates. Whenever reasonable, this software will be freely available on the lab's website.

The ILSC will serve as a cornerstone of the proposed masters in scientific computing. This masters program is being proposed by the Mathematics Department and is designed as an interdisciplinary program in scientific computing in its broadest sense. It will involve the students in research in a wide variety of computational fields, including physics, chemistry, biology, engineering, linguistics, computer science, and mathematics. The existence of the ILSC and the partnership with Sun Microsystems will also help in the planning and development of a proposal for a system-wide graduate program in computational science, which will include research in any scientific discipline which benefits from high performance computing. This program will take the ideas used to propose the masters in scientific computing and extend them to include other departments, both at UMass Dartmouth and in the entire UMass system. An important feature of both these programs will be the opportunity to work jointly with partners from Sun Microsystems.

The benefits to the UMass system will extend to the Commonwealth, as well as the Nation, through collaborations and through the training of computational scientists.

1. **UMass Dartmouth and the UMass System**.

   Scientific impact: The interdisciplinary collaborations fostered through the ILSC will lead to more sophisticated computational models which will in turn lead to improved research results. Since interdisciplinary projects tend to be broader in scope, they are generally funded at a higher level. The increased funding the ILSC brings in can be used to further support research at the university. Once the lab is up and running we plan to make it accessible to faculty at our sister UMass campuses. Thus the laboratory will not only facilitate interdisciplinary research, but also intercampus collaborations. The entire UMass system will benefit.

   Educational impact: Graduate students will benefit from the improved facilities for doing modeling and simulations and from the opportunities to work on interdisciplinary projects and to consult with experts on numerical computations. At UMass Dartmouth, the disciplines represented by the affiliates of the center have a total of 286 graduate students as of Spring 2006, and these numbers will increase as more graduate programs, including the newly approved system-wide graduate program in bioengineering, start accepting students.

   Undergraduate students will also benefit from increased opportunities to perform research in the computational sciences. At UMass Dartmouth, undergraduate education in the sciences includes many opportunities for research. Each year, close to 90 students present their research in the annual Sigma-Xi Research Exhibition on campus. Undergraduate research is an important component of computational science education. The presence of the ILSC on campus will create new opportunities for training undergraduate and graduate students, and post-doctoral researchers in computational science.

2. **The Southcoast region and the Commonwealth of Massachusetts**. Currently UMass Dartmouth faculty are actively involved in research that impacts local industries such as the fishing industry and the cranberry industry. Improved mathematical models and simulations in these fields will lead to improvements in harvesting and conservation that will have positive impact on the economy of the region. Through the ILSC our graduate students will be better trained for high tech jobs

in the region and for government jobs at facilities such as the Naval Underwater Warfare Center in Newport.

Modern biological models are extremely complex and involve massive amounts of computation. The ILSC will provide the facilities for these computations and will help train students in biological modeling. Through both research and the training of researchers the lab will contribute significantly to the build up of the biotech industry in Massachusetts. We will also train graduate students to help meet the staffing needs of the many high tech software companies in Massachusetts.

3. **The nation as a whole**. The ILSC will foster interdisciplinary research that will have a national impact. Working through the lab, faculty will be able to form federal partnerships to help meet national goals. Improved mathematical models and simulations can lead to advances in health care, improvements in homeland security, better environmental models, improved structures and construction, etc. These partnerships are offered through programs in the departments of agriculture, defense, energy, transportation and through agencies such as the National Institute of Health, the National Science Foundation, the Environmental Protection Agency, and the National Oceanic and Atmospheric Administration.