Next Generation Parallel Codes for the Simulation of Correlated Materials

M. Jarrell, J. Moreno and J. (Ram) Ramanujam

The Project. Strongly correlated electronic materials, including spintronic and high-temperature superconducting materials, have many promising applications. The 2007 International Technology Roadmap for Semiconductors stresses that these systems can enable new devices by greatly enhancing their sensitivity to different applied fields. These materials also have promising applications in energy conversion, storage and transportation. This progress relies on our ability to optimize the properties of these compounds, e.g., to make a room-temperature superconductor or a ferromagnetic semiconductor. Optimization of their properties, in turn, requires a complete theoretical understanding of these materials. Unfortunately, these compounds display complex emergent phenomena and competing phases which inhibits the use of conventional reductionist theory, so that progress will only be achieved with significant computational modeling

Strongly correlated materials are described by self-consistent field theories, which now may be parameterized by numerical calculations. The latter are used to provide complete information about the material on the short length scales while the former provides the intermediate and long length scale information. In general, such multiscale methods are more accurate when they are parameterized by *both* one (e.g., photoemission spectra) and two-particle properties (e.g., susceptibilities to different external fields). However, until recently, the solution of a complete set of two-particle equations has not been possible, since it involves the rotation and contraction of large rank-three tensors. Consider a single-band system with N=40 sites and L=100 time slices. The corresponding two-particle rank-three tensors (susceptibilities, vertices, etc.) require nearly 2TB of storage each.

However, early petascale systems, such as the national leadership class machines at ORNL (Jaguar) and on the NSF Teragrid (Kraken) will allow us to store this needed data. But, efficient computational methods for rank-three tensor rotation and contraction are needed to solve the self-consistent field theory equations. The development of low-level high-performance parallel programs for them is usually very tedious and time consuming. We have developed techniques in the Tensor Contraction Engine project to address the following: (i) use algebraic transformations to reduce the number of operations; (ii) minimize the storage requirements to fit the computation within the storage disk limits by using compiler transformations. Additional specific techniques are need to: (i) reduce memory access costs by minimizing disk-to-memory and memory-to-cache traffic; (ii) enhance parallelism, and (iii) develop a strategy to call BLAS library to implement tensor operations.

A part of this research is funded by a DOE SciDAC grant and computer time is provided by a large INCITE allocation on Jaguar and a TeraGrid allocation on Kraken. Computer resources available for this project include more than 25M hours of CPU time on petascale machines. We are presently writing an NSF CDI proposal to support this effort. The purpose of this request is to acquire the resources needed to extend this collaboration to involve additional CCT faculty and LI researchers, thereby building a truly Louisiana focused interdisciplinary collaboration. The CS will collaborate with us, the students and postdocs in our groups, and researchers throughout the state involved in a recent EPSCoR application. The CS will be a coauthor/coinvestigator on related publications and grant applications.

Required Effort. Roughly 6 months of the CS's time will be required to integrate the Tensor Contraction Engine into existing massively parallel codes, optimize the codes, and raise the level of abstraction used in order to create portable, modular and sustainable software.

How the project will benefit the LI. It will make the CCT and LI more visible on national leadership class machines thereby greatly enhancing their national and international exposure. Jarrell and Moreno recently moved to LSU, allowing this SciDAC project to acquire an LA focus. It will generate a greatly improved publically available code in an important area of research for both new devices and energy technology.

LI Proposal: Parallel Algorithms for Large Scale Data Clustering

Dr. Vijay Raghavan, Dr. Ryan Benton, Center for Advanced Computer Studies, UL Lafayette
Dr. Raju Gottumukkala, Dr. Ramesh Kolluru, NIMSAT Institute, UL Lafayette
Dr. Box Leangsuksun, Department of Computer Science, Louisiana Tech University

Clustering attempts to group similar objects together based on certain similarity. Traditional clustering algorithms were originally designed for small data sets. Given the exponential growth of data (text, images and video) obtained from advanced instruments, sensors and the Internet, the machine learning communities are exploring ways to refactor existing learning techniques to exploit the computational opportunities found through parallel processing (via cores, clusters, and clouds). In addition, it is known that no single clustering algorithm is capable of successfully grouping all types of objects; each clustering algorithm has a 'bias' that shapes how objects are perceived and grouped. A solution is to generate several sets of clusters via different techniques and select the best solution. This assumes that the user has enough knowledge to determine what the best solution is. Since this is rarely true, the use of ensembles of clusters has been proposed; ensembles of clusters begin with the generation of several different sets of clusters. Then, the predictions of the clusters are then combined (similar to merging the predictions of experts) into a final outcome. Traditionally, using ensembles is quite expensive; a potential solution is via the use of distributed computing to mitigate the impact.

One application lies in finding spatial regions in the brain that characterize stages of dementia. In this case, Positron Emission Tomography (PET) imagery is converted into a series of 3-D points, with each point representing metabolic activity. By grouping points together, based on activity and location, one could isolate regions of the brain that are indicative of progression of dementia. Other dementia-related applications lie in clustering information gathered through exams such as Functional Assessment Exam and the Mini-Mental State Exam. In addition to dementia, other target clustering applications include the spatiotemporal clustering of satellite imagery for disaster impact and characterization of highway accident data.

Goals

The main goal of this project is to investigate optimal approaches to parallelize various clustering and association mining algorithms both on LONI and a prototype cloud cluster at the Center for Advanced Computing Center (CACS). These algorithms help multiple interdisciplinary projects including Alzheimer's Disease Neuro Imaging Initiative (ADNI) datasets for characterizing and predicting dementia, spatiotemporal clustering of geospatial imagery and text mining of web data.

Effort Requested of LI Computational Scientist

We request 30 hours per month of effort for one year (FTE of 2 months) of the LONI Computational Scientist, Dr. Raju Gottumukkala's for this project. The LONI Computational Scientist would help with investigating the best approaches to improve the performance of various data mining algorithms using MPI, OpenMP, and Map/Reduce frameworks.

Benefit to LONI Institute

The research project will benefit multiple projects. In general the developed applications will provide the MPI/OpenMP implementations of data mining algorithms for LONI users upon validation. This project will also potentially advance the LONI infrastructure to support data parallelization using Map/Reduce frameworks such as Hadoop.

Parallel Optimization Algorithms for Disaster Management

Dr. Ramesh Kolluru, Dr. Mark Smith, Dr. N. Raju Gottumukkala, NIMSAT Institute, UL Lafayette

Dr. Baker Kearfott, Department of Applied Mathematics, UL Lafayette

Dr. Dileep Sule, Department of Industrial Engineering, Louisiana Tech University, Ruston

Background

Optimization refers to comparison of a number of solutions before arriving at the optimal solution. With the increase in the number of variables and the number of objectives that have to be satisfied at the same time, the evaluation of a solution can be time consuming because an algorithm has to consider all possible alternatives. Decision support applications for emergency management are time critical and need to be highly interactive. Therefore, parallelizing various optimization algorithms that are applicable in various decision support tools would significantly reduce the computation time. Existing GIS based tools like ESRI's ArcMap, Google Earth, and Hazards US (Hazus) used by emergency managers and first responders do not support analysis of large-scale and complex spatiotemporal datasets or exhaustive planning simulations. The effectiveness of disaster response and decisions are noticeable post-response when major catastrophes have already happened, hundreds of thousands of people were stranded in traffic due to unavailability of gas, and millions of dollars were spent in mismanagement of resources or not having prioritized the recovery of critical infrastructures.

Goals

The NIMSAT Institute works on multiple research and development projects from multiple emergency management agencies including the Louisiana's Governors' Office of Homeland Security (GOHSEP), Department of Natural Resources (DNR) and Department of Homeland Security (DHS). These projects have various components that require developing optimization algorithms for enhancing disaster preparedness and response. The main goal of this project is to develop parallel optimization algorithms on LONI for enhancing various disaster management applications such as evacuation models, site selection algorithms and infrastructure interdependency analysis.

Effort Requested of LI Computational Scientist

We request 6 months of FTE for the LONI Computational Scientist, Dr. Raju Gottumukkala's time for developing algorithms and parallelizing them on LONI for multiple projects at the NIMSAT Institute. The institute was recently funded through the Louisiana's DNR to develop a fuel demand model for evacuations through analyzing the historical traffic patterns, real time traffic feeds and assessing the evacuation behavior. Raju has been involved with the development of the fuel demand model and would enhance it and eventually investigate approaches to improve the performance of the algorithm by running this algorithm on LONI.

Benefits to the LONI Institute

This project involves researchers from multiple disciplines including transportation modeling, computer science, industrial engineering and applied mathematics. This project also offers a unique opportunity to improve the decision support tools of emergency managers through cyberinfrastructure as most HPC projects in the nation have been about hazard or weather prediction. In addition to current funding, this project has a great potential for federal funding from NSF, DHS and FEMA.

iLevee: Intelligent Flood Protection Monitoring, Warning and Response System

Dr. Ramesh Kolluru, Mr. Dean Mallory, Dr. N. Raju Gottumukkala,
NIMSAT Institute, University of Louisiana at Lafayette
Dr. Box Leangsuksun, Computer Science Department, Louisiana Tech University
Dr. Honggao Liu, Director of HPC, Louisiana State University, Baton Rouge, LA

Background

The State of Louisiana Department of Natural Resources, Office of Coastal Protection and Restoration (OCRP) plans to deploy a state of the art Intelligent Flood Protection Monitoring, Warning and Response System (IFPRMWRS) at strategic locations within Mississippi River flood control systems under its responsibility. In this effort, DNR has funded a pilot implementation of this project through a collaborative effort of multiple organizations including Geocomp Corporation, PB Americas, Shannon & Wilson, James Lee Witt Associates, NIMSAT Institute at the University of Louisiana at Lafayette, SMARTEC and TIE Technologies.

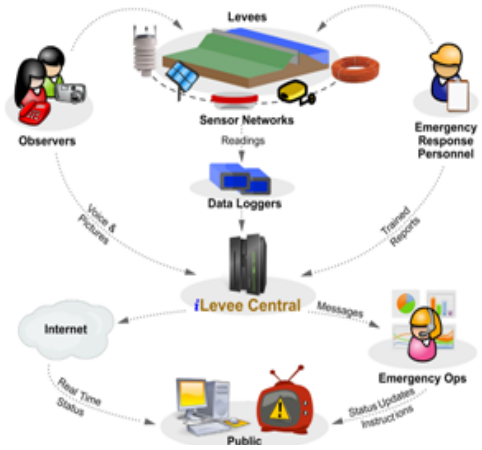


Figure 1. Concept of the iLevee System (Taken from the Proposal)

Project Description

iLevee collects data from monitoring sensors installed throughout the flood control system, Web or Mobile phone based responses from observers in the form of images, voice and text data and processes them in real time to display the health and status of the flood control system. This data is processed in real-time by decision support tools that are hosted on iLeveeCentral to assess the health of the levee and reports the status of levee health to first responders. The iLeveeCentral is the backbone of the iLevee system that consists of various hardware and software to receive and store incoming data streams through the internet, a probabilistic decision support system that runs on LONI and a GIS system that runs on a server to track and display the location of each source of data. In order to make the system highly available and avoid single points of failure, certain components of the system will be deployed at UCSD.

Time for LI Computational Scientist

The primary objective of this project is to develop and deploy an iLevee Central system that runs on LONI and UCSD. Dr. Gottumukkala was involved in writing the proposal and we request FTE of 6 months of the LONI Computational Scientist, Dr. Raju Gottumukkala's time for the next one year for the iLevee project. Dr. Gottumukkala will work with various faculty and staff from NIMSAT Institute and LONI Institute to design a highly-available iLevee Central system that will be deployed on LONI and parallelize various data processing and computation modules that will improve the response time of the iLevee Central.

Benefit to LONI

This project will be a collaborative effort across multiple universities and industry and an opportunity for the state's investements on LONI to be utilized for the state's emergency management efforts.

High Performance Computational Biology and Material Science Projects at Southern University HPC-BMSL Lab (FY: 2010~2011)

PI: Dr. Ebrahim S. Khosravi (Chair of Computer Science, SU)
Co-PI: Seung-Jong Park (Computer Science, LSU)
Co-PI: Marcia Newcomer (Biology, LSU)
Co-PI: Shengmin Guo (Mechanical Engineering, LSU)
Co-PI: Dr. Shuju Bai (Computer Science, SU)
Co-PI: Dr. Shizhong Yang (Computer Science and LONI CS)

The High Performance Computational Biology and Material Science Lab (HPC-BMSL) at Computer Science Department of Southern University proposes to perform 4 sub-projects in Fiscal year 2010(Mar. 2010 ~ Feb. 2011). The objectives of the projects are, (a). to setup and test a CRON high speed optical fiber network testbed at SU; (b) to perform ligand docking and QM/MM simulation of AA/8R-LOX, to understand and predict the electronic, optical, magnetic, and structural properties of the selected novel electronic materials; (c) to provide an infra-structured platform for systematically mentoring and training of under-graduate, graduate students, and post-doctors at Southern University and A & M College.

Dr. E. Khosravi will supervise all of the four proposed projects. The four sub-projects will synergistically address complementary tasks to dramatically enhancing our fundamental knowledge and practical applications in the biochemistry, drug design, and nano-size material science. The four research subprojects and proposed time are:
(1). The CRON high speed optical network testbed setup and testing by Dr. E. Khosravi, S. Park and S. Yang; (3 month).
(2). AA/8R-LOX docking and QM/MM simulation by Dr. Shuju Bai, Marcia Newcomer, E. Khosravi, and Shizhong Yang; (3 month).
(3). Ta doped ZrO_2 Thermal Barrier Coating (TBC) MD simulation by Dr. Shengmin Guo, E. Khosravi, and S. Yang; (3 month)
(4). Graduate, undergraduate students, and faculty HPC related research training. (1 month)

Project 1: The CRON high speed optical network testbed setup and testing
Dr. Khosravi and Dr. Park are funded by a NSF project CRON throught LSU CCT. Dr. Yang will be responsible for the high speed optical fiber network setup, including purchasing the workstation (Sun Microsystems), fiber, and network PCI card, setup the workstation and fiber network connection through the P5 LONI machine's optical port, test the connection and communication. Dr. Yang will also assist in the code developing and student training.

Project 2: AA/8R-LOX docking and QM/MM simulation
Dr. Shuju Bai and Dr. Newcomer will be the sub-project PIs. In this project Dr. Yang will be working on the AA/8R-LOX docking and QM/MM simulation. The ICM-Pro and Q-Chem will be utilized to simulate the docking site and chemical active site of the protein. The results will provide basic information for the drug design.

Project 3: Ta doped ZrO_2 Thermal Barrier Coating (TBC) MD simulation
Dr. Shengmin Guo, Shizhong Yang, E. Khosravi, and graduate students will perform ab initio MD simulation of Ta doped Yttrium (1:1) doped ZrO_2 (YSZ). This will extend current NASA-

EPSCoR supported TBC project and to enhance the closely collaboration with NASA Glenn Research Center where the Ta doped YSZ experiments are ongoing. All the simulations will be performed on LONI machines.

Project 4: Graduate, undergraduate students, and faculty HPC related research training
Southern University and A & M College is a traditionally large HBCU institution. Dr. Yang and LONI staff at SU will host an annual LONI HPC training session. Minority student and faculty will be trained by intimately engaging them in the training activities. Under Dr. Yang's funding supports, he will also support two graduate students to do the above proposed projects.

The PI, Dr. Khosravi, Chair of the Computer Science Department and LONI SU Co-PI, is currently funded by Navy, Raytheon, NSF, BoR, NIH, and NGA. Three Co-PIs, Drs. Park, Newcomer, and Guo from LSU are faculties who are experts at Computer Science, Biology, and Mechanical Engineering fields respectively. Support to the proposed LONI projects would generate new opportunities to attract more talented faculties, post-doctors, and students, secure more federal and industry funds, which without doubt fits into SU and LONI's long term development strategy.

Simulating Larval Dispersal in the Northern Gulf of Mexico

PIs: Caz Taylor (Assistant Professor, CCS & EEB, Tulane University), Richard Condrey (Associate Professor, OCS & ACE, Louisiana State University), Woody Nero (Oceanographer, NOAA National Marine Fisheries Service & Northern Gulf Institute), Erin Grey (Postdoctoral Researcher, EEB, Tulane University), Carey Gelpi (PhD Student, OCS & ACE, Louisiana State University)

Background: Commercial fishing in the Gulf of Mexico generates more than $900 million per year, representing approximately 25% of fishing revenue in the U.S. [1]. Many of the organisms fished, such as crabs and shrimps, disperse as small larvae that drift in near-shore currents. Understanding this dispersal phase is crucial for developing predictive population models. Yet such understanding has proven difficult because near-shore currents in the Gulf are influenced heavily by winds and tides [2] and because larvae often exhibit vertical swimming behaviors that can significantly alter their horizontal trajectories [3].

Project Proposal: Recently, the Naval Research Lab has developed the Northern Gulf of Mexico Ocean Nowcast/Forecast System (NGOMNFS) which incorporates wind and tide data to resolve near-shore circulation patterns at a 1.9 km, hourly scale that is sufficient for tracking larvae [4]. We proposed to simulate larval dispersal in the northern Gulf of Mexico by running a larval-behavior model through the archived NGOMNFS database (~1TB). The behavior model was written by Woody Nero at NOAA in Matlab and was parameterized for brown shrimp larvae. The current model is slow, taking 1 hour to simulate one night's release and dispersal of 500 larvae and 15 days to evaluate one year. We will develop more general and efficient software that can be easily modified to simulate the dispersal of species with different larval behaviors, and we will parallelize the model to run on the LONI cluster to speed up simulations. Postdoctoral researcher Erin Grey will use this resource to estimate blue crab dispersal for a dynamic population model for this species, Dr. Richard Condrey and Carey Gelpi will use it to look at blue crab dispersal from offshore shoals, and Woody Nero will use it to continue his research on brown shrimp dispersal. Additionally, the software will be made freely available through LONI and the Northern Gulf Institute, which is a consortium of universities led by Mississippi State University that seeks to integrate coastal science research.

Effort Requested: We request 1 month of full-time effort from a LONI Institute computational scientist who could assist Erin Grey and Carey Gelpi in compiling the larval-tracking code and then parallelizing it to run on the LONI cluster. Dr. Hideki Fujioka has already expressed willingness to work on this project, and his proximity to Dr. Grey at Tulane University would be convenient. We would, however, appreciate help from any of LONI's computational scientists.

Benefit to LONI Institute: This project represents collaboration between Tulane University, Louisiana State University, the NOAA and the Northern Gulf Institute, and would thus help LONI achieve its goal of fostering research collaborations. Furthermore, the end-result of this will be efficient, freely-available larval-tracking software that will encourage a variety of coastal scientists to utilize LONI's computational resources. Given the importance of larval dispersal to the majority of estuarine species, this LONI-based product will also greatly contribute to fisheries and ecosystem management in the Gulf of Mexico.

References

1. O'Bannon, B.K. (2001). Fisheries of the United States 2000. National Marine Fisheries Service, Office of Science and Technology, Fisheries Statistics and Economics Division. U.S. Department of Commerce. Silver Springs, MD.

2. Largier, J.L. (2003). Considerations in estimating larval dispersal distances from oceanographic data. Ecological Applications 13:S71-S89

3. Young, C.M. (1995). Behavior and locomotion during the dispersal phase of larval life. in: McEdward, L. editor. Ecology of marine invertebrate larvae. Boca Raton: CRC Press. p 249-277

4. http://www7320.nrlssc.navy.mil/IASNFS_WWW/NGOMNFS_WWW/NGOMNFS.html

Computational Infrastructure for Genome-wide Change Point Analysis at Basewise Resolution

Dongxiao □hu (UNO), □hiyu □hao (UNO), □un □hang (□avier) and Erik Flemington (Tulane)

Our goal is to develop a computational infrastructure for performing genome-wide change point analysis with base-wise resolution. Our experimental system is the detection of miRNA targets on the genome scale. miRNA is a recently discovered gene regulation mechanism that down-regulates gene transcription and translation by Watson-Crick paring with target sites. More than 80□ of human diseases have been shown to be intimately related to miRNAs. miRNA target prediction on the base-wise resolution has become a promising way to discover disease genes.

Using the very high end technology, next generation sequencing, we have identified functional target genes of the human microRNA, miR-155, a highly implicated oncogenic microRNA in a number of immune cell cancers. A genome-scale analysis on the basewise resolution will be needed to discover more novel miRNA targets. Algorithmically, we first align the short reads characterizing the transcriptome to the reference genome sequence, and then calculate genomic coverage. The targets are predicted by sudden drop outs of the coverage on the 3'-UnTranslated Region (UTR), called by low values of Schwarz Information Criteria (SIC) from a Change Point Analysis. The computational complexity of SIC based change point analysis at basewise resolution is $O(N^2 \log N)$□therefore, for $N = 3.2 \times 10^9$ potential change points in human genome, the number of calculations needed is easily to reach 10^{20} which takes hundreds of billions of seconds on a CPU with a few giga flops of computation power.

We propose to develop a computational infrastructure to run the change point analysis at basewise resolution for the human genome. The infrastructure is parallel computing in nature, since the problem size makes it infeasible on any single workstation. We request □ months of full time effort of Dr. □hao to develop a parallel algorithm for the proposed change point analysis and implement it on the LONI clusters. We have a multidisciplinary team of scientists involved in the project. Drs. □hao and □hu will collaborate on the algorithmic and computational side of this project to run the genome-wide analysis on the LONI clusters while minimizing the use for CPU time. Dr. Flemington will provide exert opinions and insights from biological domains to interpret the findings. Dr. □hang will also collaborate on the computational side and assist the outreach to researchers and educators at □avier University, a Historically Black Colleges and Universities (HBCU).

This project will benefit LONI by fostering collaborative efforts from multiple Louisiana institutions in their efforts in multidisciplinary cancer research. The proposed change point analysis on the basewise resolution is quite general, and widely applicable to diverse biological systems. For example, detection of chromosome copy number changes, *de novo* gene prediction etc. Therefore, the proposed computational infrastructure will benefit other researchers as well. This work also fits well with the goals of the state as a whole -- Louisiana is investigating significant resources in growing the biotechnology industry. Long-term, expansion in this area may interest the biotech/pharmaceutical industry and tie in with statewide emphasis on biotech.

Computational Characterization of Transcriptome on the Isoform Level

Dongxiao □hu (UNO), □hiyu □hao (UNO), □un □hang (□avier) and Erik Flemington (Tulane)

Our goal is to develop a computational infrastructure for automatic characterization of transcriptomes on the splicing isoform level. Due to inherent limitations of hybridization based expression microarray technology, it only allows for the characterization of transcriptomes at the gene level. However, in the real biological system, the ultimate effectors on the transcription level are the splicing isoforms. One gene can be alternatively spliced into different isoforms of transcripts in different tissues. Despite providing output on overall gene expression, isoform analysis using microarray data is computationally inaccessible.

The advent of a very high-end technology, next generation sequencing, has provided new opportunities to solve this problem. It provides more accurate quantitative data and it provides additional information that allows the characterization of a transcriptome at the isoform level. We have designed an iterative algorithm to detect splicing isoforms using short reads originating both from exons and exon-exon junctions. More formally, assume $i = (1,2 \ldots, \mathrm{N})$ is the short read (row) index, $j = (1,2 \ldots, \mathrm{M})$ is the isoform (column) index and $\not\Phi = (p_1, p_2, \ldots, p_M)$, where p_j is the mixture proportion for the isoform j. Initializing all the compatible p_j to be the same, and add up to 1, i.e. $\sum_{j=1}^{M} p_j I_j = 1$, where I_j is the indicator having value of 1 if the jth isoform is compatible, 0 otherwise. The proposed Expectation Maximization (EM) type algorithm is as follows□E-step□$z_{i,j}^{(k+1)} = \frac{y_{i,j} p_j^{(k)}}{\sum_{j=1}^{M} y_{i,j} p_j^k}, \forall i, j$. M-step□Let $n_j^{(k+1)} \square \sum_{i=1}^{N} z_{i,j}^{(k+1)}, \forall j$, $p_j^{(k+1)} \square \frac{n_j^{(k+1)}}{N}, \forall j$. In a proof-of-principle study, we have shown that our algorithm is able to accurately identify all splicing isoforms in mouse liver, brain and muscle.

The real challenge comes in the computational side in that the same iterative algorithm must be applied to a total of 22,000 genes annotated in human genome. The computational complexity of each gene is $O(M^2N)$, where M is typically 50 to a few hundreds, and theoretically N can be exponential to M. The problem size makes it necessary for us to develop an efficient parallel algorithm and run it on a powerful supercomputer such as those available on the LONI clusters. We request □ months of full time effort of Dr. □hao to develop a parallel algorithm for the proposed work and to implement it on the LONI clusters. We have a multidisciplinary team of scientists involved in the project. Drs. □hao and □hu will collaborate on the algorithmic and computational side of this project to run the transcriptome characterization on the LONI clusters while minimizing the use for CPU time. Dr. Flemington will provide exert opinions and insights from biological domains to interpret the finding. Dr. □hang will also collaborate on the computational side and assist the outreach to researchers and educators at □avier University, a Historically Black Colleges and Universities (HBCU).

This project will benefit LONI by fostering collaborative efforts from multiple Louisiana institutors in their efforts in multidisciplinary cancer research. The proposed computational transcriptome characterization on the isoform level is quite general and widely applicable to diverse tissues and biological conditions wherever the next generation sequencing data is available. Therefore, the proposed computational systems will benefit a wide range of researchers as well. This work

also fits well with the goals of the state as a whole -- Louisiana is investigating significant resources in growing the biotechnology industry. Long-term, expansion in this area may interest the biotech/pharmaceutical industry and tie in with statewide emphasis on biotech.

Surface plasmons in metal/semiconductor composites and devices

Dentcho A. Genov, PhD. (LI faculty, Louisiana Tech University)
Abdul Khaliq, MS. (LI CS, Louisiana Tech University)

***Background*:** The inhomogeneous metal/semiconductor composites are nanoscopic artificial materials that have unique geometrical and optical properties. Under electromagnetic wave illumination these complex materials manifest energy localization in very small spatial areas (a few nanometers) and giant enhancement of the local field intensities, which correspond to excitation of localized surface plasmon (SP) modes. At critical metal concentrations, the random films are inhomogeneous and self-similar (fractal) on any length-scale. Thus, for any incident wavelength resonating clusters exist in the composite. Such broad frequency response results in anomalous optical properties including extraordinary absorption and enhancement of nonlinear optical processes such as Surface Enhanced Raman Scattering (SERS), high order frequency generation, etc. The unique properties of the percolating films make them ideal not only for fundamental studies of light-matter interaction in disordered systems, but also for a wide range of applications in biological sensing and spectroscopy, fast optical devices, surface science, and condensed matter physics.

Proposed research: **1.** ***Numerical methods in nanoplasmonics*:** As part of this proposal we first seek to rewrite the existing FDFD codes in parallel and develop novel, highly efficient numerical methods for calculating the electromagnetic (EM) response of 2D and 3D inhomogeneous systems of metal/semiconductor nanoparticles. Additionally, we intend to use a 'memoization' method, an efficient way to do fast searches of conduction paths, providing a solution to the problem in only $O(N^{3/2})$, which is to be compared to $O(N^3)$ for the standard Gauss-Seidel method (N is the number of particles). Successful development of the numerical codes will make possible simulations on the LONI supercomputers of systems with up to 10^6 and 10^4 particles in the 2D and 3D cases, respectively. This will allow for a first time to study local and macroscopic response of real size systems and compare with existing experimental data. The developed numerical codes could be effectively applied to investigate large variety of strongly interactive, sub-wavelength ensembles of particles, including dense semiconductor quantum dots systems, periodic arrays with tunable optical properties, photonic nano-circuits and optical switches. **2.** ***SP transistor*:** Relying on the above surface plasmon related studies, we intend to develop a novel semiconductor based Surface Plasmon Transistor (SPT). The SPT promises to combine electronics with optics by excitation and active control of propagating surface plasmon modes through Si/GaAs *n-p-n* junction. The preliminary data suggest that this optoelectronic device can provide modulation bandwidth larger than 1THz, which can potentially open a new rout toward fast optoelectronics and computing. To study the device characteristics we will seek to integrate commercial software COMSOL, to perform a distributed memory, parallel parametric FDFD electromagnetic simulations on the LONI clusters.

Impact of the proposed research: The proposed research will lead to development of numerical and analytical tools for solving highly complicated problems of EM interaction with complex media. Those methods will answer standing fundamental questions concerning the nature of collective electronic excitations in metal-dielectric composites. Due to the inhomogeneous nature of the problem it is crucial that very large system sizes are investigated. Such systems cannot be studied with average computational facilities and utilizing the LONI recourses will allow to traverse new regimes of operation that have been a mystery for the last 50 years. Successful realization of the project will help to establish the LONI Institute as a top center for computational electromagnetism. Furthermore, this project will focus on the development of new type of surface plasmon transistor that is far superior compared to the conventional devices in terms of it potential bandwidth scalability. The work initiated in this proposal, is expected to serve as a basis to build on existing and establish new collaborations with theoreticians and experimental scientists within the six LONI institutions but also with other national universities including groups at UC Berkeley (Dr. X. Zhang), and Yale University (Dr. Hui Cao). The first part of this project has been included into a RCS proposal submitted on Nov. 4 to the Louisiana Board of Regents, while the SPT data will be used in a proposal to be submitted to the NSF-EPM program. The total workloads for the LI faculty and CS are 3 FTE-months per year, for total duration of the proposal of 1.5 years and expected supercomputer time allocation of 50K SUs. Also, the LI faculty will provide a PC workstation and a graduate student to work full time on the project, which will also be the subject of the student PhD thesis.

Date: 12/10/09

Improving Antibody Design by Structure Prediction, SCOP Classification and Protein - Protein Docking

Seth Pincus (spincus@chnola-research.org)
Research Institute for Children, Children's Hospital New Orleans & Louisiana State University Health Sciences Center

Zhiyu Zhao (sylvia@cs.uno.edu)
LONI Institute & Department of Computer Science, University of New Orleans

Bin Fu (binfu@cs.panam.edu)
Department of Computer Science, University of Texas – Pan American

We propose to use a structure similarity search tool that we have developed in conjunction with existing structure prediction and protein - protein docking tools to improve the structural design of an antibody. We will be able to test experimentally the computational prediction of improved design by altering the amino acid sequence and determining whether the resulting antibody has the expected immunological characteristics. Antibodies can be used for molecular measurement and to treat diseases. The efficacy of antibodies is related to binding affinity. The software we propose to develop will allow computational prediction of higher affinity antibodies; such tools are of potential value to biotechnology and pharmaceutical companies.

The computational improvement process will be automatic. It will take amino acid sequences of variable regions in an antibody as input, and output improved structures predicted to interact better with antigen. The improved antibodies, which belong to the same Structural Classification of Proteins (SCOP) domain, family, or superfamily, are structurally similar to the original one, but with specific amino acid changes. It is these differences that result in a better interaction with antigen. Computationally, our antibody improvement problem is described as: Given amino acid sequences of heavy chain variable region (V_H) and light chain variable region (V_L) in a monoclonal antibody, create *denovo* structures for the antibody complementarity determining regions (CDRs) by finding similar structures in a protein database, engrafting them onto the antibody backbone, and testing them, *in silico*, for docking with the antigen. The improved design of that antibody structure may result in a better binding with its antigen.

Instead of relying merely on sequence similarity search tools such as IgBLAST to find similar amino acid sequences in an antibody, we plan first to use a structure prediction tool to predict the structure of that antibody. Currently there are three antibody-modeling servers on the Internet: the Web Antibody Modeling (WAM), Prediction of Immunoglobulin Structure (PIGS), and RosettaAntibody (command line tool also available). We have requested the installation of Rosetta on Queen Bee. We will investigate the antibody modeling performance of these servers, by testing their predictions with antibodies whose 3D structures are known. We have obtained some preliminary results. See Figure 1 and Figure 2 for antibody modeling results of RosettaAntibody and PIGS based on protein 2DDQ (detailed interpretation of these results and other preliminary results are not presented due to page limitation and can be provided upon request). By comparing the predicted structure to the actual structure, particularly for antibodies similar in sequence to our test antibody, we will choose one as our antibody structure prediction tool. Then, since the predicted antibody structure is new to the Protein Data Bank (PDB) and its SCOP domain, family and superfamily are unknown, we will use the SCOP classification tool that we have developed to predict the class of that structure, and to retrieve from the PDB a set of similar antibody structures that belong to the same class i.e. the same SCOP domain, family or superfamily. Next, we will use a protein - protein docking tool to predict the structure of each antibody - antigen complex and observe changes in the antibody that would result in a better binding.

Figure 1: RosettaAntibody Model Rank 1 vs 2DDQ

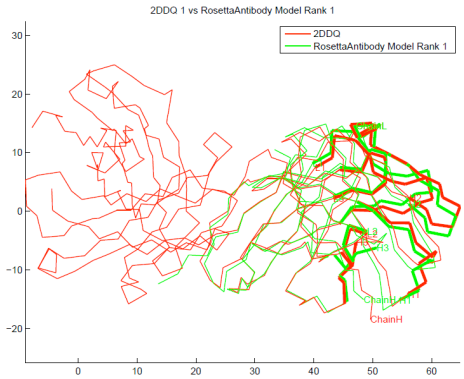
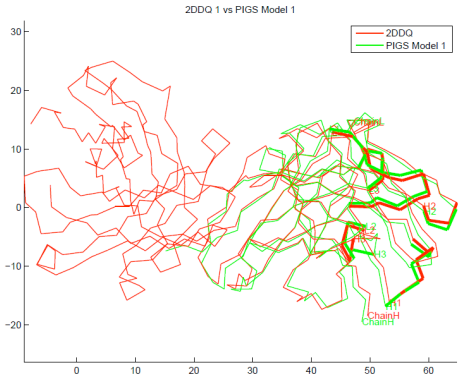


Figure 2: PIGS Model 1 vs 2DDQ



The antibody, RAC18, we will be working with is a well-studied molecule, which binds to the plant toxin ricin, a material of biodefense concern. This antibody has been shown to be of protective efficacy in animal models, even when administered hours after the toxin. We now wish to improve the affinity of this antibody. The amino acid sequences of the heavy and light chain variable regions are known, as are the amino acids that the antibody contacts on the ricin molecule. Once computational predictions have been made regarding specific alterations to improve binding to ricin, we will use genetic engineering technologies to make that antibody and test it to determine if we have improved the affinity of binding to ricin and its protective efficacy *in vivo*.

We request 6 months of full time effort of Dr. Zhiyu Zhao over a one year period to work out preliminary results using existing and homemade software tools. Dr. Zhao will then develop an antibody improvement software tool which automates the above mentioned improvement procedure. All the software development work will be performed on a LONI cluster such as Queen Bee. These specific experiments may significantly improve the function of a clinically important antibody. By addressing this theoretically important issue in a well-characterized experimental system these studies may represent the initial steps in the production of more general tools for enhancing antibody and protein design. The successful prediction of protein improvements by a computational approach would be an important accomplishment in the field of bioinformatics.

In the last a few years, we have successfully developed efficient software for alignment and similarity searches of protein structures with structures in the Protein Data Bank. The proposed project is a continuation of our ongoing development in this field. The next phase of this work is to apply the computational tools to address practical biological problems, such as this one from Dr. Pincus's group. It will also strengthen our ongoing collaboration with Dr. Bin Fu's group, which is focusing on the protein folding problem at the University of Texas-Pan American.

Computational Model of Pulmonary Small Airway Interdependence

Hideki Fujioka[1], David Halpern[2] and Donald P. Gaver[3]
[1]Center for Computational Science, Tulane University
[2]Department of Mathematics, University of Alabama, Tuscaloosa
[3]Biomedical Engineering Department, Tulane University

Project Description

Pulmonary airways are surrounded by parenchyma that consists of numerous alveoli, all of which are connected to distal airways. Therefore, the dynamics of each airway and alveolus is interdependent. As such, the behavior of one component may affect all others through parenchymal tethering.

Pulmonary epithelial cells are exposed to mechanical stresses due to the stretch of the surrounding substrate and the motion of thin liquid film over them. There are a number of situations that these stresses becomes excessive, leading to cell damage or death. For example, local reduction of surfactant concentration results in nonuniform deformation of the parenchyma and the airways, and causes high membrane strain and high fluid stress due to high surface tension. Surface tension induces liquid flows, which may cause the lung's airways to close due to the formation of a liquid plug as a result of drainage of the liquid lining coating the airways or collapse of airway due to low pressure in the liquid. Once occluded either by a short plug or an extended collapsed region, the airway must be reopened to maintain ventilation to distal regions of the lung.

Though simplified models of airway closure and reopening phenomena in single airways have led to advances in our understanding of atelectrauma, the effect of the surrounding alveoli and the interdependence of airways on the these phenomena needs to be studied.In this project, we plan to investigate the effect of parenchymal tethering on inflation/deflation mechanisms of atelectic regions of the lung. This may yield improved technologies for reducing the morbidity and mortality associated with respiratory distress syndrome and ventilator-induced lung injury.

Employing our previous method to solve the fluid mechanics[1,2], we modify the model to add parenchymal effect. A truncated-octahedron alveolus computational model[3] is employed. The displacement based finite element method is used to analyze large deformation of the alveoli surrounding airway models. The code is parallelized using MPI and PetSc library. We would like to request 6 months of an FTE of LONI Institute Computational Scientist who is experienced in pulmonary mechanics and computational modeling. This research will enhance collaboration between the Center of Computational Science and faculty members and students in the department of Biomedical Engineering at Tulane.

Benefit to LONI Institute

This project would develop a expertise in the multi-scale modeling of pulmonary system and show LI's capabilities of large scale computational modeling for fluid and solid mechanics.

References

[1] Halpern D., Naire S., Jensen O.E., Gaver D.P., (2005), J. Fluid Mech. Vol. 528, pp.53
[2] Fujioka H., and Grotberg, J.B., (2005), Phys. Fluids, Vol.17(8), 082102
[3] Dale, P.J., Matthews, F.L., and Schroter, R.C., (1980), J. Biomechanics, Vol.13. pp.865.

Project: Thermal Modeling and Thermo-mechanical Modeling of Thermal Barrier Coatings (TBCs) using Ansys Fluent Commercial Package

Stephen Akwaboa
Department of Mechanical Engineering
Southern University and A&M College
Baton Rouge, LA 70813
stephenakwaboa@engr.subr.edu

Thermal barrier coatings (TBCs) are used in gas turbine engines to achieve higher turbine inlet temperatures (TITs), improve turbine operating temperatures, reduce fuel consumption, increase components lives and thus lead to better turbine efficiency. Yttria-stabilized zirconia (YSZ), is an ideal candidate for TBCs as it has good thermal shock resistance, high thermal stability, low density, and low thermal conductivity. Traditionally, there are two main methods of fabricating TBCs: air plasma spray (APS) TBCs and electron beam physical vapor deposition (EBPVD) TBCs. The ability of TBCs to offer thermal protections to metallic substrates, to a large extent, depends on their thermophysical properties (thermal diffusivity, thermal conductivity and specific heat capacity) as well as their porosity. We have purchased laser flashline equipment (FL 5000) from Anter Corporation to characterize the thermophysical properties at high temperatures (up to 1300 ^{o}C). The porosity of the samples are measured using Poremaster, a mercury porosimetry equipment manufactured by Quantachrome Incorporated. Operating modern turbines at high temperatures is crucial to most turbine researchers since it is the main potential source of improving engine efficiency. In order to avoid overheating of the metallic components and also to prevent corrosion and oxidation at elevated temperatures, thermal barrier coatings (TBCs) are the ultimate choice. By using TBCs, turbine inlet temperature can be increased by 200°C. It is also known that the use of TBCs promotes enhanced component life, reduces fuel cost, and combustion gases emitted into the atmosphere.

Our ongoing research involves the use of TBCs on superalloy IN 738 substrate materials for elevated temperature gas turbine applications. In principle the focus of our research involve thermophysical properties characterization of TBCs, thermo-mechanical analyses of the TBCs and micro-scale thermal modeling of in-service performance studies of the TBCs. By using the thermophysical properties as inputs to the thermal modeling in Ansys Fluent CFD Commercial package, the temperature distribution in the metal substrate and the composite TBCs are determined. The Fluent solution is exported into Ansys to perform the thermomechanical anaylsis utilizing the temperature distribution as thermal loads. Analysis in Ansys gives the thermal stress distribution in the metal substrate as well as the in-plane stresses at the various interfaces of the TBC system (top coat/TGO, TGO/bond coat and bond coat/ substrate interfaces) which are crucial in identifying high stress areas in the TBC system.

Even though we have purchased Ansys Fluent products which are used to do most of our thermal and structural modeling work, we are constrained by the number of nodes we can use for our analysis. Ansys , for instance, allows only 256 kilo-nodes for its structural analysis package. This cap on nodes greatly compromises the quality of work that we can produce and publish. We need high resolution models in Fluent and Ansys to be able to obtain good solutions with high reliability. This means we need computers with superior computing power, hence my request for this allocation so good quality work can be produced for publications in a timely manner. We request the services of a Computational Scientist in Spring of 2010 (50 hrs) to help us to run our code on the supercomputers for quick solutions.

Large-Scale First Principles Computation and Simulation of Catalytic Properties of Nitrogen Doped Carbon Nanotubes for Dioxygen Reduction
(A LONI Proposal for FY 2010 - 2011)

PI: Dr. Guang-Lin Zhao, Professor, Department of Physics, Southern University and A & M College, Baton Rouge, Louisiana 70813. (Phone) 225-771-4491; (fax) 225-771-2310; (email) Guang-Lin_Zhao@subr.edu.
Co-PI: Dr. Ebrahim Khosravi, Professor, Computer Science Department, Southern University and A & M College, Baton Rouge, Louisiana 70813
Co-PI: Dr. Shizhong Yang, Computational Scientist and Assistant Professor, Computer Science Department, Southern University and A & M College and LONI CS, Baton Rouge, Louisiana 70813

Proposal Description

We propose to perform first principles density functional calculations for the catalytic properties of nitrogen doped carbon nanotubes (CNTs) for dioxygen reduction, by requesting **two month**s of Dr. Shizhong Yang's research time in fiscal year 2010-2011 (Mar. 2010 - Feb. 2011). We aim to understand, (i) the stable structure of nitrogen doped CNTs; and (ii) the electronic and catalytic properties of the N-doped CNTs for dioxygen adsorption and reduction.

Precious platinum (Pt) catalyst is a key ingredient in fuel cells, which produce electricity and water as the only byproduct from hydrogen fuel.[1] However, platinum is rare and expensive. Reducing the amount of Pt loading by identifying new catalysts is one of the major targets in the current research for the large-scale commercialization of fuel cells. Specifically, developing alternative catalysts to substitute platinum for the oxygen reduction reaction (ORR) in the fuel cell cathodes is essential, because the slow kinetics of this reaction causes significant efficiency losses in the fuel cells. Recent intensive research efforts in reducing or replacing Pt-based electrode in fuel cells have led to the development of new ORR electrocatalysts, including carbon nanotube–supported metal particles.[2,3]

In 2006, Ozkan and coworkers reported that nitrogen-containing nanostructured carbons and nanotubes have promising catalytic activity towards ORR.[4, 5] In a 2008 report, Yang *et al.* at Argonne Laboratory showed that the vertically-aligned carbon nanotube (CNT) arrays, which are functionalized through nitrogen and iron doping by a chemical vapor deposition (CVD) process, can be electrocatalytically active toward ORR.[6] They further identified FeN_4 sites, which are incorporated into the grapheme layers of aligned carbon nanotubes, being electrocatalytic active.

In a 2009 publication in *Science*, Gong *et al*.[7] reported that vertically aligned nitrogen-containing carbon nanotubes (VA-NCNTs) can act as a metal-free electrode with a much better electrocatalytic activity, long-term operation stability, and tolerance to crossover effect than platinum for oxygen reduction in alkaline fuel cells. The functionalized CNTs show promise properties as an alternative non-Pt electrocatalyst with a unique nano-architecture and advantageous material properties for the cathode of polymer electrolyte membrane fuel cell (PEMFC). They also performed hybrid density functional theory (DFT) calculations for the hydrogen edge-saturated (5, 5) CNT, in which a nitrogen atom doped in the middle of the nanotube. However, according to our recent *ab initio* simulation, nitrogen atoms prefer to stay at the open-edge sites of single wall (10, 0) CNT.[8]

In order to understand the fundamental mechanism of the catalytic properties of the N doped CNTs for O_2 reduction, we need to perform chemical reaction path simulations. The PI, Khosravi, and Yang will perform the reaction barrier simulations using the Vienna Ab-initio Simulation Package (VASP), Q-Chem package (which is a recent developed quantum chemistry fast software package), and some supplemental data processing codes developed at SU HPC group. We had tested the exchange-correlation interaction potentials of the many electron system both in local density approximation (LDA) and in the generalized gradient approximation (GGA) with the same model and same parameters and found that they give the consistent results in the stability studies of nitrogen doped CNTs. We will develop an efficient and reliable model for structure optimizations. The calculated results from the two software packages will be compared carefully for verification of the calculated catalytic properties of nitrogen doped CNTs for dioxygen reduction reaction. We expect that the reduced-reaction-barrier quantity for dioxygen reduction reaction on nitrogen doped CNTs can be calculated. All the simulation will be performed on LONI machines.

The proposed project will lead a fundamental understanding of the novel non-precious-metal catalysts. One graduate student will be involved in the project. The success of the project will also increase the future success in acquiring DOE funding support.

Corresponding PI's email address: Guang-Lin_Zhao@subr.edu

Co-PI's email addresses: Ebrahim_Khosravi@subr.edu, Shizhong_yang@subr.edu

References:

[1]Subramani, V.; Gangwal, S. K. *Energy & Fuels* **2008**, *22*, 814..

[2]Che, G.; Lakshmi, B. B.; Fisher, E. R.; Martin, C. R. *Nature* **1998**, *393*, 346.

[3]Kongkanand, A.; Kuwabata, S.; Girishkumar, G.; Kamat, P. *Langmuir* **2006**, *22*, 2392.

[4]Matter, P. H.; Ozkan, U. S. *Catal. Lett.* **2006**, *109*, 115.

[5]Matter, P. H.; Zhang, L.; Ozkan, U. S. *J. Catal*. **2006**, *239*, 83.

[6]Yang, J.; Liu, D. J.; Kariuki, N. N.; Chen, L. X. *Chem. Commun*. **2008**, *3*, 329.

[7] Gong, K.; Du, F.; Xia, Z.; Durstock, M.; Dai, L. *Science* **2009**, *323*, 760.

[8] S. Yang, G.L. Zhao, and E. Khosravi, *J. Phys. Chem.*, submitted for publication.

Leveraging LONI Workflow Developments onto TERAGRID To Enable High Performance High Throughput Molecular Dynamics Simulations.

PI: **Thomas C. Bishop (CCS, Tulane University)**

bishop@tulane.edu

504-862-3370

Senior Investigators: Joohyun Kim and Nayong(CCT, LSU)

Requested: LI Computational Scientists (CSs) Time: 6 months FTE-months

Project Description:

In an NIH funded study (R01GM076356) Bishop's goal is to investigate sequence dependent variations in nucleosome stability using molecular dynamics simulation techniques. Nucleosomes are the fundamental structural unit of chromatin and can be formed from any 146 basepair segment of DNA. In theory there are some 4^{146} possible sequences of nucleosomal DNA. In practice the ~12 million basepairs in the yeast genome position only about 60,000 nucleosomes (Jiang, et al., 2009). The degree of positioning ranges from highly positioned to nearly random positioning. Unfortunately, all available x-ray structures of the nucleosome have utilized nearly the same 146bp sequence of DNA (Luger, et al., 1997). We therefore have limited information on sequence induced structural variations in the nucleosome superhelix.

Progress to Date:

In order to investigate sequence dependencies, we implemented with the support of LI staff, a high-performance high-throughput molecular dynamics work flow on LONI (Bishop, 2010). Our workflow strategy was inspired by NAMD-G (Gower, et al., 2006), but was structured so as to take advantage of Louisiana Optical Network Initiative (LONI) resources. We used 4 LONI machines and Petashare (Wang, et al., 2009) to accomplish these simulations by farming them out in groups of 4 to Oliver, Louie, Poseidon, and Painter. Each simulation utilized 64cores (16nodes) and all tasks were run in 1ns increments (approx. 16hrs run time per ns of simulation). In total there were 16*16 = 256 tasks to be completed. Each task began by fetching via Petashare's pget command the necessary initial coordinates, velocities and parameter files from Petashare and ended by depositing the simulation's outputs (extended system file, final coordinates, velocities, trajectory and velocity files) into Petashare via the pput command. During run time trajectory data was copied on a daily basis from Petashare to Bishop's local computing resources for purpose of analysis and visualization. This avoided a lengthy post-production trajectory transfer. In this manner we were able to complete all simulation tasks during two separate 10 day periods. The first period produced 8ns for each of the 16 systems, and the second period produced an additional 8ns. This workflow enabled us to efficiently utilize nearly 300,000SUs at 4 different LONI sites (sometimes concurrently), and manage approximately 1.5Tb of data via PetaShare. However it still required significant user intervention and manual monitoring of progress.

Need for LI Support:

Based on results from this LONI study of 16 nucleosomes, we were recently awarded a TERAGRID allocation (8,000,000SU) to conduct a larger study that is expect to produce some 40Tb of data. The TERAGRID effort requires that we port our LONI workflow to the TERAGRID environment and that we further automate the workflow. Our workflow makes use of LONI's PetaShare resources and requires advanced scheduling capabilities. Both Petashare and the scheduling tools function differently in the LONI environment than in the TERAGRID environment. Thus developing a workflow solution that works optimally in both the LONI and TERAGRID environments requires collaboration between experts with knowledge and mastery of each.
From a broader perspective developing these tools provides an important/critical bridge between LONI and TERAGRID usage scenarios. The overall goal is to leverage LONI developments efforts for achieving high-performance high throughput molecular dynamics simulations onto the TERAGRID in such a away that we and other users can realize a lower time-to-solution by effectively managing distributed resources where ever they may reside. Molecular dynamics studies are historically the largest consumer of supercomputer resources thus any tools developed for this project will have a significant impact on both LONI and TERAGRID users.

Effort Requested and Involvement of Computational Scientists:

We estimate this will require 6 months of FTE.

Benefit to LONI Institute:

Advanced User Support was requested in Bishop's TERAGRID allocation request. The award received the highest recommendation for AUS possible. Thus LI support for this project provides a unique opportunity for LI and AUS personnel to collaboration in leveraging technology developed on LONI onto the TERAGRID and thereby demonstrate proof-of-concept in the LONI-TERAGRID development cycle.

References Cited:

- **Bishop T. C.** The Nucleosome Energy Landscape as Revealed by All Atom Molecular Dynamics [Journal] // J. Biomolecular Structure and Dynamics. - 2010. - Vol. Submitted.

- **Gower Michelle [et al.]** Managing Biomolecular Simulations in a Grid Environment with NAMD-G [Journal] // TCBG Publication. - 2006. - pp. 1-7.

- **Jiang Cizhong and Pugh B. Franklin** A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome. [Journal] // Genome Biol. - 2009. - Vol. 10. - p. R109.

- **Luger K. [et al.]** Crystal structure of the nucleosome core particle at 2.8 A resolution. [Journal] // Nature. - 1997. - Vol. 389. - pp. 251-260.

- **Wang X. [et al.]** Semantic Enabled Metadata Management in PetaShare [Journal] // International Journal of Grid and Utility Computing. - 2009. - Vol. 1. - pp. 275-286.