LONI Institute Project Proposal – Spring 2011:

Large-Scale Combustion Simulations with Detailed Reaction Chemistry

PI: Ingmar M. Schoegl, Assistant Professor – ✉ ischoegl@lsu.edu / ☎ (225) 578-4332
Mechanical Engineering, Louisiana State University, Baton Rouge

Background

Numerical simulations of reacting flows with detailed chemistry are an important field of research, as they reveal details of chemical processes that cannot be assessed with experimental methods. Detailed combustion simulations have practical value as they illuminate the formation of environmental pollutants, e.g. soot and its precursors. Also, they yield valuable insights for the efficient conversion of hydrocarbon fuels in power generation, which is an important sector of Louisiana's economy. Combustion simulations with detailed chemistry are computationally challenging, as modern reaction mechanisms involve hundreds of molecular species participating in thousands of chemical reactions. In computational fluid dynamics (CFD), concentrations of individual species add to the parameter space, producing large amounts of data even for relatively small computational grids. Furthermore, combustion occurs in thin reaction layers, which requires suitable grid refinement strategies.

In earlier numerical studies, the PI employed serial codes that typically required several weeks of computational time although both domains and reaction mechanisms were relatively small (1-D transient code with 300 grid points and 107 species;[1] 2-D transient CFD code with 300x12 grid points and 49 species[2]). Results of these studies, however, revealed that the number of species/reactions was insufficient to produce acceptable numerical data at extremely fuel-rich conditions. The expected computational effort for larger mechanisms clearly indicates that a parallel computational approach is necessary.

Goals

The proposed project involves the parallelization of a serial combustion code that is part of Cantera, an open-source chemistry package.[3] The PI has in-depth knowledge of the existing C++ code base; LI expertise would help build the framework for large-scale parallel combustion simulations.

The scope of the project is to investigate laminar, premixed combustion on a one-dimensional domain using large reaction mechanisms (>100 species and >1000 reactions *in each grid point*). Numerical results will enhance an ongoing Louisiana BoRSF RCS grant (LEQSF(2010-2013)-RD-A-04; "Tomographic imaging of combustion zones"; $139,365 over three years) that studies combustion at fuel-rich conditions. Once preliminary numerical results become available, federal funding will be sought for model extensions to 2-D and 3-D domains.

Effort Requested of LI Computational Scientist

Roughly 30 hours of effort per month is requested of a LI computational scientist for a year (FTE of two months). The LI CS would help a PhD student (Mohsen Ayoobi) identify the best approaches for building a parallel framework for large scale combustion simulations that is suited for 1D, 2D and 3D domains.

Benefits to the LONI institute

The project would increase LI's national visibility and prove its capabilities of large scale combustion simulations. The significant computational effort required for multi-dimensional simulations with large reaction mechanisms has only recently become manageable, and there is little published material in the literature. Furthermore, the proposed project may be suited for integration with the Cactus-based CFD toolkit within LONI's cybertool environment.

References

[1] M. J. Dixon, I. Schoegl, C. B. Hull, J. L. Ellzey, Experimental and numerical conversion of liquid heptane to syngas through combustion in porous media, Combust. Flame 154 (2008) 217–231.

[2] I. Schoegl, J. L. Ellzey, Numerical investigation of ultra-rich combustion in heat exchangers, Combust. Sci. Tech. 182 (2010) 1413–1428.

[3] D. G. Goodwin, An open-source, extensible software suite for cvd process simulation, in: Proc. of CVD XVI and EuroCVD Fourteen, Electrochem. Soc., 2003, pp. 155–162.

Thermodynamics and Kinetics in H_2 Storage Systems

Weizhong Dai, Ph.D. (Professor, Louisiana Tech University)
Abdul Khaliq, Ph.D. (LI CS, Louisiana Tech University)

Background: Among the barriers that hinder the use of hydrogen as a clean alternative to hydrocarbon fuels are absorption/desorption rates, volume/weight ratios, hydride stability, and desorption temperatures of current H_2 storage materials. We will use novel multiscale MD and MC simulations, kinetic MC, finite element, and finite difference modeling to predict rates of hydrogen uptake/release over time scales reaching 10^3 s and extending over use X-ray tomography to probe these materials over the same length and distance scales. The goals of this focus area are to predict the influence of catalytic additives in enhancing mobilities and desorption rates in metal hydrides and to explore a wider range of potential hydrogen storage materials. The payoff will be an improved ability to design materials of hydrogen storage.

Proposed Research: We will study diffusion and reactions using kinetic MC and employ finite element and finite difference methods for extending length and time scales.

- Employ the LSU X-ray synchrotron tomography beamline and the Argonne Advanced Photon Source nanotomography beamline to study in situ and ex situ de/rehydrogenation reactions, to assess 3D microstructures and interphase boundaries, and to quantify the distribution and size of single phase domains.
- CTCI visualization toolkits will be utilized since atomic diffusion rates may limit the desorption, Browne (LSU) will use Fick's Law diffusion and nonequilibrium chemical models to model solid state diffusion and phase transitions.
- The temperature-dependent diffusion coefficients will be determined using diffusion and diffusion-reactions equations and experiments.

Impact of the Proposed Research:

The results of these simulations, calculations, and experiments will be the prediction of optimum conditions and materials for hydrogen release and uptake.

Computational Scientist Time Requested:

12 Months

This is a funded project of Louisiana Alliance for Simulation-Guided Materials Applications (LA- SiGMA).

Reaching the terahertz range with optoelectronics: surface plasmon transistor

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

Background: The proposed research aims to develop novel optical transistor based on Surface Plasmon (SP). The preliminary data suggest that this optoelectronic device can provide modulation bandwidth up to 1THz, which can potentially open a new direction toward fast optoelectronics and computing.

Research Objectives: In this proposal we propose to study numerically the optoelectronic response of 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 a Si/GaAs *n-p-n* junction. To study the device characteristics we will perform a distributed memory, parallel parametric 3D (three dimensional) finite difference frequency domain (FDFD) electromagnetic simulations on the LONI clusters. Our research objectives rely and expand on the information already gained in the 2D case as part of the research performed by the PI's for a previously supported by the LI proposal (Title: *Surface plasmons in metal/semiconductor composites and device*s, 2010).

Research Accomplished: The electron generation in a three layered Si/GaAs device was analyzed with the state of the art semiconductor device simulator "MEDIC". In the device a p-type Si/GaAs is sandwiched between two layers of highly doped n-type semiconductor. Under forward bias the electron concentration within the p-type semiconductor is increased to match the electron concentration in the adjacent n-type semiconductor, which allows transmission of the SP modes through the device. The obtained electron profile (MEDIC) was incorporated in a 2D FDFD code used to calculate the Surface Plasmon propagation throughout the entire device (see Fig.1a). The power flux calculated vs. applied voltage (Fig. 1b) shows clear optical switching with response time surpassing 200GHz.

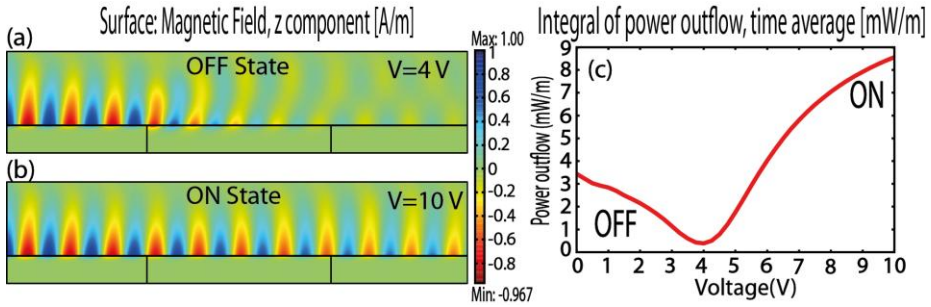


Figure 1. Optoelectronic transistor: (a) *on* and *off* states for a surface plasmon mode propagating on the device surface are established for (b) forward bias of 4V and >10volts, respectively.

Future Work: Here, we will seek to expand our initial work by developing and characterizing the SPT for a realistic 3D configuration. A full size model of the device will be implemented with the MEDIC software package and the results for the electron concentration integrated in an existing 3D FDTD code.

Impact of the Proposed Research: This project is focused on the development of a 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. The knowledge gained will be also used to apply to the NSF- Electronic and photonic Materials (EPM) program with submission targeting the September 1, 2011 - October 31, 2011 window.

Computational Scientist Time Requested and Other Relevant Resource Allocations: The total workloads for the LI faculty and CS are 6 FTE-months per year, for a total duration of one year and expected allocation time 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.

COVER PAGE

LONI-Assisted modeling of brain tumor cell growth in 2- and 3-dimensions

PI: Mark A. DeCoster, Associate Professor, Biomedical Engineering and Institute for Micromanufacturing, Louisiana Tech University

Lead grad students: James McNamara (Ph.D. candidate); Vivek Dutta (M.S. student), Louisiana Tech.
Suggested Loni Computational Scientist: Dr. Abdul Khaliq, Louisiana Tech.

Current PI funding:

1. "A Center for Advanced Materials and Nanotechnology in AMRI at the University of New Orleans"-subcontract to Louisiana Tech, Board of Regents (State of Louisiana), PKSFI Contract No. LEQSF (2007-12)-ENH-PKSFI-PRS-04, $550,000.
   07/2007-06/2012.

2. "Louisiana's Research Infrastructure Improvement Strategy (EPSCOR: Cybertools)"- PI for Louisiana Tech subcontract through Tulane University. National Science Foundation, Award #0701491, $420,000.
   10/2007-03/2011.

3. "Nanomaterials Safety Lab: Research Integrated with Service and Education (RISE)"- PI. Board of Regents (State of Louisiana), $76,430.
   07/2010-07/2011.

4. "Mathematical Modeling of Biological and Biomedical Engineering Processes"- Co-PI. National Science Foundation, Award #1032176, $99,997.
   09/15/2010-08/31/2012.

LONI-Assisted modeling of brain tumor cell growth in 2- and 3-dimensions

(PI: M. DeCoster)

The proposed project will greatly facilitate modeling of brain tumor cell and normal brain cell growth in 2- and 3-dimensions (2D, 3D) using dynamic cell culture systems for experimental input and assisted by digital imaging and analysis software. Experimental input will come from Louisiana Tech University (LaTech) which has LONI (LI) nodes, and in partnership, we propose to work with LI Computational Scientists (LICSs) to develop computational methods for better visualization and modeling of both normal and diseased brain cell growth. We suggest Dr. Abdul Khaliq as our LICS partner as we have recently published a manuscript together and thus have an excellent and productive history of collaboration **(1).** The genesis for this proposal comes from recently discovered techniques in our laboratory which allow for sustained 3D growth of both normal and tumor brain cells in contiguous cell groups or sphere-like structures (objects). Thus, in our model, each sphere-like structure is maintained as a single object which may be measured in 2 dimensions on a daily basis by non-destructive digital microscopy and image analysis. To date, cultures have been maintained and grown in this system for at least 14 days. In 11 samples analyzed thus far, we have measured on average more than a 300% greater difference in growth area rate for tumor cells vs. normal brain glial cells over a 7 day period. Furthermore, we have measured a large difference in object shape for normal vs. brain tumor cells using our new model. Our computational needs are twofold: 1) computational methods are needed to predict the dynamics of brain tumor object area and shape changes so that potential anti-cancer drug efficacy may be better evaluated; 2) while object growth can be monitored and captured digitally in a non-destructive manner in 2 dimensions in our model, the brain tumor object growth in our model (as well as in reality) is occurring in 3 dimensions, thus computational methods are needed to calculate, infer, or predict volume dynamics from area changes measured and obtained experimentally over time.

We envision a team of researchers at LaTech assisted by LICSs that will work together to iterate experimentally obtained digital images showing a progression of different brain cell growth over time with computational models for prediction and visualization. LICSs will be an integral part of the research team, and we anticipate that if accepted, this partnership will result in the professional development of LICSs to include co-authorship on meeting abstracts, peer-reviewed publications, and funding in future grant proposals to local and national funding agencies. We therefore request the effort of 4 FTE months for partnership with one or more LICSs on this project, suggesting Dr. Khaliq as a possible candidate for this position. As indicated on cover page, the PI has adequate current funding to support ongoing studies for this project which will ensure productive collaboration with the LICS. In addition to our initial publication together, we are including Dr. Khaliq on a grant proposal (requesting 3 years of funding) to be submitted in mid February 2011 to NSF on this topic of computational methods for brain tumor growth. A second proposal will be submitted in May to NIH for clinical applications using this model and Dr. Khaliq will play a key role there as well. We anticipate that linkage to LONI resources will aid in the overall strength of the proposals to both of these agencies, increasing our chances for external funding.

(1). Khaliq, A., Jenkins, F., DeCoster, M., Dai, W. "A new 3D mass diffusion-reaction model in the neuromuscular junction". *Journal of Computational Neuroscience* pages 1-17 (**2011**) In Press.

A High Performance Data Analytics framework for Large Scale Data Analysis

Dr. Ramesh Kolluru, Dr. Raju Gottumukkala CBIT, UL Lafayette
Dr. Philip Caillouet, Louisiana Center for Health Informatics
Dr. Lee Bairnsfather, LSU Health Sciences Center, Shreveport
Dr. Vijay Raghavan, Dr. Ryan Benton, CACS, UL Lafayette
Dr. Paul Miller, Medical Director for Kidney Consultants of LA, Dialysis Clinics Incorporated; Kim Bahrami, Evolvent; Joe Block, McKesson; Doug Menefee, The Schumacher Group

Introduction

Several industries and government agencies, ranging from healthcare, bioinformatics to homeland security are faced with an increased demand to analyze data large scale data in near-real-time to enable insightful and informed decision making. Existing data analysis using statistical tools such as data mining algorithms typically run on standalone workstations. However, the increasing availability of large-scale multidimensional data and the demand to extract knowledge for real-time decision support has driven the need to develop a High Performance Data Analytics framework.

Project Significance

This project is of great interest to the state of Louisiana in its efforts to implement the Louisiana's Health Information Exchange (LaHIE). This will benefit two major collaborative projects at the University of Louisiana at Lafayette.

1. We have recently submitted two proposals to provide an operational technology to allow for rapid and successful implementation of Louisiana's Health Information Exchange (LaHIE) to the Louisiana Health Care Quality Forum (LHCQF) in partnership with various healthcare companies (amount: $10.6M). The second proposal was submitted to the Office of National Coordinator (amount: $2.3M) in partnership with LHCQF.
2. We are also establishing an NSF I/UCRC Center for Visual Decision Informatics at the University of Louisiana at Lafayette in partnership with Drexel University. UL Lafayette has been awarded a planning grant (of $10K) and has conducted a preliminary workshop for industry members in November 2010. We are working to submit a full proposal to NSF in summer of 2011.

The project will also leverage the work on two previous LI projects titled "Parallel Algorithms for Large Scale Data Clustering" and "Parallel Optimization Algorithms for Disaster Management during which we experimented a suite of parallel algorithms to parallelize some clustering and data mining algorithms.

Goals

The main goal of this project is to develop a distributed data management component and a data analytics component on the federated LONI cyberinfrastructure. This includes the development of a gateway to provide an easy to user portal based graphical user interface, investigate middleware components to manage federated data, development of a suite of data analysis components that leverage the distributed cyberinfrastructure for decision support. This will enable analysis of

Effort Requested of LI Computational Scientist

We request 9 months of the LONI Computational Scientist, Dr. Raju Gottumukkala's for this project. The LONI Computational Scientist would help with investigating the extending of existing suite of data mining tools and libraries and also develop gateway to LONI and the TeraGrid. The LI Scientist will be supported by two Master's students and a system administrator.

Benefit to LONI Institute

The LaHIE aligns with the Blue Ocean Strategy of the State of Louisiana. This research project also aligns well with the LI metric goals in terms of supporting the establishment of I/UCRC centers in the state. The development of high performance data analytics package would advance the LONI infrastructure as these tools will be available for industry and academic users of LONI.

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. iLevee is a collaborative project that includes Geocomp Corporation, PB Americas, Shannon & Wilson, James Lee Witt Associates, NIMSAT Institute at the University of Louisiana at Lafayette, SMARTEC and TIE Technologies.

Granting Agency: LA-OCPR
Grant Duration: 2010-2013
Funding Amount: $2.9M

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.

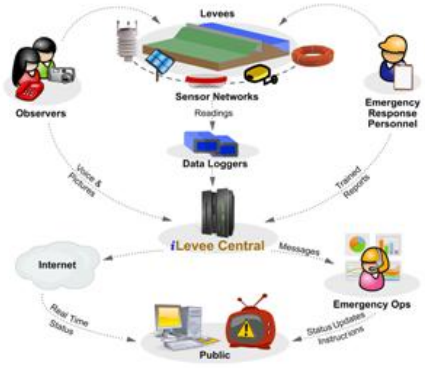


Figure 1. Concept of the iLevee System (Taken from the Proposal)

Time for LI Computational Scientist

The primary objective of this project is to develop and deploy an iLevee Central system that runs on LONI. Dr. Gottumukkala was involved in writing the proposal. This project is an extension of previously submitted LI 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. An NSF Proposal is also currently under development with Dr.Kara Fadi from NJIT.

Leveraging Many Jobs
To Enable
High Performance High Throughput Simulations
on LONI and the TERAGRID.

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

bishop@tulane.edu

504-862-3370

LONI CS Request:

Hideki Fujioka LONI Computational Scientist, Tulane University

Time: 6 months FTE-months

Project Description: In our NIH funded study (R01GM076356, Feb. 2006-Jan. 2012) one aim is to investigate sequence dependent variations in the nucleosome using molecular dynamics simulations. because all available x-ray structures of the nucleosome utilize nearly the same 147 bp sequence of DNA (Luger, et al., 1997.) In theory there are some 4^{146} possible sequences of nucleosomal DNA. In practice the ~12 million basepairs in the yeast genome positions about 60,000 nucleosomes (Jiang, et al., 2009). We have adopted a forward looking workflow approach to our simulations that enables us to conduct 100's of simulations now and that enables us to utilize $Peta^2$ (Petabyte by Petaflop) resources in the future. IBM developers project such resources will be widely available by 2015 (Jordan, 2010).

Progress to Date: We initially implemented a high-performance high-throughput molecular dynamics workflow on LONI to conduct 16 simulations(Bishop et al., in prep.). Our workflow was inspired by NAMD-G (Gower, et al., 2006), and structured to take advantage of LONI's computational and data storage resources. In total there were 16*16 = 256 tasks. Each task begins by fetching the necessary inputs and ends by depositing outputs into Petashare(Wang, et al., 2009). This first manual implementation produced 16ns trajectories for each of the 16 systems, ~1.5Tb of data, and required nearly 300,000SU at four LONI sites. With support of LI Computational Scientist Hideki Fujioka we automated the process by implementing the ManyJobs concept (Luckow 2009) using phyton scripting and secure shell authentication. ManyJobs.py is a portable, lightweight tool. It has been used by Dr. Fujioka to conduct a series of simulations parameterizing a computational model of pulmonary small airways for the Gaver Lab at Tulane University. The Bishop Lab utilized the same ManyJobs.py tools to conduct a second molecular dynamics study of nucleosome positioning. This study included 42 different systems each contained ~150,000 atoms. Each system was simulated for 20ns. In total the study required 840 tasks, generated ~3.5Tb of data, and required over 250,000SU. The limiting factor in achieving higher throughput was primarily data management within the LONI system. Several Petashare servers proved unstable during the last quarter of 2010.

Proposed Developments by LONI-CS.The specific objectives for this year are to: 1) incorporate more robust fault tolerance into the workflow. This enables us to address the file management bottleneck and provides other benefits. 2) Integrate advanced security features available via SAGA into ManyJobs.py. This enables seamless access to LONI, TERAGRID, and other resources including alternate storage options. 3) Enhanced usability. Improve documentation, distribution & sharing (svn) and extend the number of examples.

Need for LI Support: We currently have large TERAGRID (8,000,000SU, June 2010 to July 2011) and LONI (2,000,000SU Jan 2011 to Dec. 2012) allocations that are expected to produce some 50Tb of data. The proposed developments seamlessly integrate local (e.g. Tulane's cluster), LONI, TERAGRID and even cloud resources. It also provides access to numerous data storage resources. Dr. Fujioka has the necessary skills and experience to accomplish this task. We estimate 6 months of FTE.

Benefit to LONI Institute: ManyJobs.py has been utilized by the Bishop Lab and Gaver Lab at Tulane to accomplish entirely different computing studies. Two LONI associated faculty Randy Hall, Chemistry at LSU and Lawrence Pratt Chemical Engineering at Tulane have begun working with ManyJobs.py for other purposes. We expect the proposed improvements will benefit all LONI users, be of particular use to LA-SIGMA computational researchers, and be an integral part of the statewide effort to establish a computational center of excellence in materials science.

References Cited:

- **Gower Michelle [et al.]** Managing Biomolecular Simulations in a Grid Environment with NAMD-G TCBG Publication. - 2006. - pp. 1-7. http://www.ks.uiuc.edu/Publications/Papers/
- **Jordan, Kirk.** Data and Analytics Transforming Healthcare for Smarter Healthcare & Life Science, 8th Annual Rocky Mountain Meeting, International Society of Computational Biologist, Snowbird, CO. Dec, 2010.
- **Jiang Cizhong and Pugh B. Franklin** A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome. Genome Biol. - 2009. - Vol. 10. - p. R109.
- **Luckow A. [et al.]** Adaptive-Distributed Replica Exchange Simulations. Phil. Trans of the Royal Society A - 2009. Vol. 367 – pp 2595-2606.
- **Luger K. [et al.]** bCrystal structure of the nucleosome core particle at 2.8 A resolution. Nature. - 1997. - Vol. 389. - pp. 251-260.
- **Wang X. [et al.]** Semantic Enabled Metadata Management in PetaShare International Journal of Grid and Utility Computing. - 2009. - Vol. 1. - pp. 275-286.

Proposal

LONI Institute Computational Scientists (LI CSs) Support

Researcher: DQSI, LLC, a Louisiana-based Small, Woman-owned Business (Dr. Shelly Stubbs)
Project: "Distributed GIS Computing for High Performance Simulation and Visualization"
Sponsor: NASA (Under SBIR Program), $100k Phase I Grant; $2.5k LED Phase 0 Grant
Project Period: February 1, 2011 – July, 2011.
Partners: University of New Orleans (Environmental Research/CFD Modeling)
Vanderbilt University (Logistical Networking)

Abstract: Today, the ability of sensors to generate geographical data is virtually limitless. Although NASA now provides a capability for accessing, manipulating, and visualizing these data, an enormous gap still exists between the sensor data and the visualization capability. The challenge addressed in this research is to develop new methods for automating the access and processing of raw sensor data, sharing data and methods among researchers, and integrating visualization throughout the process. To meet this challenge, we will develop two new innovations: 1) Geo-Cloud, and 2) COAST HPGIS. **Geo-Cloud** is the distributed network environment that maintains connectivity between a geo-browser and Geo-Resources (sensor data, GIS datasets, models, simulations, and metadata). The Geo-Cloud maintains all metadata (and ontology) required to efficiently catalog, locate, store, access, and update Geo-Resources. **COAST HPGIS** is an enhanced version of the NASA COAST product which is based on the NASA World Wind geo-browser. COAST HPGIS includes plug-ins, overlays, and interfaces for COAST in order to enable interaction with the Geo-Cloud for real-time visualization of temporal and parametric simulations and models.

Computational Needs: The following tasks provide a description of the support required from LONI CSs.
Task 1 – Consultation. Consultations for Visualization tools, Cloud computing, and advanced GRID techniques for the project.
Task 2 – HPC/CFD. Provide assistance and HPC programming to port two computational programs to the LONI HPC environment and optimize parallel processing.

- FVCOM or ECOMSED (CFD Models – (Fortran90, netCDF interfaces)
  - http://fvcom.smast.umassd.edu/FVCOM/index.html
  - http://www.hydroqual.com/ehst_ecomsed.html
- TETRACORDER/SPEC-PR (Spectroscopy analysis tool) - (ratfor, fortran77)
  http://speclab.cr.usgs.gov/tetracorder.html

Task 3 – Visualization. Provide assistance with image processing tools for the generation of outputs from Task 2 for GIS visualization – (MATLAB, PLAPACK, OpenGl, etc.).

- 3D visualization and manipulation of dynamic metadata structures
- Rendering of geographical features (from raw data or simulated outputs).

Timeframe/FTEs: We had identified two LI CSs who could support this project. Their initial feedback was positive for this support. Scheduling of FTEs is flexible during the project period.

- Hideki Fujioka – HPC/CFD (Tasks 1 & 2) - 3 FTE months
- Zhiyu (Sylvia) Zhao – Visualization (Tasks 1 & 3) - 3 FTE months

LI Benefits: LI will benefit through the continued professional development of CSs, participation in the development of Louisiana high-tech development, and furthering scientific understanding of advanced networking and HPC within the earth-based science community. It is expected that several papers will be published as a part of this research. LI CSs, as well as collaborators from UNO and Vanderbilt, are expected to co-author papers related to their participation.

Improving Antibody Design by Structure Prediction: Comparison of Computational and Molecular Approaches

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

David K. Worthylake (dworth@lsuhsc.edu)

Department of Biochemistry and Molecular Biology,
Louisiana State University Health Sciences Center, New Orleans

This project is for a second year of support to allow Zhiyu Zhao (Sylvia) to continue with the improvement of antibody (Ab) design, a problem of great scientific interest. It is also an important year for the development of Sylvia as a computer scientist who also has a deep understanding of biology. The research addresses the design of high affinity Abs, which can be used as therapies for human disease, in biodetectors and diagnostics, or as a well-characterized model for general protein design. The project studies a monoclonal Ab that neutralizes the toxin ricin, a molecule of biodefense concern. Although it is already an effective Ab, we postulate here that it can be made better by rational design (other projects in the lab are taking more "random" combinatorial approaches). As a first step, we must solve the 3-D structure of this Ab as it binds to ricin. During the first year of the project, Sylvia has utilized various computational approaches, including Ab modeling and protein docking. In the laboratory, she has learned to produce, purify, and characterize monoclonal Abs. She has prepared sufficient Ab to allow us to proceed to the next step: Ab crystallization.

Computational Approaches

Sylvia has developed a primitive computational tool which combines homemade programs with publically available resources to allow for computational tasks such as CDR loops identification of Abs, sequence search, structure search, structure alignment, Ab structure modeling, Ab-Ag docking and binding energy estimation. Using this tool, she has been predicting the 3-D structure of the anti-ricin Ab, RAC18, and the complex structure of RAC18 binding to ricin A. She has also been using her sequence and structure search programs to find similar proteins structurally similar structures to RAC18 and docking them with the known structure of ricin to predict sites of contact, and how protein-protein interactions may be improved to produce higher binding affinity. Once the real structures of RAC18 and the Ab-ricin complex are determined, we will compare computational models with real structures and determine the accuracy of different models. Knowing the real structures will allow refinement of structure prediction and docking models.

Biological Experiments

Bioinformatics and computational biology are closely related fields that are undergoing a rapid expansion. Unfortunately, the number of people able to understand the complexities of both biomedicine and computer science are limited. It is our belief that the best way to attain expertise of this sort is to expose computer scientists to the rigors of molecular biology, bioassays, and analysis of such data. Over the past year, Sylvia has learned a variety of immunological and biochemical techniques, and prepared 20 mg of functional Fab fragments. The unaggregated monomeric state of these Fab fragments is now being studied by gel-filtration chromatography. The next stage of the project will be to screen for crystallization conditions. These studies will be performed in the laboratory of David Worthylake, a crystallographer at LSUHSC-NO. Complexes of Ab-ricin will be made in high concentration. High throughput screens, using the hanging drop method in microwells, will be used to establish optimal crystallization conditions. When these are identified, large scale crystallization will be performed. If sufficient high quality crystals are obtained, then X-Ray diffraction studies will be performed.

Longterm goals

Once the 3-D structure of the Ab binding to ricin is obtained, we will compare the results with those from computational modeling. Because Ab is such a well-studied molecule, the chance for informative results is high. These results can be used to refine existing computational models for the 3-D structure and antigen binding interfaces of Abs. The next step is to use this 3-D model to computationally predict amino acid changes that will produce a higher-affinity interaction between Ab and ricin. These changes can then be designed, synthetic genes made, and mutant Ab produced and tested for biological function. The ability to rationally improve Ab binding to Ag is a highly sought goal, and can be generalized to the improvement of other protein-protein interactions.

Computational needs

We request 6 FTE months of Sylvia to use the LONI's HPC resources to accomplish large-scale Ab modeling and docking tasks using tools such as the Rosetta Software Suite and AutoDock which are available on the LONI clusters. Sylvia will also perform the crystallization of RAC18 and its Ab-Ag complex.

Benefits to the LI

Besides its importance in improving antibody design, this interdisciplinary and inter-institutional research proposal well meets the merits of LI. The project will broaden the scientific impact of LI in the participating institutions, has great potential to attract external funding support, and the developed computational tools will benefit researchers who are interested in designing novel proteins for efficient interactions. We have been regularly broadcasting a joint bioinformatics seminar with UNO's bioinformatics faculty via the Access Grid.

Other support for this project

This project has been supported by NIH grant AI059376, Children's Hospital (a student stipend and supplies), and the Louisiana Vaccine Center. Dr. Worthylake is the recipient of NIH grant GM084072, and was the prior recipient of an RCS grant from Louisiana Board of Regents.

Dear LONI officials:

I hereby submit our collaborative proposal to continue and significantly expand our proposed research on human transcriptome identification and quantification. In the last year's project, we proposed computational infrastructures for transcriptome quantification at the isoform-level using the second-generation RNA sequencing data (RNA-seq). We have developed and implemented an Expectation-Maximization (EM) type algorithm to quantify the whole catalogues of transcripts in Alternative Splicing and Transcription Diversity (ASTD) database. We have also demonstrated significant biological applications. Our past work has led to significant advance in the computational side of this research. In particular, our GUI software engine has recently been published in *Source Code in Biology and Medicine* (Xu et al. 2011, 6:2) and our transcriptome quantification algorithm and analysis have been published in *Nucleic Acids Research* (Deng et al. 2011, 10.1093/nar/GKR042).

Compared with transcriptome quantification, identification is considered as a more challenging task. Transcriptome is highly diverse, overlapping, complex, and dynamic; therefore, identification of context (tissue, cell line or time) specific transcriptome is a central to ultimately decipher many human diseases. Computationally, given tens of millions of short reads generated from different locations of a transcriptome, we aim to develop an accurate, fast and scalable algorithm and software to infer the structure of the whole transcriptome. If successful, the software will identify the structures of hundreds of human disease transcriptomes and significantly expand the current catalogue of transcripts.

We have extensive experience in developing and applying statistical models, computational algorithms and software to transcriptomic research. We have access to the latest second generation sequencing platforms; including an Illumina Genome Analyzer and support from well-established research facilities at the Tulane Health Sciences Center (Dr. Zhu is a program member of Tulane Cancer Center). Selected results from our bioinformatics analyses will be tested and validated experimentally. Because we are addressing important, timely and vexing research and clinical problems, we believe that the progress we will make in carrying out the proposed informatics research will have immediate and far-reaching impact in the broader area of biomedical research and foster improved clinical trials.

Finally, we have strong experience in the proposed research domain, demonstrated by four recent research manuscripts pertaining to transcirptome research using RNA-seq data (Xu *et al.*, 2010, *RNA*, 16(8), 1610-1622; Lin *et al.*, 2010, *Journal of Virology*, 84(24), 13053-13058; Xu *et al.,* 2011, *Source Code for Biology and Medicine*, 2011, 6:2; Deng *et al.,* 2011, *Nucleic Acids Research,* 10.1093/nar/GKR042). Our current bioinformatics research is supported by a number of extramural research grants from NIH (R2LM010137, 09/01/09-08/31/11, will request one-year extension), Tulane Health Sciences Center (07/01/10-06/30/11, renewable) and Children's Hospital at New Orleans (01/01/11-12/31/11, renewable).

Sincerely,

Dongxiao Zhu, Ph.D.

Assistant Professor Department of Computer Science

University of New Orleans

Isoform Level Human Transcriptome Identification using RNA-seq

Dongxiao Zhu (UNO) and Zhiyu Zhao (UNO)

Our goal is to develop a computational infrastructure with a Graphical User Interface (GUI) for identification context-specific human transcriptome at the isoform level and to make it freely available for biomedical researchers to substantially accelerate the disease transcriptome research. Next generation sequencing has provided unprecedented opportunities to *de novo* identify transcript structures and quantify transcript abundance at the isoform level. Despite improvements in sequencing methods and bioinformatics advances allowing *de novo* construction of transcriptoms, the existing approaches are often not sufficient to detect certain transcripts.

We propose a new statistical model to explain how the observed base-wise coverage signal is accumulated from a mixture of sibling isoforms. The relationship between *observed* base-wise signal from RNA-seq experiments, *latent* isoform proportion and *latent* gene expression score can be modeled as:

$$E = W * R + \varepsilon,$$

where $E \in R^{MXN}$, $W \in R^{MXL}$, $R \in R^{LXN}$, M represents the length of the transcriptional unit, L represents the number of annotated isoforms and N represents the number of samples. When the gene expression abundance is unknown, we propose a new constrained optimization algorithm to estimate isoform proportion and gene expression level simultaneously. We start with a random guess of W, subject to sum-to-one constraint, denoted as $\widehat{W}$. The constraint is translated into the fact that sum of all isoform proportions of the same gene adds up to 1. The above object function can be decomposed into the sum of the object function on each sample, i.e. $min||E - W * R||^2 = \sum_j^n min||E - W * R||^2$. With this decomposition, in the Step 1, we will estimate the gene expression abundance for the i-th sample as $\hat{r}_j = \frac{E(:,j)\widehat{W}A}{||\widehat{W}A||^2} \geq 0$. In the Step 2, we will use the closed-form solution derived in the step 1 to update the estimate of $\widehat{W}$. Step 1 and step 2 alternate until the $\widehat{W}$ converges.

Thus the real challenge comes in the computational side in that our 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 *100* to a few hundreds, and theoretically *N* can be up to *22,000*. 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.

This project will benefit LONI by fostering collaborative efforts from Louisiana institutions in their efforts in multidisciplinary human health related research. The proposed computational infrastructure is quite general and widely applicable to diverse human disease tissues wherever the RNA-seq data is available. Therefore, the proposed computational systems will benefit a wide range of researchers at large. This work also fits well 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.

We request six months of full time effort of Dr. Zhiyu Zhao to develop a parallel algorithm for identification of human transcriptomes using RNA-seq, and implement it on the LONI clusters. Dr. Dongxiao Zhu has considerable experience working with Dr. Zhao to develop computational algorithms and software for human transcriptome identification and applications.

Design a novel environmental barrier coating from *ab initio* simulation

PI: Shengmin Guo (sguo2@lsu.edu Mechanical Engineering Department, Louisiana State University)
Co-PI: Shizhong Yang (shizhong_yang@subr.edu Computer Science, Southern University)

Introduction: This project will support a current NASA project, lead by Drs.Guo and Yang (award No. NASA/LEQSF (2009-2012)-Phase3-03, $1.5 M, Oct. 2009 ~ Sept. 2012). Ceramic coatings are used in the hot section of rocket and jet engines to safeguard them under extreme working temperatures, ~1600ºC, and in highly oxidizing and corrosive environments. To improve the engine durability and efficiency, new coatings are required for the hot section of future gas turbine engines. Ceramic coating is usually prepared from ceramic powders, such as yttria-stabilized zirconia (YSZ), using a plasma spray process. However, the current YSZ based coatings and Ni-based super alloys cannot meet the DOE targets for next generation gas turbine engines. Due to the high melting point, Nb-based alloy has the potential to meet those stringent NASA/DOE targets. Thus currently, there is a strong effort by NASA/DOE scientists to examine how to improve the Nb-based alloys, especially how to overcome the rapid high temperature oxidation problem. Dr. Guo's group has setup all the required instruments for experimental study on novel MAX phase materials and his group has obtained very encouraging data on Ti_2AlC MAX materials. The PIs believe special layered Nb_2AlC MAX phase materials have the potential to become an excellent environmental barrier to mitigate the oxidation of Nb based alloys. This proposed project will simulate the stability and optical property of Y-doped Nb_2AlC and the interface of Nb_2AlC/Nb, based on the interface energy simulation. The simulations will be performed on SU workstations and LONI supercomputers.

Proposed research: The goal of this proposal is to develop and test an efficient environmental barrier coating(EBC) design method based on *ab initio* simulation and interface energy calculation, to enhance the close collaboration among LSU, SU, and NASA research centers, and train under-represented minority students in nano-material high performance computing. Using *ab initio* method, the electric structure, optical properties, lattice dynamics, and phase transformations for bulk Y-Nb_2AlC will be calculated. *Ab initio* Density Functional Theory (DFT) based molecular dynamics simulations will be employed to explore and to optimize the electronic structures and physical properties of EBCs. Both surfaces of Nb alloy and Nb_2AlC will be simulated and the relative stability will be compared. Then the interface energy will be simulated based on the stable surfaces. The outcome of the computation, simulation, and analysis of the properties of the targeted materials will aid us to choose the material compositions and the microstructures. Dr. Yang will use *ab initio* MD methods to fully relax the bulk Y-Nb_2AlC and Y-Nb_2AlC/Nb interface structures. He will then calculate the electronic structure (DOS, charged state and band structure *etc.*) and optical properties. Dr. Yang plans to spend **3.5 months** to study the electrical and optical properties of Y-Nb_2AlC/Nb interface. It is planned as follows: 0.7 month --- build and test the interface model; 1.8 month --- test the Modules of the MedeA code and run MD, optical property, interface energy simulation jobs on LONI machines; 1.0 months --- analyze the MD and optical property results, adjust the dopant, temperature and pressure parameters from Dr. Guo's sample tests, finalize the optimized parameters and write summary and submit papers and/or patents.

Broad impact of the proposed research: The proposed project will develop and test an effect EBC design method with intensive use of LONI HPC facilities. A significant amount of LONI CPU time is required to complete this challenging and high rewarding project. Currently, Dr. Yang's group has six graduate students and one undergraduate student working on a variety of funded projects. The proposed project will train under-represented minority students in high performance computing. This project will promote a close collaboration between LSU, SU, and NASA research centers to perform advanced research. Its outcomes have the potential to lead to long-lasting research, educational, and outreach efforts, which will reach far beyond the proposal period. The activities will foster cross-disciplinary interactions between LSU and SU, and serve the needs of students across campus. Successful implementing the proposal will lead to submit new proposals to NSF, DoE, DoD and NASA with potential big impacts to gas turbine industry.

Molecular dynamics simulation of the interactions of apocynin, 5-Nitroapocynin, and diapocynin with 1K4U subsection of human neutrophil NADPH oxidase system

Rao M. Uppu (PI: rao_uppu@subr.edu) Environmental Toxicology, Southern University, Baton Rouge, LA 70813; Shizong Yang and Ebrahim Khosravi (Co-PIs: shizhong_yang@subr.edu; ebrahim_khosravi@subr.edu) Computer Science, Southern University, Baton Rouge, LA 70813

Introduction: This project is supported by funding from NSF SBIR IIa program of NCRR [grant number IIP-0956877, $505,014, 12/1/2010-3/31/2012]. Apocynin (Apo) is a potent phenolic antioxidant isolated from plants belonging to the apocyanaceae family. Apo has long been recognized for its potential to treat inflammatory and degenerative conditions. It is generally believed that Apo inhibits the plasma membrane NADPH oxidase system by interfering with translocation of key cytosolic components $p40^{phox}$, $p47^{phox}$ and $p67^{phox}$. In order to gain an insight into the nature of ligand- binding sites, we performed flexible docking of Apo and DiApo with 1K4U subsection of NADPH oxidase system using ICM-Pro (Molsoft) software. The study also included docking of 5- nitroapocynin (5-NitroApo), a nitration product known to be formed in reactions of Apo with nitric oxide-derived oxidants namely peroxynitrite and its CO_2 adducts. We have finished the ICM docking simulation and find the possible docking sites. The molecular dynamic simulations are needed to detail the atomic level interactions of Apo and its metabolites with $p47^{phox}$ and $p67^{phox}$ subunits of the NADPH oxidase system. We propose **three months** of Dr. Yang's research time to set up models, build force field, perform HPC simulation, analyze the results, interactions and properties.

Proposed research: Plasma membrane NADPH oxidase of phagocytic cells is a multi-component system consisting of three cytosolic proteins ($p47^{phox}$ and $p67^{phox}$, and $p40^{phox}$) that translocate from the cytosol to the flavocytochrome in the plasma membrane upon stimulation of the cells. The enzyme system is the major source of superoxide anion which subsequently is converted to much stronger oxidants as hypohalous acids and peroxynitrite. While these oxidants play a pivotal role in bactericidal and fungicidal activities of phagocytic cells, the same oxidants can induce inflammation and oxidative stress in surrounding cells and tissues. Often, it is regarded as a kind of collateral damage that superoxide-producing cells need to incur in addition to their own loss due to oxidant-induced apotototic and/or necrotic cell death. Apocynin, a naturally-occurring phenolic antioxidant has long been known for its anti-inflammatory properties. The underlying mechanism(s) of apocynin and/or its metabolites inhibiting NADPH oxidase system has not been elucidated yet. We have shown that apocynin can also undergo nitration in the peroxynitrite/CO_2 reactions and form 5-nitroapocynin, and the possible crystal structure of 5-nitroapocynin has also been elucidated. Herein we attempt to simulate the interactions of apocynin, diapocynin and 5-nitroapocynin of the $p47^{phox}$ subsection (PDB: 1K4U) of NADPH oxidase system using NAMD software and analysis tools plus Yang's group's property calculation codes. Dr. Yang will spend one month each on the following three parts of the research respectively: setup the model and force fields, perform HPC MD simulation on LONI machines, and calculate, analyze, and compare the three types of systems.

Broad impact: The proposed project will build up the three types of force field for the MD simulation and calculate the interactions and properties of the three systems with intensive use of LONI HPC facilities. Currently, Dr. Uppu has a post doc and a Ph. D. student working on the experimental part of the research. Dr. Yang has six graduate students and one undergraduate student working on his projects. The proposed project will train under-represented minority students in high performance computing. Its outcomes have the potential to lead to long-lasting research, educational, and outreach efforts, which will reach far beyond the proposal period. We believe successfully implement the project would lead to further NSF/NIH funding support.

First principles molecular dynamics simulation on the formation mechanism of a novel core-shell copper-carbon nano-particle for fuel cell and wood treatment applications

Kun Lian (PI: klian@lsu.edu) CAMD, Louisiana State University, Baton Rouge, LA 70806; Shizhong Yang (Co-PI: shizhong_yang@subr.edu) Computer Science, Southern University, Baton Rouge, LA 70813

Introduction: This project will support the funding from USDA-CSREES (Kun Lian's grant #2008-38814- 04771, $448,877, Sept. 2008 ~ Aug. 2012] and LONI NSF LASiGMA project (grant number #EPS-1003897, 20 M, Oct. 2010 ~ Sept. 2015, Shizhong Yang's SU ongoing project). Dr. Kun Lian's group has recently discovered a novel copper metal-carbon core-shell nano-particles(US patent approved in 2009), which have the potential to be used in wood treatment and in fuel cell catalyst to achieve high reactivity, tunable selectivity, enhanced poisoning resistance, self-repairing capability, long-term stability, and low cost. The advantages of this technology are: (1) simple one-step process; (2) the nano-particles are very stable in aqueous environment; (3) the nano-particles have netlike carbon shells and can be surface engineered after forming the metal nano-particles; (4) a simple treatment, such as flash heating, can recover the partially oxidized nano-particles to the original metal nano-particles. Although we have experimentally verified the core-shell structure to be copper-carbon core-shell, we do not know how the core-shell structure is formed under the temperature of 350~400 ºC. In this proposed project, Dr. Yang will use density functional theory (DFT)-based molecular dynamics (MD) simulations to reduce the cost and time needed for understanding the nano-particle formation mechanism. We propose **four months** of Dr. Yang's research time to set up models, perform HPC simulation, analyze the results, and study the nano-copper formation mechanism.

Proposed research: The core-shell structure of copper nano-particle covered by carbon layers has shown potential for fuel cell, corrosion protection applications. Coating copper nano-particles with a carbon layer appears to protect the copper against oxidation, while allowing the copper nano-particles to retain useful properties. The general method for producing novel metal-core carbon-shell nano-particles comprises soaking a natural fibrous material (cellulose) with a solution containing metal ions, removing the solvent, and then carbonizing the impregnated fibers at a temperature sufficient to generate metallic cores encased in carbon shells. We will perform first principles DFT simulation on the copper atoms/cellulose segment model (two cellulose units and bound with Cu atoms to each electron-rich oxygen atoms in the cellulose unit). The Vienna *Ab-initio* Simulation Package (VASP) integrated into MedeA will be utilized to efficiently perform simulation. In the mean time, we will test and optimize the efficiency of the HPC simulation on the LONI machines. We will analyze the results and propose copper binding model to explain the possible core-shell formation mechanism. Dr. Yang will spend four months on the following research: setup the model and test on LONI machine with both VASP and MedeA on three layers of applications, i.e. laptop-job server-supercomputer, perform HPC MD simulation on LONI machines, calculate, analyze the results, and propose the formation mechanism.

Broad impact: This project will partly support current LONI LASiGMA project and help enhance fundamental understanding in the area of manufacturing science and technology: nanofabrication and self-assembly. The proposed project will build up the efficient HPC working environment for the material design with intensive use of LONI HPC facilities. Dr. Yang has six graduate students and one undergraduate student working on his projects. The proposed project will train under-represented minority students in high performance computing. Its outcomes have the potential to lead to long-lasting research, educational, and outreach efforts, which will reach far beyond the proposal period.

A novel *ab initio* molecular dynamics method for Cr-Y high temperature alloy simulation

PI: Shizhong Yang (Shizhong_Yang@subr.edu Computer Science, Southern University)
Co-PIs: Ebrahim Khosravi(Ebrahim_Khosravi@subr.edu, Computer Science, Southern University)
Shengmin Guo(sguo2@lsu.edu Mechanical Engineering Department, Louisiana State University)

Introduction: This project will support two current DOE projects, lead by Drs. Yang and Guo respectively (award No. DE-FE0004734, $199,596, June 2010 ~ May 2012, and award No. DE-FE0003693, $500,000, Oct. 2010 ~ Sept. 2013) and in collaboration with NETL Principal Material Scientist M. Gao in high temperature alloy simulation research. The goal of this project is to develop reliable interatomic potentials from the highly accurate *ab initio* molecular dynamics (MD) calculation for Cr-Y system. Clean coal power generation is critical for the United State economic growth and international competitiveness. The key component for advanced coal power generation is a durable high temperature alloy. Chromium(Cr), a refractory metal, based high temperature alloys show considerable promise due to their relative low cost, low density, and good high temperature strength. We will use the developed interatomic potential to optimize the Cr-Y composition and processing environment through enhanced high temperature microstructures, melting points, elastic constants, diffusion coefficients, activation energy, and oxidation and sulfate corrosion resistance simulations. We propose **three months** of Dr. Yang's research time to develop the potential generating codes, perform property test, and develop HPC MD code and perform simulation on LONI HPC machines.

Proposed research: In this project, we will integrate *ab initio* calculations with MD simulations to enhance capabilities for materials modeling and prediction. As is well known, an appropriate potential of the interatomic interaction is the key to the success of the MD simulations. Unfortunately, not much data is available for complex materials. Ideally, a microscopic theory of quantum calculation that can produce such a potential would be very beneficial. Therefore we will utilize *ab initio* quantum computations to provide the input information of the interatomic potentials that will be used in MD simulations. The resulting computer codes will have the capability of MD calculations with the reliability of the *ab initio* method. We will develop the interatomic potentials for Cr, Y, and Cr-Y. Then the NVT ensemble will be chosen to simulate the high temperature elastic constants and diffusion coefficients and activation energy. The typical NVT MD run will be about 30,000 MD step with t=0.002 ps for the simulation. The small size test jobs, to test the parameter settings, will be performed on our workstations. After the scalable code tests, the HPC high temperature molecular dynamics simulation will be performed on Louisiana Optical Network Initiative (LONI) and TeraGrid supercomputers.

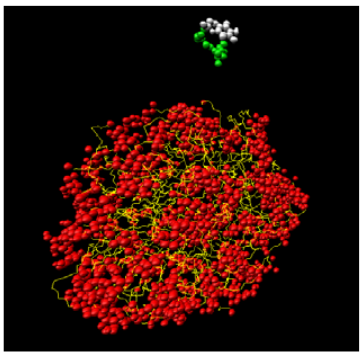
Broad impact: The proposed project will develop a novel ab initio MD simulation method and perform test on the Cr-Y alloy using LONI HPC facilities. The project will enhance our current collaboration with LSU and NETL lab to perform advanced research. Currently, Dr. Yang's group has six graduate students and one undergraduate student working on a variety of funded projects. The proposed project will train under-represented minority students in high performance computing. Its outcomes have the potential to lead to long-lasting research, educational, and outreach efforts, which will reach far beyond the proposal period. Successful implementing the proposal will lead to secure more fundings from DOE and NASA with potential big impacts to material science and especially for gas turbine industry.

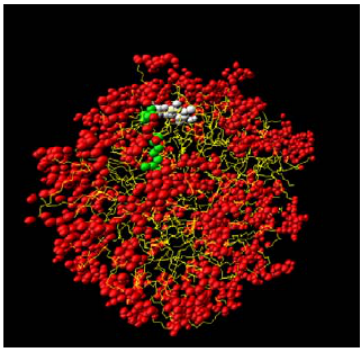
Development of all-atom force fields for bio-molecular systems

PI: Dorel Moldovan
Department of Mechanical Engineering, Louisianan State University, Baton Rouge, LA

Co-PI : Hank Ashbaugh
Department of Chemical and Biomolecular Engineering, Tulane University, New Orleans, LA

Co-PI: Bhupender Thakur,
Center for Computational Technologies, Louisianan State University, Baton Rouge, LA





Part of Dr. Moldovan's research is currently supported by grants from USDA (grant # 2009-35603-05055; $325,000, 10/1/2008-9/30/2012) and NSF – LA-SiGMA (grant # EPS-1003897; $20,000,000, 10/1/2010-9/30/2015)

Motivation: Bio-molecular systems such as membranes, DNA, and proteins have unique attributes that make them suitable for a large number of applications of great importance in pharmaceuticals and biosciences. Much of the advancements in these applications are hindered by our limited understanding of the fundamentals of interactions of these bio-systems with various chemicals and surfaces. Most of our recent investigations are focused on using large-scale molecular dynamics (MD) simulations, to address a host of open research problems pertaining to the areas of cryopreservation, DNA sequencing, bio-molecular machines, and drug delivery. It is well established that, apart from using appropriate simulation conditions, the quality of the force fields used are crucial ingredients in obtaining reliable simulation results.

Project Description: The empirical force field used in MD simulations describes the interaction potential energy of a system of atoms or molecules with certain functional terms representing covalent interactions between atoms (such as bond-stretching, bond-angle bending, improper and proper dihedral-angle torsion) on one hand and non-bonded interactions (van der Waals, electrostatic) on the other hand and a set of parameters characterizing these interactions terms. Despite of extensive use, and recent developments, not all molecules of interest can be parameterized directly by direct reference to the available, well-established, force field packages, such as AMBER, CHARMM, OPLS, GROMOS, etc. Therefore there is great deal of interest in expanding and streamlining the procedures and methodologies for the development of force fields and this is the goal of the proposed project. Given the current interest of the PIs in the mechanisms of self-assembly of various amphiphiles into structures that may serve as drug delivery vehicles, the development of a coherent strategy for parameterization of the interactions in such systems is of great relevance.

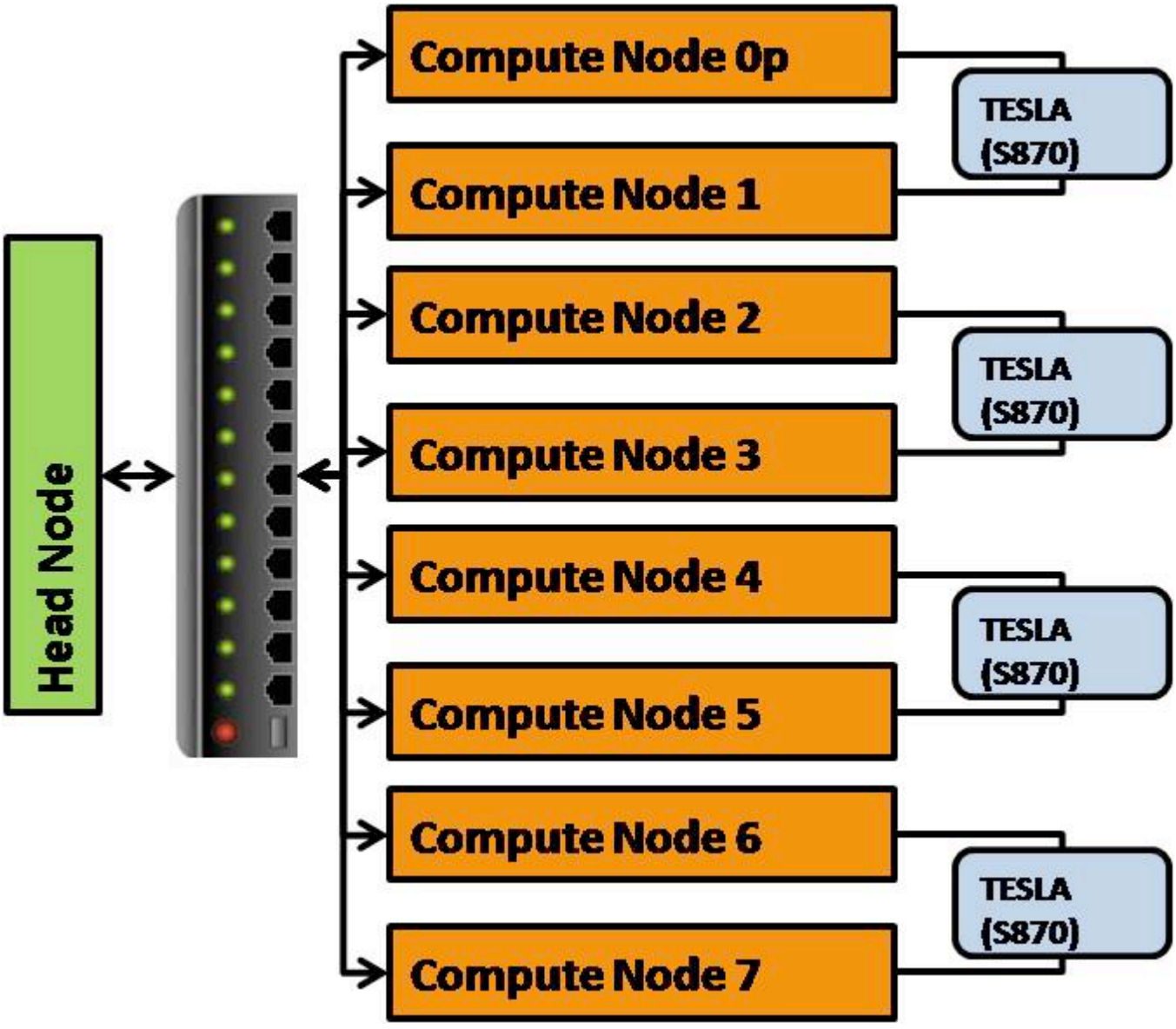
Goals: The goal of this project is to develop, and document, an optimized methodology for obtaining accurate force field parameters for a large class of bio-molecules such as: Span80, Phospholipids, bile salts (cholate, glycocholate, taurocholate), etc. The results obtained together with the codes and algorithms developed in this context will be archived and disseminated through publication for greater use by researchers working in this area. Various simulations are currently being performed using GROMACS 4.0 and Amber 10 on LONI computer clusters.

Effort requested: To help us achieve our goals, we request 4 months of FTE of LI scientist Bhupender Thakur. We should mention that Bhupender is familiar with the technical aspects and the software packages used for force-field development. His task would be to help us build an optimized methodology for obtaining reliable force field parameters, for the molecules of interest for our research that were previously mentioned. He will work together and provide technical expertise and guidance to Kumuditha Ratnayake a second year graduate student assigned to this project.

Benefits to LONI Institute and scope within LA-SiGMA: The development of the formalisms and algorithms for force field development will benefit multiple projects at LSU and Tulane, including the collaborative NSF funded project LA-SiGMA. We aim to further archive the results and provide the algorithms, to all LONI users upon validation. An online resource comprising a compute facility and a database is also being considered for a greater benefit to a larger community.

Next Generation GPU-Based Codes for Materials Discovery

Th. Bishop (Tulane) and D. Mobley (UNO)
M. Jarrell, J. Ramanujam, J. Moreno, B. Chen, and R. Hall (LSU)
S. Dua (LA Tech)
B. Thakur (LONI Institute)



Need Due to energy usage, the present generation of parallel computing cannot simply be scaled to the exascale. Nearly all proposals for beyond petascale computing involve heterogeneous computers. In most cases, this involves the use of the General Purpose Graphics Processing Units (GPGPU or GPU) generally running Cuda or OpenCL. Current national leadership class machines at the National Center for Computational Sciences include GPU clusters, and BlueWaters at NCSA will also likely be at least partially heterogeneous. Indeed, GPU acceleration is ideally suited for the Monte Carlo (MC) and Molecular Dynamics (MD) codes used by many investigators. Thus, in order to remain internationally competitive, it is essential for us to both develop GPU accelerated codes and to learn how to use the local and later the national GPU hardware to enable new discovery.

Rationale: The PIs are particularly well qualified to accomplish the goals of this request. CCT recently purchased a GPU Fermi cluster and the recently funded LI LA-SiGMA project (insitute.loni.org/lasigma) will purchase a larger GPU cluster. Efficient migration to and utilization of existing codes on next generation heterogeneous computers is a central component of LA-SiGMA. Drs. Jarrell and Hall have supervised the effort, with help from Dr. Thakur, of 6 undergraduates students over the past 5 months devoted to investigating the issues related to migrating MC codes to GPU machines. Dr. Bishop has strong ties with the creators of NAMD and is in position to evaluate different modes of operation of NAMD on GPUs. Dr. Mobley is involved with a molecular dynamics free energy code, "Yank", for binding free energy calculations on GPUs, and interacts closely with the developers of the OpenMM library which includes GPU support. LA-SiGMA is composed of 3 "Science Drivers" (correlated materials, energy materials, and biomolecular materials) which are closely coupled with a computational effort which includes migration to GPUs. Thus, involvement of Dr. Thakur in this effort will have an immediate and beneficial affect on the a large number of researchers within the LI. All of the investigators in this project are also LA-SiGMA investigators. LA-SiGMA will provide funding for graduate and undergraduate students and postdocs to work on this project.

Details and Required Effort. The PIs have existing homegrown MC codes and the source code for NAMD (which can use GPGPUs, but this feature is relatively untested) is available. Migration and testing of the MC and NAMD codes requires both direct coding and the organization of the effort of postdocs, graduate, and undergraduate students. Dr. Thakur will be responsible for coordinating this effort. Roughly 3 months of FTE will be required to develop and debug CUDA and OpenCL versions of the codes. Further 3 months will be required for continued optimizations and enhancements in the abstraction level of the source codes. Additional effort might entail configuring installations and developing patches for improving performance of canned codes on desired systems. In total, we request 6 month FTE aimed at creating portable, modular and sustainable software.

How the project will benefit the LI. This project is an essential part of the new LA-SiGMA collaboration, and is part of its mission to graduate LA researchers to the next generation of cyberinfrastructure. The project will benchmark existing codes and algorithms, and help discover tools which can benefit greatly from the use of such infrastructure. It will generate greatly improved publicly available codes in an important area of research for new devices, energy technology, and biomaterials. Utilizing existing CUDA enabled code, e.g. NAMD, will give the LI immediate practical experience in utilizing GPU hardware. Analysis of timings and results between CPU and GPCPU runs will help LI personnel develop their performance expectations/metrics and provide experience in numerical analysis.