HPC Day 2017 at UMass Dartmouth

UMass Dartmouth, 285 Old Westport Road, North Dartmouth, MA

Thursday, 25 May 2017

08:00-09:00 — Breakfast (Sponsored by Dell)

09:00-09:45 — **Keynote.**

Sushil Prasad (NSF)

Title: *Developing IEEE TCPP Parallel and Distributed Computing Curriculum and NSF Advanced Cyberinfrastructure Learning and Workforce Development Programs.*

Abstract: Parallel and distributed computing has become ubiquitous to the extent that even casual users depend on parallel processing. This necessitates that every programmer understands how parallelism and distributed programming affect problem solving. Thus, teaching only traditional, sequential programming is no longer adequate. In recognition of the importance of the issue coupled with its challenges, in 2012 the IEEE Technical Committee on Parallel Processing (TCPP) released its Curriculum on Parallel and Distributed Computing for undergraduates. This curriculum had significant influence on the ACM/IEEE CS2013 curricula. I will share my experience developing the TCPP curriculum and its related activities.

I will also briefly discuss National Science Foundation (NSF) programs related to Advanced Cyberinfrastructure's (OAC) multi-disciplinary learning and workforce development such as CAREER, CRII, and REU site programs, and the recent addition, the CyberTraining solicitation.

10:00 - 11:40 — Session I

Arghavan Louhghalam (U. Mass Dartmouth)

Title: *Quantitative Engineering Sustainability: Integration of Mechanics-based Models in Lifecycle Footprint*

Abstract: The sustainable development of our nations transportation infrastructure requires quantitative tools to translate pavements condition and design into lifecycle energy use and greenhouse gas (GHG) emissions. Recent developments in mechanistic pavement-vehicle interaction (PVI) models aim at providing such engineering estimates. I will discuss an integrated framework that involves models that range from material scale to structural scale all the way up to network scale to characterize the performance of pavement infrastructures. While the mechanics-based models within this framework establish the link between material characteristics and structural scale performance, the scale-up to network level estimates is accomplished using big data. I will also discuss the impact of aging in the form of evolution of material properties over time due to shrinkage, alkali silica reactions and other environmental stressors on durability and ultimately the life cycle environmental footprint of pavement infrastructure.

Ofer Cohen (U. Mass Lowell)

Title: *Computational Plasma Physics in the Solar System and Beyond*

Abstract: The universe is composed almost entirely of ionized gas a.k.a. plasma, where the dynamics of space plasmas is governed by the interaction between the plasma and magnetic fields in space. Developing computer models to study plasma dynamics is challenging since plasma dynamics involves extremely complicated physics, which includes many physical processes operating together, and a strong coupling between large and small scales. I will briefly describe the basics of computational plasma physics and its applications to solar and space physics, astrophysics, and extrasolar planets.

Giovanni Widmer (Tufts)

Title: *Visualizing complex bacterial populations in animal models*

Min Hyung Cho

Title: *Fast computational method for wave scattering*

Abstract: Many modern electronic/optical devices rely on wave such as solar cells, antennae, radar, and lasers. These devices are mostly built on a patterned layered structure. For optimizing and characterizing these devices, numerical simulations play a crucial role. In this talk, integral equation method in 2- and 3-D layered media Helmholtz equation will be presented. In 2-D, boundary integral equation with periodizing scheme is used. This method uses near- and far-field decomposition to avoid using the quasi-periodic Greens function. By construction, far-field contribution can be compressed using Schur complement with minimal computational cost. The new method solved the scattering from a 1000-layer with 300,000 unknown to 9-digit accuracy in 2.5 minutes on a workstation. In 3-D, a Lippmann-Schwinger type volume integral equation is used with layered media Greens function to include interface condition between layers and reduces the problem to only scatterers. In both 2- and 3-D layered media, a fast integral operator application is required because integral equation methods usually yield a full dense matrix system. A heterogenous fast multipole method (H-FMM) is developed. This is a hierarchical method and uses recursively-generated tree-structure. The interactions from far fields are compressed with free-space multipole expansion. All the spatially variant information are collected into the multipole-to-local translation operators. As a result, many free-space tools can be adapted directly without any modification to obtain an optimal O(N) algorithm for low frequency. This is a joint work with Jingfang Huang (UNC), Alex Barnett (Dartmouth College), Duan Chen, and Wei Cai (UNC Charlotte)

11:40 - 12:40 — Lunch break & posters (Sponsored by Microway)

12:40 - 02:20 — Session II

Marinos Vouvakis (U. Mass Amherst)

Title: *High-Performance Electromagnetic Computations: The Domain Decomposition Paradigm*

Abstract: Full-wave computational electromagnetic (CEM) kernels are integral part of modern high-frequency simulators and electronic design automation (EDA) tools for antennas, microwave/mm-wave/terahertz systems, radio frequency and mixed signals integrated circuits (ICs) and high-speed electronic packaging. Over the past decade, it has become clear that the ever increasing need for faster and higher fidelity/precision EDA tools can not rely solely on faster processors, but new computational algorithms that leverage the trends in many-core CPU and parallel computing architectures.

Domain Decomposition (DD) computational paradigms have a emerged as the leading approach to such distributed computations. They attempt to divide a large and/or complex, often intractable, computational domain into smaller suitably chosen regions, termed domains, that are easier solved independently in each core. The overall solution is then recovered by communicating portions of these solutions across processors often in an iterative fashion. Although this divide-and-conquer strategy offers distinct advantages, such as straightforward parallel programing and balanced usage of multi/many-core processing power, when difficult CEM models are encountered, no guarantees on the speed and success of the overall iterative convergence exist.

In this talk the audience will be introduced to the broader DD paradigm, and its most successful adaptations in the context of Finite Element Methods (FEM) for CEM. We will then identify the root causes of the most troubling reliability and scalability problems encountered in many iterative DD methods, and propose some solutions that have been explored by our research group. These solutions rely on advanced preconditioning schemes in the case of iterative DD methods, and a novel direct matrix factorization based solution that relies on DD. The main novelty of this work arises from the purely algebraic nature of both approaches which makes the transferable to other

computational disciplines, and result in high reliability at modest computational overhead. Results on challenging real-life antenna arrays, microwave engineering devices and scattering structure will presented. The talk will conclude with some personal reflections and the future of CEM computations.

Paul Whitford (Northeastern)
Title: *How HPC is transforming the study of biological machines*
Abstract: Continued growth of computing capacity has enabled the study of a range of complex biomolecular systems. In the mid-2000s, it was considered a heroic feat to simulate a large molecular assembly, such as the ribosome. In those early studies, dedicated use of world-class supercomputers was required to perform each simulation. Combined with the development of new theoretical models, modern HPC resources allow us to perform hundreds of simulations of these massive systems simultaneously. With such a dramatic increase in capability, it is possible to now address precise physical questions about these massive machines. In this talk, I will describe the trajectory of HPC applied to biomolecules, while highlighting how these efforts are providing insights into the physical principles that govern biology.

Randy Paffenroth (Worcester Polytechnic Institute)
Title: *Musings on Exabyte Scale Principal Component Analysis*
Abstract: Modern mathematical algorithms, when wed with high performance computing resources, allow one to treat large scale problems in Data Science. In particular, dimensionality reduction is one of the cornerstones of data analytics, and such ideas form the foundation for many other approaches commonly used in industry and academics. In this talk, we will describe how techniques such as low-rank matrix completion can be used to reduce the dimensionality of large scale data sets in an efficient and distributed fashion.

Samuel Isaacson (Boston University)
Title: *Jump Process Approximation of Particle-Based Stochastic Reaction-Diffusion Models*
Abstract: High resolution images of cells demonstrate the highly heterogeneous nature of the nuclear and cytosolic spaces. We are interested in understanding how this complex environment influences the dynamics of cellular processes. To investigate this question we have developed the convergent reaction-diffusion master equation (CRDME), a lattice particle-based stochastic reaction-diffusion method that can model the spatial transport and reactions of molecules within domains derived from imaging data. In this talk I will introduce the CRDME, and explain how it is similar in spirit to the popular reaction-diffusion master equation (RDME) model. The CRDME allows for the reuse of the many extensions of the RDME developed to facilitate modeling within biologically realistic domains, while eliminating one of the major challenges in using the RDME model.

02:20 - 03:10 — Education Roundtable
Title: *New Voices: Perspectives on Joining the HPC Community from students and members of non-traditional disciplines*
Abstract: High Performance Computing has been around for decades, primarily developed by computer scientists and engineers and used by physicists, mathematicians and engineers. As data volumes grow, it is clear that HPC resources and tools are useful to members of disciplines that have not traditionally used these systems, such as biology, chemistry, political science, economics. This years educational roundtable focuses on hearing from members new to the HPC community to better understand the application requirements and adoption challenges across a broader spectrum of research areas. Our goal is to use the new perspectives coming from this roundtable conversation as a starting point to develop strategies to address adoption challenges and expand our HPC education and training to support new researchers.

03:10 - 03:45 — Break (Sponsored by Dell)

03:345 - 5:00 — Session III

Ryan McKinnon (Massachusetts Institute of Technology)
Title: *Cosmological Simulations of Galaxy Formation*
Abstract: Galaxies in the universe form and grow over time in a complicated nonlinear fashion. Recent advances in supercomputing ability make it possible to numerically model the essential physics and evolve a "mock" universe from shortly after the Big Bang to the present day, producing a fairly realistic population of galaxies. In this talk, I will highlight the key topics in physics that govern galaxy formation, display visualizations from state-of-the-art astrophysics simulations, and discuss the supercomputing resources needed to simulate the universe.

Alfa Heryudono (U. Mass Dartmouth)
Title: *Polygonal Brain, Conformal Transplant, and Alzheimer's Disease*
Abstract: Systems of reaction-diffusion PDEs have been commonly used as one of many models to understand and simulate Alzheimer's disease in human brains. The PDEs, which simulate the spread of amyloid plaques, are numerically solved on irregular domain that mimics two-dimensional cross section of the brain. With a classical technique called conformal transplantation to map complex geometry to a simpler one, the simulation can be equivalently done on a unit disk domain using standard pseudospectral collocation solver. Numerical preliminary results are shown and couple of drawbacks of the method will be discussed. This is joint work with Tiffany Ferreira and Sigal Gottlieb.

Devesh Tiwari (Northeastern)
Title: *What is stopping us from getting to exascale computing and what should we do about it?*
Abstract: Continued increase in computing power and faster storage subsystems have enabled scientists to expedite the process of scientific discovery. As a result, emerging data-intensive workloads (e.g., scientific simulations, data analytics, etc.) are continuing to produce and analyze unprecedented amount of data on high performance computing systems. Over the past 50 years, we have gone from MFLOPS peak performance to PetaFLOPS peak performance. Unfortunately, it is likely that this trend will not continue on the same trajectory and will not lead us to the exascale era. In this talk, I will discuss what does it mean to build a "meaningful" exascale system? What are the new challenges in building an exascale system? How can we mitigate these challenges? And, finally, how can we continue to keep an exascale system "sustainable" and "meaningful" after it has been built?

05:00-05:45 — **Keynote.**

Luke Kelley (Harvard)
Title: *Predictions of future Gravitational Wave Observations using Simulations of the Universe*
Abstract: Last year, the LIGO collaboration announced the first direct detection of Gravitational Waves (GW) produced by the merger of two stellar-mass black holes. The computational techniques used to simulate these types of mergers, and the gravitational waveforms they emit, are remarkable and required for making future detections. In the next decade, astronomers will use the incredibly precise radio flashes from pulsars to measure GW from supermassive black hole binaries. I will discuss the state-of-the-art, cosmological hydrodynamic simulations (Illustris) which our group use to study the formation of galaxies, stars, and supermassive black holes in the universe. We use the results to make predictions of the GW that Pulsar Timing Arrays (PTA) will observe in the near future. In this talk, I will highlight some of the computational techniques we use in our models and simulations, in addition to presenting our most recent results on how long until PTA make their own GW detections, and what physical insights those detections will offer.

5:45 - 6:00 — Closing Remarks & Student poster prizes awarded (Prizes sponsored by Microway, Dell, SIAM, and MathWorks)