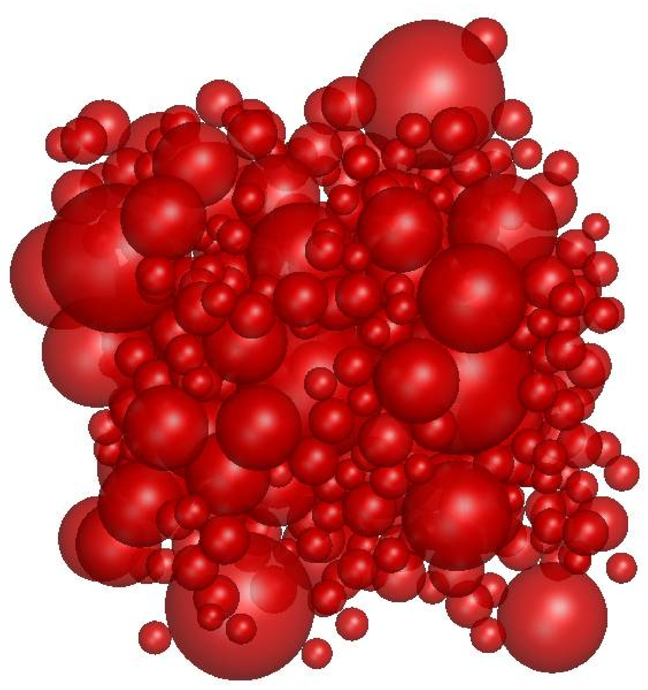
High-Performance Parallel Computing for Scientific Applications

Edmond Chow, Georgia Institute of Technology
School of Computational Science and Engineering
Atlanta, Georgia

Large-Scale Simulations of Macromolecules in the Cell

Proteins and other molecules are modeled by spheres of different radii. Stokesian dynamics is used to model both near-range and far-range hydrodynamic interactions. Scientific goal is to understand diffusion and transport mechanisms in the crowded environment of the cell.



Data from T. Ando

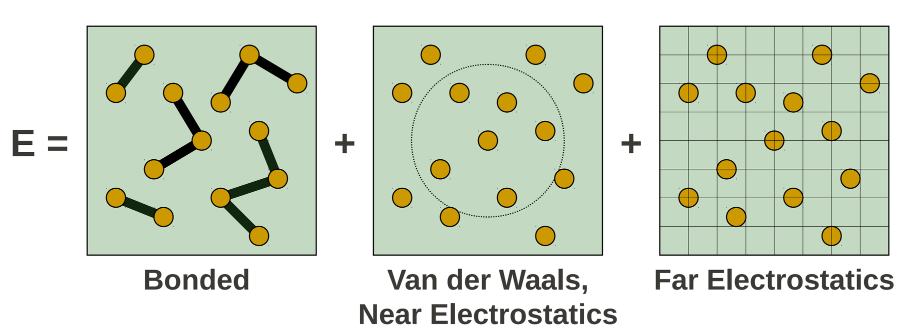
Computationally intensive components:

- Calculation of near-range interactions with size-dependent cutoffs
- Calculation of resistance matrices
- Linear solves with resistance matrices
- Calculation of Brownian force vector

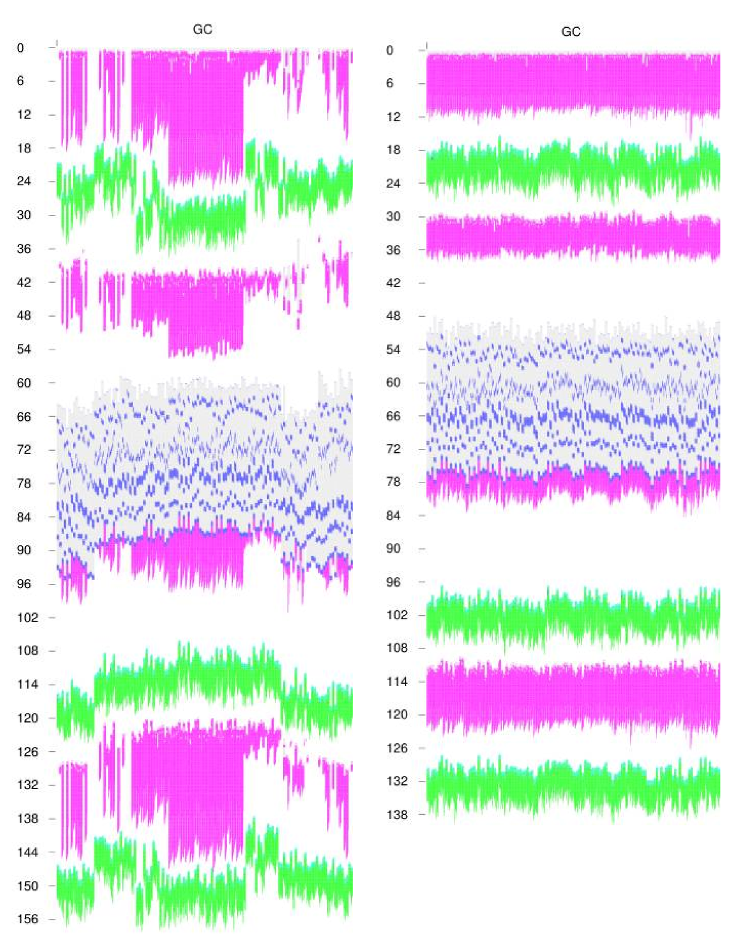
Technologies:

- Distributed-memory MPI parallelization
- GPU implementation
- High-level parallel programming abstractions and parallel computing infrastructures

Load-Balancing Heterogeneous Applications

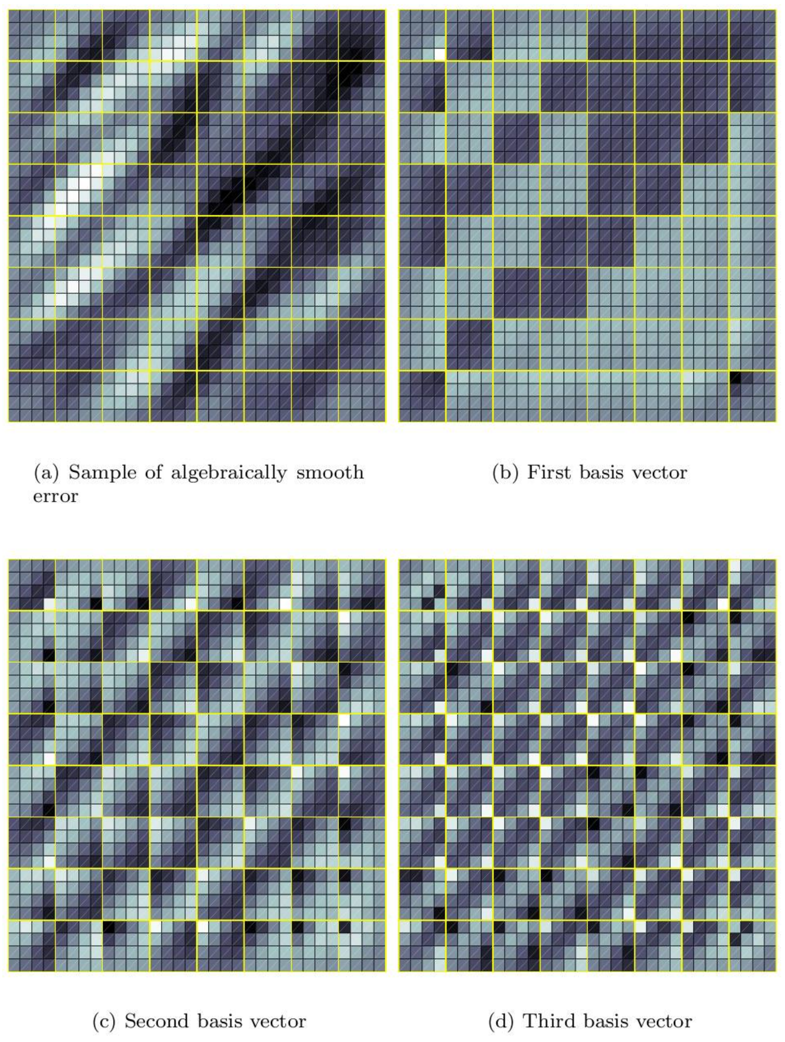


- The force calculation in molecular dynamics is composed of heterogeneous components, each requiring different communication patterns
  - Task-based parallelization with work stealing
- Particles and bonds may be distributed unevenly in space, but their geometric coordinates are known
  - Diffusion-based repartitioning



Activity graph, before and after load balancing; 2048 cores along x-axis; time along y-axis.

Multilevel Algorithms for Large-Scale Applications



(a) Sample of algebraically smooth error (b) First basis vector (c) Second basis vector (d) Third basis vector

Multilevel algorithms compute and combine solutions at different scales. They are necessary for solving many types of problems in a numerically *scalable* fashion.

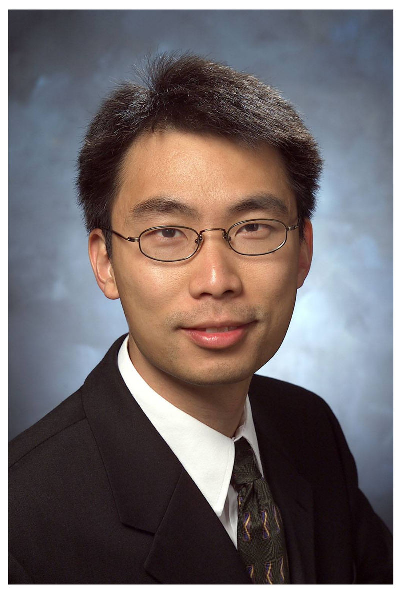
Large high-performance machines also have multilevel structure: memory hierarchies and parallel resources at the core, chip, node, and cluster levels.

Multiscale physical problems have different phenomena operating at different scales.

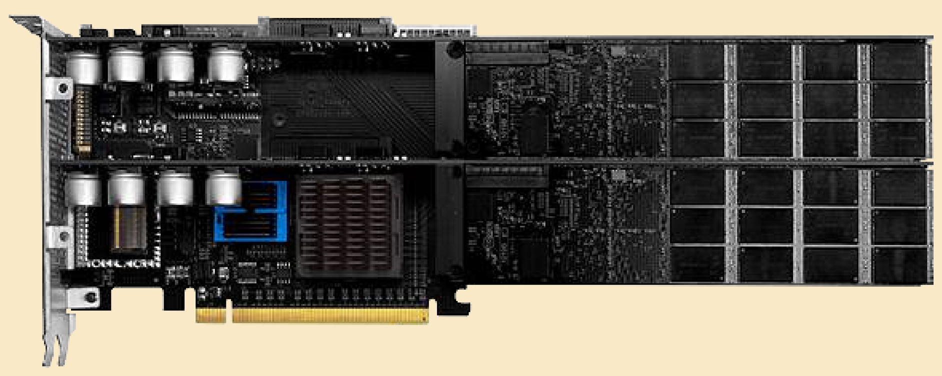
Goal is to link the structure of the physics to the structure of the algorithms and parallel computer to achieve high performance.

Background

- Associate Professor
  Georgia Institute of Technology, 2010-present
- Columbia University, 2009-2010
- D. E. Shaw Research, 2005-2010
- Lawrence Livermore National Laboratory, 1998-2005
- University of Minnesota, PhD 1998
- Contact: echow@cc.gatech.edu



Quantum Chemistry with Flash Memory Computing



Fusion-io ioDrive Octal, using PCI Express x16, up to 6 GB/s bandwidth, up to 5 TB capacity

Non-volatile memory devices are faster than disk and larger than DRAM and can accelerate many types of data-intensive applications.

Electronic structure calculations require huge amounts of intermediate memory for storing “two-electron” integrals: $O(N^4)$ for N basis functions.

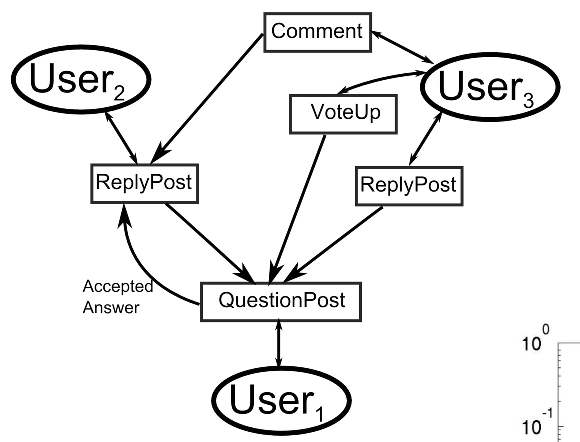
Many algorithms and codes must store these integrals on disk, rather than recompute these integrals “on-the-fly.”

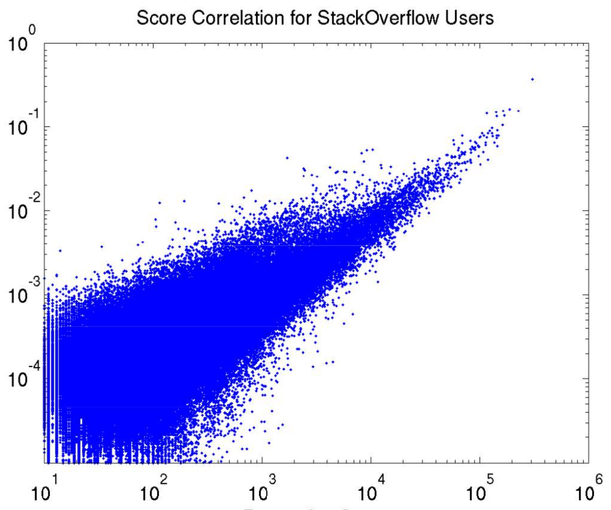
Challenge: understand application behavior and reformulate algorithms to exploit flash memory.

Data-Intensive Computing with Graphical Data

Small, important pieces of information may be hidden in vast amounts of graphical data.

We apply algorithms and machine architectures for analyzing unstructured data such as massive graphs.





A graph is constructed that describes the interactions between users on the StackOverflow web site. There is a strong correlation between the user Markov score (e.g, PageRank) and the user “reputation” score. Outliers in this plot signify users who may be “gaming” the system.

Additional Opportunities for Collaboration

- High-performance computing for uncertainty quantification.
- Accelerating dense matrix computations for quantum chemistry (tensor contractions).
- Software tools for high-performance computing.

Georgia Tech College of Computing
Computational Science and Engineering

China-US Software Workshop – Bilateral Collaboration for Software Development