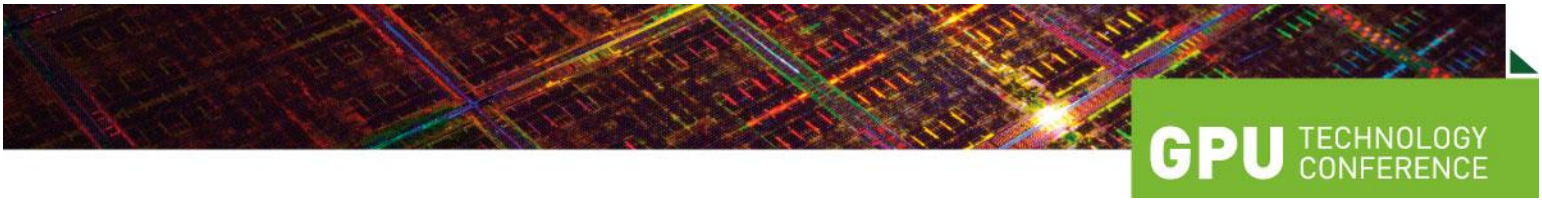


GPU Technology Conference, May 14-17, 2012
McEnery Convention Center, San Jose, California
www.gputechconf.com

Sessions on Life Sciences (subject to change)

IMPORTANT: Visit http://www.gputechconf.com/page/sessions.html for the most up-to-date schedule.

S0337 - High-Throughput Epistasis Screening Using GPUs
Mark Seligman (Insilicos LLC)
Day: Tuesday, 05/15 | **Time:** 9:00 am - 9:25 am

Topic Areas: Bioinformatics; Life Sciences; Supercomputing; Cloud Computing
Session Level: Intermediate

Epistasis is the interaction of two or more genes in coding for a biological property. Epistasis is believed to be an important factor in an individual's susceptibility to disease, and the search for epistasis is a major component in the development of personalized approaches to genomic medicine. Statistical tests for epistasis are typically confounded by the multiple-testing problem, that is, the aggregated loss of precision incurred through repeated hypothesis testing. One way to circumvent this problem is to simulate a false-discovery rate via resampling. We report success in using GPUs to accelerate these highly compute-intensive resampling techniques.

S0519 - GPU Accelerated Bioinformatics Research at BGI
Bing Qiang Wang (BGI)
Day: Tuesday, 05/15 | **Time:** 2:00 pm - 2:25 pm
Topic Areas: Bioinformatics; Life Sciences; Algorithms & Numerical Techniques; Supercomputing
Session Level: Intermediate

After digitizing DNA double helix by sequencing, computation is the key connecting raw sequences with life science discoveries. As massive data is generate, how to process and analysis as well as storage them in an efficiently manner turns out to be a major challenge. By developing GPU accelerated bioinformatics tools and integrate them into pipelines, BGI researchers now run analysis pipelines in several hours instead of several days. These tools include SOAP3 aligner, SNP calling and tool for population genomics. The speed up is generally around 10-50x comparing with traditional counterparts.

S0351 - Strong Scaling for Molecular Dynamics Applications
Sarah Tariq (NVIDIA)
Day: Tuesday, 05/15 | **Time:** 2:30 pm - 2:55 pm
Topic Areas: Molecular Dynamics; Cluster Management; Life Sciences
Session Level: Intermediate

In this session we will talk about how to improve strong scaling for molecular dynamics applications. Using the NAMD molecular dynamics code as our primary case study, we will discuss the types of issues that can impede scaling, how to use already available and custom tools to discover such issues, and how to build a model to help analyze and predict scaling performance. Although this session is primarily focused on molecular dynamics applications, most of the lessons can be applied equally well to many other areas and applications.

S0008 - Algorithms and Tools for Bioinformatics on GPUs
Bertil Schmidt (Nanyang Technological University)
Day: Tuesday, 05/15 | **Time:** 4:00 pm - 4:50 pm
Topic Areas: Bioinformatics; Life Sciences
Session Level: Intermediate

Learn how to use GPUs to accelerate compute- and data-intensive applications and algorithms Bioinformatics. High-throughput techniques for DNA sequencing and gene expression analysis with microarrays have led to a rapid growth in the amount of digital biological data, e.g. the NCBI Sequence Read Archive (SRA) houses raw sequence data generated by next-generation sequencing (NGS) technologies which succeeds 25 trillion base-pairs. Therefore, modern bioinformatics tools need to be scalable; i.e. they need to deal with an ever growing amount of data. GPUs and CUDA provide the opportunity to significantly reduce the runtime of many biological algorithms on inexpensive hardware.

S0638 - Lenovo ThinkStation Accelerates Medical Research with Beckman Coulter *(Presented by Lenovo)*
Scott Ruppert (Lenovo), Tanmay Dharmadhikari (Beckman-Coulter)
Day: Tuesday, 05/15 | **Time:** 4:00 pm - 4:50 pm
Topic Areas: Computer Graphics; Life Sciences
Session Level: Intermediate

Lenovo ThinkStations utilize NVIDIA Maximus technology to accelerate mission critical applications across multiple industries, including manufacturing, media & entertainment, and Life Sciences. Discover how GPUs are used to accelerate medical research from product experts with Lenovo and Beckman Coulter. Beckman Coulter has utilized Nvidia GPUs to reduce software development and test cycles by 50% with their Kaluza software. Kaluza is a revolutionary flow cytometry analysis software solution that provides visualization tools, speed and an innovative simplicity to the flow community. See how Kaluza allows users to analyze 10 million cells in real time. Session attendees will receive a drawing entry to win a brand new ThinkPad Tablet.

S0156 - Towards Computing the Cure for Cancer
Wu Feng (Virginia Tech)
Day: Tuesday, 05/15 | **Time:** 5:00 pm - 5:50 pm
Topic Areas: Bioinformatics; Life Sciences; Supercomputing; Algorithms & Numerical Techniques
Session Level: Intermediate

Learn about how to create "designer" genomic analysis pipelines as part of the "Compute the Cure" for cancer initiative from NVIDIA Foundation. Get an overview of an open-source framework that enables the creation of customized genomic analysis pipelines. Discover how different plug-ins from the "mapping/realignment/discovery" repositories, respectively, can be composed to form a genomic analysis pipeline. Learn to use next-generation sequencing data to characterize previously undetectable genetic changes between normal and malignant cells. Find out how you can contribute to the "Compute the Cure" cause.

S0072 - GPU-Enabled Spatiotemporal Model of Stochastic Cardiac Calcium Dynamics and Arrhythmias
M. Saleet Jafri (George Mason University), Hoang-Tron Minh Tuan (George Mason University)
Day: Wednesday, 05/16 | **Time:** 9:00 am - 9:50 am
Topic Areas: Life Sciences; Bioinformatics
Session Level: Beginner

Calcium ions play a central role controlling the contraction of the heart to pump blood. This requires tight regulation of cellular calcium dynamics which depends upon over 1,000,000 calcium channels that open and close

stochastically and have a very specific spatial arrangement. In the School of Systems Biology at George Mason University, CUDA technology coupled to novel algorithms for Monte Carlo simulation have made possible this computationally expensive spatiotemporal model of calcium dynamics in the heart muscle cell to study the regulation of calcium dynamics and what aberrations leads to cardiac arrhythmia.

S0010 - Towards Routine Microsecond Molecular Dynamics Simulations on Commodity Hardware
Ross Walker (University of California San Diego)
Day: Wednesday, 05/16 | **Time:** 9:00 am - 9:50 am
Topic Areas: Molecular Dynamics; Life Sciences
Session Level: Advanced

The original AMBER 11 provided performance on one GPU equivalent to an 8 node cluster and almost 60ns/day for 8 GPUs running the JAC production benchmark without additional approximations outstripping the performance of all conventional supercomputers. Here we describe further optimization of the code, coupled with hardware and software advances on the part of NVIDIA, that provides performance of >50ns/day on a single GPU with multiple GPUs providing simulation rates on systems the size of DHFR approaching a microsecond per day. This brings performance levels on desktops and commodity hybrid clusters to levels previously only considered possible using custom silicon.

S0262 - GPU-Accelerated Model-Based Drug Development
Chee Ng (Children Hospital of Philadelphia/University of Pennsylvania)
Day: Wednesday, 05/16 | **Time:** 10:00 am - 10:25 am
Topic Areas: Life Sciences; Algorithms & Numerical Techniques; Bioinformatics
Session Level: Beginner

Explore how GPUs can be used to improve the efficiency of drug development. Drug development is a very time-consuming, complex and expensive process that has low successful rate. A model-based drug development paradigm has been proposed as a possible solution to overcome these problems. A key challenge is to develop computational intensive drug and disease-specific models from a large quantity of highly complicated preclinical and clinical data. This session will describe how GPUs can and will play a key role in shortening the model development times and improving the efficiency of model-based drug development.

S0272 - GPU GWAS - CUDA Based Genome Wide Association Studies
Tim Bi (Johns Hopkins University / George Mason University)
Day: Wednesday, 05/16 | **Time:** 10:30 am - 10:55 am
Topic Areas: Life Sciences; Bioinformatics
Session Level: Intermediate

We have developed a CUDA based GWAS analyzer that has achieved a 10x analysis speed-up per GPU. Genome wide association studies scans through millions of SNP markers across the human genome seeking the genetic basis of life threatening diseases such as coronary artery disease and prostate cancer. The prospect of the $1,000 genome heralds a potential new scale of GWAS involving hundreds of thousands of patients. We will discuss how we utilized the Python, R, and C languages to produce a robust GWAS algorithm that can be extended to multiple GPUs and GPU clusters.

S0327 - Large and Sparse - Mass Spectrometry Data Processing in the GPU
Jose de Corral (Waters Corporation)
Day: Wednesday, 05/16 | **Time:** 2:00 pm - 2:50 pm
Topic Areas: Life Sciences; Bioinformatics
Session Level: Intermediate

Learn how the GPU helps identify millions of ions in datasets of several billion points of four-dimensional sparse data. The data is first reduced to 3D to locate regions of dense data, and then only those regions are processed in 4D. Processing involves combining several steps of convolution filters in three axes, finding local maximums in volumes of data, and extracting information from the data around each local maximum.

S0269 - Accelerating 3D-RISM Calculations using GPUs
Yutaka Maruyama (Institute for Molecular Science), Fumio Hirata (Institute for Molecular Science)
Day: Wednesday, 05/16 | **Time:** 3:00 pm - 3:25 pm
Topic Areas: Life Sciences; Algorithms & Numerical Techniques; Computational Physics
Session Level: Intermediate

The three-dimensional reference interaction site model (3D-RISM) theory, is a powerful tool to investigate biomolecular processes in solution. Unfortunately, 3D-RISM calculations are often both memory intensive and time-consuming. We sought to accelerate these calculations using GPUs. To work around the problem of limited memory size in GPUs, we modified the less memory-intensive Anderson method for faster convergence of 3D-RISM calculations. Using this method on C2070, we reduced the computational time by a factor of eight compared to Intel Xeon (8 cores, 3.33GHz) with the conventional method.

S0103 - Accelerating Protein Sequences and Classification using GPU-HMMER Search
Mahesh Khadtare (I2IT, Pune University)
Day: Wednesday, 05/16 | **Time:** 3:30 pm - 3:55 pm
Topic Areas: Life Sciences; Bioinformatics
Session Level: Intermediate

In this paper we present the results of parallelizing HMMer, which is a widely used tool for protein sequence homology detection, as well as functional annotation of homologous protein sequences, and protein family classification. The HMMer program is based upon a Viterbi algorithm coded in C, and is quite time consuming. We modify the Viterbi algorithmic logically to port it on GPGPU. We test multiple enhancements in our GPU kernels in order to demonstrate the effectiveness of each strategy. Our implementation cuda_hmmsearch achieves overall up to 30x speedup over intel single core CPU.

S0363 - Efficient Molecular Dynamics on Heterogeneous GPU Architectures in GROMACS
Szilárd Páll (KTH Royal Institute of Technology), Berk Hess (KTH Royal Institute of Technology)
Day: Wednesday, 05/16 | **Time:** 4:00 pm - 4:25 pm
Topic Areas: Molecular Dynamics; Computational Physics; Life Sciences
Session Level: Intermediate

Molecular Dynamics is an important application for GPU acceleration, but many algorithmic optimizations and features still rely on code that prefers traditional CPUs. It is only with the latest hardware and software we have been able to realize a heterogeneous GPU/CPU implementation and reach performance significantly beyond the state-of-the-art of hand-tuned CPU code in our GROMACS program. The sub-millisecond iteration time poses challenges on all levels of parallelization. Come and learn about our new atom-cluster pair interaction approach for non-bonded force evaluation that achieves 60% work-efficiency and other innovative solutions for

heterogeneous GPU systems.

S0516 - The Advantage of GPU Computation for Analyzing Complex Traits
Jun Zhu (Zhejiang University)
Day: Wednesday, 05/16 | **Time:** 4:00 pm - 4:50 pm
Topic Areas: Bioinformatics; Life Sciences
Session Level: Intermediate

Most import agriculture traits and human diseases are complex traits which are controlled by gene network with gene by gene interaction (epistasis) and gene by environment interaction (GE). New statistic methods and software are developed for analyzing genetic architecture for complex traits based on genome-wide association study (GWAS). When deal with large mapping population and huge amount of molecular information, GPU computation has an advantage over CPU computation. We will demonstrate the newly developed GPU based software QTLNetwork V3.0 and GWAS-GMDR for mapping genes with epistasis and GE interaction for complex traits of human, crops, and mouse.

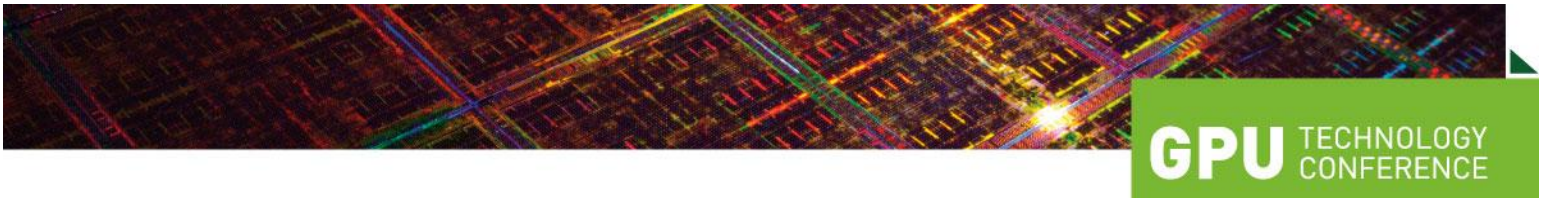
S0121 - Software Architecture to Facilitate CUDA Development
Peter Shenkin (Schrodinger), K. Patrick Lorton (Schrodinger)
Day: Wednesday, 05/16 | **Time:** 4:30 pm - 4:55 pm
Topic Areas: Development Tools & Libraries; Life Sciences
Session Level: Intermediate

We describe a workflow architecture and its use in developing Schrödinger's core-hopping application. The application supplies the stages as callbacks. A stage may have multiple implementations; for example, CUDA and CPU. An implementation can be assigned a maximum number of simultaneous threads. When any stage completes, a scheduling algorithm determines which implementation of which stage will be launched next. The application may detect "special" environments, such as CUDA, and set up its stages accordingly, or it may allow specification of which implementation of each stage to run. This makes it easy to develop and debug CUDA stages flexibly and incrementally.

S0382 - Hybrid System Architectures for High-Speed Processing in Optical Coherence Tomography
Brian Applegate (Texas A&M University Department of Biomedical Engineering), Brian Applegate (Texas A&M University Department of Biomedical Engineering)
Day: Thursday, 05/17 | **Time:** 10:00 am - 10:25 am
Topic Areas: Medical Imaging & Visualization; Life Sciences; Application Design & Porting Techniques; Development Tools & Libraries
Session Level: Intermediate

Several factors are spurring the development of hardware and software to accomplish high-speed processing for Optical Coherence Tomography (OCT), e.g. ultrahigh speed (>1 MHz) volumetric imaging and clinical applications (e.g. intravascular imaging). The computation power of GPUs ensures that it will be an essential part of the solution. We are exploring the development of a hybrid system in which the computational burden is shared between GPUs and other processors. This will make it possible to extract crucial diagnostic information in real or near real time. Technical challenges and recent progress will be discussed.

S0315 - Probing Bio-Nano Interface Structure from Microsecond Molecular Dynamics on GPUs
Olexandr Isayev (Case Western Reserve University)
Day: Thursday, 05/17 | **Time:** 10:00 am - 10:25 am
Topic Areas: Molecular Dynamics; Life Sciences
Session Level: Advanced



Using the latest algorithmic development in molecular dynamics on multiple GPUs over MPI, and technologies like GPUDirect it is now possible to address problems of interaction at bio-nano interface via large scale atomistic simulations. This talk will discuss the aspects of DNA-nanotube interactions and SWCNT induced conformational changes in DNA nucleosome structure. We will also address technical challenges upon porting and tuning AMBER 11 code on Condor GPU cluster at AFRL.