Conghui He

Phone: (+86) 15311775057 Email: heconghui@gmail.com Beijing, China

Education

Stanford University, visiting student researcher in Stanford Exploration Project (SEP) 7/2016 – present

Tsinghua University, Ph.D. Candidate in Department of Computer Science and Technology, High Performance Geoscience Computing Group 9/2013 – present

Sun Yat-sen University, Bachelor in Software Engineering 9/2010 – 6/2014

Social Experiences

Software engineer in National Supercomputing Center in Wuxi, China 2/2016 – present

Intern of Seismic Imaging R&D Group in Statoil (Beijing) Technology Service Co, Ltd. 7/2014 – 9/2014

President of IEEE Tsinghua Student Branch 9/2013 – 9/2014

Skills and Interests

Experience in leveraging HPC for geoscience applications

Skillful in parallel program profiling and optimization

High Performance Computing: pthread, OpenMP, MPI, CUDA, FPGA

Solid understanding of Linux system management, Bash scripting

General: C/C++, Python, Java, Linux, Bash scripting

Selected Awards

Schlumberger Scholarship for Computing Earth Science, 2014

First place in IEEE/IBM International Smarter Planet Challenge, 2013

Fourth Place in ISC12 International Super Computing Challenge, 2012

Sun Yat-sen's First Prize Student Scholarship, 2011

IBM Outstanding Student Scholarship, 2010

National Scholarship, 2009

Research/Project Experiences

My research interest includes Computational Geophysics and Parallel Algorithms. I'm experienced in the parallel algorithms design on modern computer architectures like GPU, multi-core CPU, and FPGA processors to solve the computational challenge raised from exploration geophysical applications. Participated projects include:

Refactoring and Optimizing the Community Atmosphere Model (CAM) on the New Sunway Manycore Supercomputer *7/2015 – 5/2016*

We refactor and optimize the Community Atmosphere Model (CAM) on the new Sunway many-core supercomputer of China, which is the rank 1 supercomputer in the latest Top 500 announcement. It uses a many-core processor that consists of management processing elements (MPEs) and clusters of computing processing elements (CPEs). To tackle the major challenges of mapping the large code base of CAM to the millions of cores on the Sunway system, we take OpenACC-based refactorization as the major tool, and apply source-to-source

translator tools to generate the most suitable parallelism for the CPE cluster, and to fit the intermediate variable into the limited on-chip fast buffer. For single kernels, when comparing the originally ported version using only MPEs and the refactorized version using both the MPE and CPE clusters, we achieve up to 22x speedup for the compute-intensive kernels. For the 25km resolution CAM global model, we manage to scale to 24,000 MPEs, and 1,536,000 CPEs and achieve a simulation speed of 2.81 model years per day.

Accelerate the Market Data Server on Reconfigurable platform *5/2015 – 6/2016*

Cooperative project with China Financial Futures Exchange (CFFEX), targeted at developing a high performance and low-latency market data server on a reconfigurable platform. The system, which is deployed on MAX4 platform, receives/sends trading packets from/to 10Gbs Ethernet on FPGA and process them directly on FPGA without any interference from CPU, thus gaining very low latency. Our work on FPGA includes parsing the packets in XTP protocol, maintaining a sorted order book and publishing the reconstructed market data. After careful design and optimizations, our design supports more than 200 instruments per FPGA and reach a latency of 3 us, reducing the latency by a factor of 33 compared to the existing software implementation.

Ensemble Full Wave Inversion with Source Encoding *9/2014 – 6/2015*

Full wave inversion (FWI) suffers from convergence toward local minima because of the inaccuracy of the initial model and the lack of low frequency data. Noises in seismograms further deteriorate the imaging quality. To relax the dependency on high-quality low-frequency data, we present an ensemble full waveform inversion method with source encoding (EnFWI), which is an ensemble approximation of the total inversion proposed by Tarantola. The method refines the velocity model iteratively by incorporating the observation, while the nonlinear evolution of the covariance is approximated by ensemble covariance. Encoded simultaneous-source FWI (ESSFWI) is applied to improve the representation for the low rank ensemble approximation, and to increase the rate of convergence. Experiments show that EnFWI achieves larger convergence range and better tolerance to data noise with less computational costs than traditional FWI methods.

A Fully-Pipelined Expectation Maximization Hardware Design for Gaussian Mixture Model 4/2014 – 11/2015

Gaussian Mixture Models (GMMs) are powerful tools for probability density modeling and soft clustering. They are widely used in data mining, signal processing, and computer vision. In many applications, we need to estimate the parameters of a GMM from data before working with it. This task can be handled by the Expectation-Maximization algorithm for Gaussian Mixture Models (EM-GMM), which is computationally demanding. In this project, we present our FPGA-based solution for the EM-GMM algorithm. We propose a pipelined-friendly EM-GMM algorithm, a variant of the original EM-GMM algorithm that can be converted to a fully-pipelined hardware architecture. To further improve the performance, we design a Gaussian probability density function evaluation unit that works with fixed-point arithmetic. We also come up with an idea to convert a large dimension or/and component EM configuration to multiple pieces of smaller dimension or/and component in case of resource limitation. In the experiment, one test case that spends more than 2 minuses on a single core CPU only spends hundreds of milliseconds in our data flow engine, enabling it to be a low-latency and high performance application. Our FPGA-based solution also generates accurate enough results while

achieving a maximum of 939 times speedup over a CPU-based solution, and 55 times speedup over a GPU-based solution.

A CPU-GPU Hybrid Parallel Design for Beam Migration *9/2013 - 12/2014*

The Kirchhoff beam-stack migration is quite popular in production with both better image quality and faster speed compared to Kirchhoff migration. Meanwhile, continuous HPC developments offer new opportunities for the industry to further enhance the efficiency of beam migration methods. We present a design of a highly efficient GPU-based beam migration. By parallelizing both the ray tracing and the beam mapping kernels with millions of GPU threads and using an asynchronous IO scheme, we derive a parallel beam migration design that fits current CPU-GPU hybrid clusters. Then, we test our GPU-based beam migration on the SEG/EAGE salt model and the SEAM salt model for different generations of GPU architectures, presenting accurate imaging results with 2-6 times speedup compared to a parallel 16-core CPU design.

Accelerating the Global Vegetation-Precipitation Correlation Algorithm *9/2013 - 11/2013*

Startup Project for Ph.D. candidate cooperated with a Professor in Remote Sensing field, aiming to accelerate the algorithm taking months to finish. Optimization strategies for it include modifying the algorithm to reduce I/O accessing by utilizing local buffer, adding a memory pool to reduce frequent memory allocation/destruction, overlapping I/O transferring and computing. It gained 20x speedup in the end.

Reverse Time Migration Implementations on Reconfigurable Platform *3/2013 - 6/2013*

Bachelor final year project, accelerating the Reverse Time Migration method on FPGA. The biggest computational challenge is the simulation of the acoustic wave equation on a big 3D-mesh. Several algorithms like finite difference method which conduct to a stencil operator were accelerated and researched.

Publications

Haohuan Fu, Conghui He, etc. “Refactoring and Optimizing the Community Atmosphere Model (CAM) on the New Sunway Manycore Supercomputer”. Under review

He, Conghui, Haohuan Fu, Ce Guo, Wayne Luk, Guangwen Yang. “A Fully-Pipelined Hardware Design for Gaussian Mixture Models”. Under review

He, Conghui, Haohuan Fu, Bangtian Liu, Huabin Ruan, Guangwen Yang, Hui Yang, and Are Osen. "A GPU-based Parallel Beam Migration Design." In 2015 SEG Annual Meeting. Society of Exploration Geophysicists, 2015.

He, C., Y. Chen, H. Fu, and G. Yang. "Ensemble Full Wave Inversion with Source Encoding." In 77th EAGE Conference and Exhibition 2015. 2015.

Clinton, Nicholas, Le Yu, Haohuan Fu, Conghui He, and Peng Gong. "Global-Scale Associations of Vegetation Phenology with Rainfall and Temperature at a High Spatio-Temporal Resolution." Remote Sensing 6, no. 8 (2014): 7320-7338.