Vasudevan Rengasamy

Personal Data

Address:	265 Blue Course Drive, Apt G3, State College, PA 16803, USA
Phone:	+1 (814) 3215454
email:	vxr162@psu.edu

About me

I am a PhD student at The Pennsylvania State University. My current research is in the area of High Performance Computing focusing on parallel algorithm development for scientific workflows and application optimization. I'm currently looking for full time employment opportunities.

Education

Aug 2014-Present

PhD in Computer Science and Engineering
Pennsylvania State University, USA
Advisor: Dr. Kamesh Madduri
Key Courses: Algorithms, Compiler Construction, Multiprocessor Architecture, Manycore Programming
Research work:

Optimizing Word2Vec: Word2Vec is a widely used word embedding technique which learns low dimensional vector representation for words in a training corpus. The original implementation of Word2Vec uses memory-bound vector products. Recently, Li et al. introduced minibatching approach to transform vector dot products into compute-bound matrix multiplications. My work further improved the floating point throughput while reducing overhead involved in dense matrix creation by processing multiple contexts together thus obtaining $1.3\times$ speedup over Li et al.'s implementation. Source code: https://github.com/vasupsu/IA3_Paper16_ArtifactEvaluation

Efficient Preprocessing for Metagenome Assembly: Metagenome samples consist of genome sequences of many different species. Metagenome assembly process can be parallelized by constructing read graph from metagenomic sequence data and assembling individual graph components concurrently. I have developed a workflow called MetaPrep for performing parallel and memory-efficient identification of connected components from read graphs constructed from metagenome sequences. Source code: https://github.com/vasupsu/MetaPrep

Accelerating Genetic Variant Detection: Genetic variant detection pipeline is compute and I/O intensive and takes few days to complete if executed serially. I have developed a tool SPRITE to optimize different stages of the pipeline by exploiting *intra-node and inter-node parallelism and avoiding intermediate I/O* by in-memory data processing. SPRITE has a novel counting based variant calling algorithm called PARSNIP which is very fast and has good accuracy for sequencing data with high coverage. Source code: sprite-psu.sourceforge.net

Gpa: 3.68/4

Aug 2011-Jun 2014	MSc(Engineering) in Computer Systems **Indian Institute of Science**, India Advisor: Dr. Sathish Vadhiyar Key Courses: Algorithms, Operating Systems, Computational Methods of Optimization Thesis: "A Message Driven Adaptive Framework for Irregular Applications on GPUs" Developed a framework G-Charm to improve the performance of irregular applications like molecular dynamics, N-body simulation etc. by effectively utilizing heterogeneous GPU and CPU cores. The project aims to overcome the challenges such as communication latency between CPU and GPU, resource idling etc. with minimum programmer overhead. Gpa: 6/8
Aug 2004-May 2008	B.Tech in Computer Science and Engineering **SASTRA University**, India Gpa: 9.62/10

Areas of Interest

High Performance Computing, Computational Biology, Operating Systems, Distributed Systems

Work Experience

May 2016-Aug 2016	Research Associate at HP Labs, Princeton, NJ. Worked towards identifying potential security related events in an enterprise network from anomalous user/system behavior. In the first part of this work, I created time-series models to model each user's normal behavior and then identified time slots corresponding to anomalous behavior. In the second part, I worked towards determining if there is a correlation between the detected anomalies and security incidents.
Oct 2008-June 2011	Systems Engineer at Tata Consultancy Services Limited, India Worked in mainframe applications development, managing a portion of the deposits system of a bank. The work involved source code analysis, monitoring daily batch runs and making enhancements to the existing systems as requested by business partners.
Aug 2007-Jan 2008	Intern at Klinik Kloster Paradiese, Soest, Germany Indexed nearly 2000 articles and magazines related to cancer research and developed a web interface and back-end to query the indexed data.

Other projects

Pintos	Modified Pintos kernel as part of Operating Systems course project. This involved implementing priority and MLFQ schedulers, adding system call interface for user programs, virtual memory management including frame allocation, clock page eviction algorithm, lazy page loading and file system management.
MORCO	Contributed to the development of MORCO framework which is a middleware for executing long-running multi-component applications with execution times much greater than execution time limits of batch queues.

Scholarships and Awards

Best Outgoing Student Award for the 2008 batch of B.Tech - Computer Science and Engineering of SASTRA University.

Merit scholarships for being among top 2% students in terms of academic performance in the years 2004, 2005.

Publications

Nov 2017	Vasudevan Rengasamy, Tao-Yang Fu, Wang-Chien Lee, Kamesh Madduri, **Optimizing Word2Vec Performance on Multicore Systems**, Accepted IA^3 workshop 2017
June 2017	Vasudevan Rengasamy, Paul Medvedev, Kamesh Madduri, **Parallel and Memory-Efficient Preprocessing for Metagenome Assembly**, HiCOMB 2017
June 2016	Vasudevan Rengasamy, Kamesh Madduri, **SPRITE: A fast parallel SNP detection pipeline**. In Proceedings. ISC High Performance, 2016
Oct 2015	K. Madduri, V. Rengasamy, and P. Medvedev, **SPRITE: A fast parallel SNP detection pipeline** poster presentation at the American Society of Human Genetics (ASHG) Annual Meeting, Oct. 2015
June 2013	R. Vasudevan, Sathish S. Vadhiyar, and Laxmikant V. Kal*é*. **G-Charm: an adaptive runtime system for message-driven parallel applications on hybrid systems**. In Proceedings of the 27th international ACM conference on International conference on supercomputing (ICS '13)

Programming Skills

C, C++, Java, MPI, openMP, CUDA, Charm++, R, Python