

Saurabh V. Pendse

Computer Scientist
Software Developer

1000 Kiely Blvd, Apt. 52

Santa Clara, CA - 95051

(774) 232 4060

✉ svpendse@ncsu.edu

🌐 <http://www.saurabhpendse.com>

Education

- Aug '11–May '13 **M.S. in Computer Science**, *North Carolina State University*, Raleigh, USA, GPA 4.0/4.0.
○ Relevant courses: Operating Systems, Database Management Systems, Algorithms, Automated Learning and Data Analysis, Performance Modeling, Automata Theory, Software Engineering, Internet Protocols, Artificial Intelligence II, Management, Decision and Control Systems.
- Jul '07–Jun '11 **Bachelor of Engineering (B.E.) in Computer Science**, *The Maharaja Sayajirao University*, Baroda, Gujarat, India, GPA 4.0/4.0, Gold Medalist.

Work Experience

- Jun '13–Pre **Software Engineer**, *Apple Inc.*, Cupertino, CA, USA.
- May '12–Aug '12 **Summer Intern**, *iOS, Apple Inc.*, Cupertino, CA, USA.
○ Worked on production-level Platform services, Information Retrieval systems, and Internal APIs for the iOS 6 release. My work involved optimizing inverted index compression and query processing on data interchange formats like Protocol Buffers, XML and JSON.
- Aug '11–December '12 **Research Assistant**, *North Carolina State University*, Raleigh, NC, USA.
○ Designed and implemented algorithms and enabling technologies for peta-scale data analytics, performance modeling, causality estimation in large-scale temporal and spatio-temporal data, high performance computing, data compression and query processing.
○ Worked on the **Cray Titan** supercomputer at **ORNL**, towards developing enabling middleware technologies for scientific simulation throughput optimization, parallel in-situ indexing of simulation data, failure prediction, root cause analysis and data assimilation on the **IBM Blue Gene/P** supercomputer at **ANL**.

Computer skills

- Languages, C, C++, Java, Python, Obj. C, Cocoa, Assembly, Prolog; CUDA, MPI, OpenMP, STL, Boost, APIs
GitHub Enterprise, Subversion, Maven, NumPy.
- OS Mac OS, Linux, Windows.
- IDEs Xcode, Eclipse, Visual Studio, Netbeans.
- Build GNU & IBM XL Compilers, CMake, Jenkins.
- Networking Socket Programming, OPNET, ns-2, Wireshark.
- Software MATLAB, R, Mathematica, GNU Octave, L^AT_EX, OpenOffice, MS Office, iWork, Weka.
- Data, I/O Protocol Buffers, XML, JSON, Avro, ADIOS, HDF5, NetCDF, Hadoop FS.
- Databases Oracle, (PL/SQL), IBM DB2, MongoDB, MySQL.
- Web Dev. HTML, CSS, PHP, JSP, JavaScript, jQuery, AJAX, Drupal.

Selected Publications

- HPDC 2013 S. Lakshminarasimhan and D. A. Boyuka II, **S. V. Pendse**, X. Zou, J. Jenkins, V. Vishwanath, M. E. Papka, N. F. Samatova. Scalable In Situ Scientific Data Encoding for Analytical Query Processing. *The 22nd International ACM Symposium on High Performance Parallel and Distributed Computing* [pdf],[Best paper award]

- HPDC 2012 **S. V. Pendse**, E. R. Schendel, J. Jenkins, D. A. Boyuka II, Z. Gong, S. Lakshminarasimhan, Q. Liu, S. Klasky, R. Ross, N. F. Samatova. ISOBAR Hybrid Compression-I/O Interleaving for Large-scale Parallel I/O Optimization. *The 21st International ACM Symposium on High Performance Parallel and Distributed Computing* [pdf] (Nominated for best paper)
- SIAM SDM 2012 **S. V. Pendse**, I. K. Tetteh, F. H. M Semazzi, V. Kumar, N. F. Samatova. Toward Data-driven, Semi-automatic Inference of Phenomenological Physical Models: Application to Eastern Sahel Rainfall. *2012 SIAM International Conference on Data Mining* [pdf],[Talk slides]
- CLUSTER 2012 N. Shah, E.R.Schendel, S. Lakshminarasimhan, **S. V. Pendse**, T. Rogers, N. F. Samatova. Improving I/O Throughput with PRIMACY: Preconditioning ID-Mapper for Compressing Incompressibility. *2012 IEEE International Conference on Cluster Computing* [DOI]

Selected Software Projects (Details available on website)

- December 2012 **Scalable In-situ Scientific Encoding for Analytical Query Processing.**
Developed a parallel in-situ data indexing technique on the Blue Gene/P supercomputer that enables transformation of scientific simulation output into a query-efficient form, with negligible overhead to the simulation run. Implemented a novel in-network merging strategy, together with highly optimized index compression and machine learning based system tuning to achieve up to 10x speed up in query response versus alternative state-of-the-art techniques.
- April 2012 **Stacked Denoising Autoencoders and Deep Belief Networks.**
Implemented from scratch the seminal deep learning structures in Matlab and Python for learning useful representations from real-world temporal and spatio-temporal datasets, applicable to a variety of classification and prediction problems.
- February 2012 **Hybrid Compression-I/O Interleaving for Large-scale Parallel Systems.**
Developed a framework for compression based Compute-I/O process interleaving on the Titan and Blue Gene/P supercomputers, and integrated in the production S3D simulations.
- January–April 2012 **Thread library, Buffer Overflows, Bag of tasks problems, Distributed Systems.**
Developed a complete user-level thread library analogous to POSIX threads, studied buffer overflow attacks and related system vulnerabilities and developed a sandbox-based approach for mitigating them; a CPU-GPU framework using CUDA for flavors of the bag of tasks problems, designed a fault-tolerant voting protocol on a distributed client-server architecture.
- October 2011 **Causality Estimation Framework.**
Developed a causality estimation framework based on LASSO regression methods in MATLAB for inference of phenomenological models from temporal and spatio-temporal climate datasets.

Academic Achievements

- 2013 Inducted into the **Phi Kappa Phi Honor Society**.
- 2012 Peer reviewer for the **Cluster Computing** Journal.
- 2012 SIAM student travel award for **SDM'12**, Anaheim, CA.
- 2011 Cleared the **PhD Written Prelim Exam** in December, 2011.
- 2011 Graduate Teaching Assistantship awarded from NC State University.
- 2011 Awarded Gold Medal from MS University for academic excellence.
- 2010 Secured **99.19%ile** in CAT 2010, and admits at 3 Indian Institutes of Management (**IIM**).
- 2008–2010 Winner of Codewrath 2009, 2010, Trinity 2009, Bytecode 2010, programming events held at **Footprints** (a national level technical fiesta (India)).
- 2007–2008 Selected in the Extended Merit List of the **IIT** (Indian Institute of Technology) Examination in 2007 (out of 3,00,000 students) and the Computer Science program at **BITS**, Pilani, 2008.
- 2007 Amul Vidhyabhusan, Swami Vivekanand Memorial awards for securing **2nd** rank (Science subjects) in Board Examinations.
- 2004 **Ramanujan** all India Mathematics Examination - Stood **8th** in Gujarat State in 2004.