

Babak Behzad, Ph.D.

CONTACT INFORMATION	Cell Phone: (217) 419 1834 E-mail: babakbehzad@gmail.com Website: http://www.babakbehzad.com/
POSITION	Senior Software Engineer at SAP, through acquisition of Altiscale Inc.
PROFESSIONAL INTERESTS	Big Data, Deep Learning, Software Engineering
EDUCATION	University of Illinois at Urbana-Champaign , Urbana, Illinois USA Ph.D. Computer Science, August, 2015 Thesis: <i>Optimizing Parallel I/O of HPC Applications</i> Advised by Marc Snir University of Illinois at Urbana-Champaign , Urbana, Illinois USA M.S. Computer Science, May, 2013 Thesis: <i>Autotuned optimized parallel I/O for GIScience and spatial applications</i> Advised by Marc Snir and Shaowen Wang Sharif University of Technology , Tehran, Iran. B.S. Computer Science, May, 2010
WORK EXPERIENCE	SAP, Palo Alto, CA <i>Senior Software Engineer</i> September 2016 - Present On September 27, 2016, SAP officially announced its completed acquisition of Altiscale, which provides a high-performance, scalable Big Data-as-a-Service (BDaaS) solution that uniquely includes full operational services. Altiscale Inc., Palo Alto, CA <i>Performance Lead</i> August 2015 - Present Ensure Altiscale is constantly providing high-performance Big Data services to customers. Development of an infrastructure for a fully multi-tenant cluster using Docker containers for Hadoop. Altiscale Inc., Palo Alto, CA <i>Summer Intern</i> May 2014 - August 2014 Design and implementation of a new architecture for the Data Pipeline. The design using Kafka proposed and implemented in this project for its scalability and flexibility. National Center for Atmospheric Research (NCAR), Boulder, CO <i>Research Intern</i> May 2013 - August 2013 Evaluation and optimization of <i>PIO Library</i> . <i>PIO Library</i> is a parallel I/O library designed to be used in different climate simulation applications in order to increase I/O performance. The HDF Group, Champaign, IL <i>Research Intern</i> May 2012 - August 2012 Design and implementation of <i>A framework for Autotuning Parallel I/O Parameters for HPC Applications</i> . This framework uses heuristics to optimize for I/O layers and their parameters without requiring to change anything in the scientific application.

Developing a fully-functional OCR system for Persian with Tensorflow

(Personal Interest)

- A fully functional CNN for recognizing the most popular Persian words.
- A fully functional RNN model with CTC for recognizing characters and the context.
- An Inference system based on Tensorflow serving running as a cloud API.

A Truly Elastic Big Data Infrastructure using Containers

(with Altiscale Inc.)

- An abstraction for a host that enables elastic, cloud-based big data analytics.
- Scheduler responsible for moving compute containers between clusters within a multi-tenant data center.
- The impact of this scheduler on performance, utilization, and cost reduction within our data centers.

Autotuning for High-Performance I/O

(with LBL and Marc Snir)

- Developed an autotuning framework for parallel I/O helping scientists to obtain up to 100X speedup in I/O.
- Showed the applicability of this framework on all the leading-edge HPC platforms for different applications.

PGAS Programming Languages for Irregular Applications

(with ANL and Marc Snir)

- Identified and solved current performance bottlenecks of PGAS languages for irregular applications.
- Applied this scheme to multicore clusters for obtaining better performance.

High-Performance Computing for Geospatial Sciences

(with Shaowen Wang and Marc Snir)

- Parallelizing different geospatial sciences applications such as map reprojection.
- Explored different high-performance solutions for geospatial sciences such as cloud computing.

Power and Performance Studies of Applications on Different Platforms

(with Josep Torrellas)

- Detailed study of the performance and power consumption of different applications on Intel SCC.
- Comparison of Intel SCC with Intel Core i7, Intel Atom, and NVIDIA ION2.

High-Performance Cloud Computing for Big Data Platforms

(with Indy Gupta)

- Detailed analysis of metrics and events of Hadoop clusters for improving performance.
- Evaluation and analysis of I/O intensive HPC applications on Amazon cloud.

- ◇ Multi-Tenant Elastic Big Data Infrastructure Utility Patent No. 15/352,250. filed on Nov. 15, 2016.
- ◇ Multi-Tenant Elastic Big Data Infrastructure Provisional Patent No. 62/353,898 filed on Jun. 23, 2016.

SELECTED
PUBLICATIONS

Journal

- ◇ Junchao Zhang, [Babak Behzad](#), and Marc Snir. “Design of a Multithreaded Barnes-Hut Algorithm for Multicore Clusters”, *IEEE Transactions on Parallel and Distributed Systems*
- ◇ Eric Shook, Michael E Hodgson, Shaowen Wang, [Babak Behzad](#), Kiumars Soltani, April Hiscox, Jayakrishnan Ajayakumar. “Parallel cartographic modeling: a methodology for parallelizing spatial data processing”. *International Journal of Geographical Information Science*

Conference

- ◇ [Babak Behzad](#), Hoang-Vu Dang, Farah Hariri, Weizhe Zhang, and Marc Snir. “Automatic Generation of I/O Kernels for HPC Applications”, Nov. 2014, *Parallel Data Storage Workshop 2014 (PDSW 2014)*
- ◇ [Babak Behzad](#), Surendra Byna, Stefan Wild, Prabhat, and Marc Snir. “Improving Parallel I/O Autotuning with Performance Modeling”, Jun. 2014, *ACM/IEEE High Performance and Distributed Computing (HPDC 2014)*
- ◇ [Babak Behzad](#), Huong Vu Thanh Luu, Joseph Huchette, Surendra Byna, Prabhat, Ruth Aydt, Quincey Koziol, and Marc Snir. “Taming Parallel I/O Complexity with Auto-Tuning”, Nov. 2013, *ACM/IEEE Supercomputing (SC 2013)*
- ◇ Junchao Zhang, [Babak Behzad](#), Marc Snir. “Design of a Multithreaded Barnes-Hut Algorithm for Multicore Clusters”, Nov. 2013, *Poster Session. ACM/IEEE Supercomputing (SC 2013)*
- ◇ [Babak Behzad](#), Indranil Gupta. “Measurement and Analysis of I/O Intensive Workloads on Cloud Computing Platforms”, Nov. 2013, *Poster Session. 2013 Amazon PhD Symposium.*
- ◇ [Babak Behzad](#), Huong Vu Thanh Luu, Ruth Aydt, Marianne Winslett. “A Multi-Level Approach for Understanding I/O Activity in HPC Applications”, Oct. 2013, *Workshop on Interfaces and Abstractions for Scientific Data Storage 2013 (IASDS 2013)*
- ◇ Yan Liu, Michael P. Finn, [Babak Behzad](#), and Eric Shook. “High-Resolution National Elevation Dataset: Opportunities and Challenges for High-Performance Spatial Analytics”. Abstract accepted for presentation in the Special Session on “Big Data” Mar. 2013, *American Society for Photogrammetry and Remotes Sensing Annual Conference*
- ◇ [Babak Behzad](#), Yan Liu, Eric Shook, Michael P. Finn, David M. Mattli, Shaowen Wang. A Performance Profiling Strategy for High-Performance Map Re-Projection of Coarse-Scale Spatial Raster Data, Sep. 2012, *The International Symposium on Automated Cartography (AutoCarto 2012)*
- ◇ Ehsan Tottoni, [Babak Behzad](#), Swapnil Ghike, Josep Torrellas. Comparing the Power and Performance of Intel’s SCC to State-of-the-Art CPUs and GPUs, Dec. 2011, *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS 2012)*

PROFESSIONAL
SERVICES

- ◇ Journal reviewing: IEEE TPDS Journal, Cluster Computing Journal
- ◇ Conference reviewing: Mass Storage Systems and Technologies (MSST), International Conference on Supercomputing (ICS), International Workshop on Data-Intensive

Scalable Computing Systems (DISCS)
◇ Graduate student application reviewing: Computer Science Dept. of UIUC

TECHNICAL
SKILLS

Programming Languages: C/C++, Java, Python, Ruby, R, JavaScript

INTERESTS

Robotics, volleyball, dancing, coffee, running