

Milo Polte

762 Shotwell St
San Francisco, CA 94110
Phone: 607-280-3615

milo.polte@gmail.com
github.com/sqrl
www.linkedin.com/in/milopolte

Work Experience

- 2012-2015 *Software Engineer and Manager/Tech Lead for [WibiData, Inc](#) (San Francisco).*
- Fourth employee at a Big Data startup that grew to over forty.
 - Survived one name change and three changes to the logo.
 - Developed a platform for personalized applications on top of Hadoop, HBase, and Cassandra.
 - Our platform was used in production to deliver real-time product recommendations at two large retail customers, including holiday traffic.
 - Project/tech lead on the effort to open source our software stack as the [Kiji Project](#).
 - Broke down the effort into manageable, prioritized tasks for the team, delivering tested, first releases on schedule.
 - Kiji permitted us to better engage with our community with our own meet-ups, etc.
 - Tech lead on our earliest real time scoring component, [Kiji Scoring](#) that served our first real time recommendations.
 - Transitioned to a dual technical-managerial role and managed ten direct reports during my tenure (a maximum of five at one time).
 - Tech Lead/Manager of the platform team.
 - Responsible for the foundational layers of our stack.
 - Developed and help design internal build tools and scripts that reduced compile time and improved engineer productivity.
- 2011 (3 months) *Software Engineer internship at [Panasas](#) (Pittsburgh, PA).*
- Project work towards instrumenting a FreeBSD kernel driver to implement QoS I/O scheduling in network attached storage devices for our parallel, distributed filesystem.

Technologies and Skills

- Python My most recent projects have been in Python.
- Authored devtools and simple but helpful build components, e.g. [packrat](#) which we used internally at WibiData to cache and share build artifacts.
- Java Experienced coding distributed applications and map-reduce programs in Java. I've also used Scala but am less familiar with it.
- Hadoop Stack I've used a variety of technologies in the Hadoop ecosystem for model training and analytics: MapReduce, Yarn, HBase, Hive, Scalding, etc.
- Ported our ecosystem to different versions of the Cloudera Distribution of Hadoop and others, dealing with the bugs and changes along the way.
 - Debugged numerous grungy issues in HBase ecosystem, such as ZooKeeper references, unblocking coworkers.
- SCRUM and hygienic coding I was one of the first scrum masters at WibiData. Everyone does it differently.
- Led Sprint Planning for platform team, working with Product to prioritize tasks.
 - Experienced in a culture of code reviews, code style guides, etc.
 - Owned the effort on writing our Python style guide.

Education

2006-2011

Masters Degree in Computer Science with partial work towards a PhD at Carnegie Mellon University (Pittsburgh, PA)

Advisor: Garth Gibson.

- Conducted research in CMU's Parallel Data Lab group under the programs of the [Petascale Data Storage Institute](#) and the [Institute For Reliable High Performance Information Technology](#).
- Benchmarked Flash-based SSDs and evaluated cost efficiency of different hybrid configurations. Papers: <http://www.pdl.cmu.edu/PDL-FTP/PDSI/simsa-pdsw08.pdf> and <http://www.pdl.cmu.edu/PDL-FTP/PDSI/Polte.pdf>. Included work inside the Linux kernel.
- Collaborated with Los Alamos National Laboratory in the design and implementation of the [Parallel Log-structured File System](#). Papers: <http://institutes.lanl.gov/plfs/plfs.pdf> and <http://www.pdsi-scidac.org/events/PDSW09/resources/pdsw09-final9.pdf>
- Investigated the design of parallel filesystems for HPC built on top of BigTable-like software. Paper: <http://www.cs.cmu.edu/~svp/2009hotcloud-tablefs.pdf>

1999-2005

Bachelor of Arts and Master of Engineering in Computer Science at Cornell University(Ithaca, NY)

Emphasis on courses in Distributed Systems, Networks, and Theory.

- Independent Research with Prof. Emin Gun Sirer - (2001- 2002) Researched and developed an anonymizing peer-to-peer overlay network based on dining cryptographer nets called 'Herbivore'. White paper available: <http://www.cs.cornell.edu/People/egs/herbivore>
- Cornell University Neurobiology Department - (2005) Software engineer for audio experimentation and education cross platform workbench program called "Koé", providing functionality similar to professional synthesis applications, such as Reaktor.

Hobbies and Other Activities

- [Hiking](#)
- [Vegetarian Cooking](#)
- Losing at [Dota](#)
- My pet [rabbit](#)