

Milo Polte

762 Shotwell St
San Francisco, CA 94110
Phone: 607-280-3615
www.linkedin.com/in/milopolte

milo.polte@gmail.com
github.com/sqrl (mostly professional)
github.com/hyena (mostly hobbies)

Work Experience

May 2015 - Present *Software Engineer at [Dropbox](#) in the Data Infrastructure Team* (San Francisco).

- Tech lead on new data-infra APIs, replacing single point of failure web-forms with programmatic interfaces providing the backends for a new suite of analytic tools.
 - Query Service: A RESTful API for Hive and Presto queries.
 - ETL Service: A RESTful API for safely managing job pipelines.
- Member of an on-call rotation maintaining an infrastructure that runs thousands of jobs crunching petabytes of log data.

2012 - Feb 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

I've used a variety of technologies and programming languages, but here I highlight a few and describe what I've used them for in work and play. For any language or technology I probably have "strong opinions weakly held" about what's great and awful about them.

Python	<p>I've used python at just about every company for scripting as well as systems coding.</p> <ul style="list-style-type: none">• Authored devtools and simple but helpful build components, e.g. packrat which we used internally at WibiData to cache and share build artifacts.• Worked on large Python-based backend systems such as Dropbox's Python based blockstore service and our 'Drone' job scheduler for Hadoop jobs.
Java	<p>I am experienced in coding distributed applications and map-reduce programs in Java. For example, it's what I use primarily at Dropbox for our new Data APIs written as DropWizard applications.</p>
Hadoop Stack	<p>I've used a variety of technologies in the Hadoop ecosystem for model training and analytics: MapReduce, Yarn, HBase, Hive, Scalding, etc.</p> <ul style="list-style-type: none">• Ported Wibi's 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.• Put in yet more grunge-work hours on Dropbox's giant Hive cluster, migrating namenodes on a live cluster, debugging subtle Zookeeper locks, etc.
Scrum and hygienic coding	<p>I was one of the first scrum masters at WibiData. Everyone does it differently.</p> <ul style="list-style-type: none">• Led Sprint Planning for Wibi's platform team, working with Product to prioritize tasks.• Established much of our culture of code reviews, tooling, etc.• Owned the effort on writing our Python style guide.
Go-lang and Rust	<p>Still learning these technologies, but I'm really excited about their potential. * Currently re-implementing Dropbox's legacy and feature-creeped blockstore with a Go-based solution for reasons of performance, safety, and cleanliness. * Coded a number of silly slack bots to practice Go and Rust: * https://github.com/sqrl/slack-quoter Bot to save and replay humorous slack quotes * https://github.com/hyena/professor_sloak Bot that gives people random pokemon on request, fortune cookie style. * https://github.com/sqrl/slack-logger Super simple bot that logs slack conversations</p>

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](http://institutes.lanl.gov/plfs/plfs.pdf). 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: <http://www.cs.cornell.edu/People/egs/herbivore>
- Cornell University Neurobiology Department - (2005) Worked as a software engineer on an audio research and education workbench program called "Koé", providing functionality similar to professional synthesis applications, such as Reaktor.

Twitch Specific Information

- My stream: [twitch.tv/cyberskunk](https://www.twitch.tv/cyberskunk) Mostly features me losing at PvP in various games.
- Favorite streamers currently:
 - [twitch.tv/sheevergaming](https://www.twitch.tv/sheevergaming) First streamer I supported. Did so much for women in Dota casting.
 - [twitch.tv/almightysavo](https://www.twitch.tv/almightysavo) Always fun to watch. First Twitch Creative I followed after Ross.
 - Any Dark Souls stream, really.
- Favorite games of 2015/2016: Undertale, Dark Souls III, Stellaris, Guild Wars 2

Hobbies and Other Activities

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