

Technical Snapshot

Synopsis: Ph.D. ABD in computational mathematics, software engineer in HPC, data scientist & engineer in ad tech.

Expert Knowledge In: Python, Bash, Cython, C, Sage, MySQL, Cloudera Hadoop, Impala/Hive, HDFS, OpenMP, SES

OS: RHEL-based and Debian-based Linux (sysadmin experience on both)

Work Experience

SpotXchange May 2014–present Westminster, CO data scientist

- Designed and implemented data ETL pipeline [Python/Impala/httpfs/Luigi/Parquet] for RTB analytics. Incoming data (every incoming/outgoing call, bid, impression, etc.) averages one million lines per second.
- Evaluated big data solutions relative to business requirements. [CDH, Hive/Impala, Spark, Drill, Cassandra, Oozie, Luigi, Sqoop, Parquet, httpfs, Flume, Kafka]
- Worked alongside brand safety team (fraud/bot avoidance) to ETL and analyze in-house and third party data.

SGI (Silicon Graphics International) May 2013–April 2014 Longmont, CO software engineer

- Led development of SCSI Enclosure Services (SES) backend [C/Cython] and frontend [Python] environmental monitoring utilities for the Modular Infinite Storage scale-out storage and compute platform.
- Wrote Massive Array Idle Disk (MAID) drive power [C] and udev/sysfs hardware utility [C++/Cython] libraries for the Infinite Storage Gateway platform.
- In charge of data path related system event logging for notification and system management. Wrote syslog event service [C++/Boost] and led event definition and cataloging management.

More work experience can be found at: www.linkedin.com/in/hilljb

Education

Ph.D. Mathematics (ABD) – computational group/graph theory University of Colorado Boulder

M.A. Mathematics (2010) – computational group theory, representation theory University of Colorado Boulder

M.S. Mathematics (2006) – computational number theory University of Vermont

B.S. Mathematics (2004) University of Michigan-Flint