

## Summary:

Accomplished Engineering Leader in Data Infrastructure with substantial experience scaling large data systems. Solid track record of designing, building, and maintaining scalable architectures, infrastructure, and technologies.

**Experience:** Start-ups, Distributed Systems, Scalability, Cluster Management, Distributed Storage

**Looking for:** Opportunities at early stage to mid-sized companies in San Francisco. Possible projects or roles:

- o CTO or Software Architect for funded startup, full-time or advisor.
- o Improving open-source data stores, such as large scale Key-Value Stores (NoSQL, C++).
- o Web-scale text analysis: web crawlers, pipelined analysis, webmap.
- o Management of large scale machine learning pipelines.
- o Build SaaS/PaaS products for start-ups and small companies.

## Work Experience Highlights:

6/2017 to present

### CTO

[ConnectWith \(Notify.io\)](#)

Connecting Businesses with customers, using mobile messaging. We've built a managed messaging platform specifically for professionals, out of our experience in Conversational Commerce for The Gap, The Wall Street Journal, many startups, and professionals. SMS, Facebook Messenger, Twitter.

App Store: [TextWith - Superhuman Texting](#)

Technologies: Django, DRF, Postgres, Kubernetes, ReactNative, DialogFlow.



7/2016 to 7/2017

### Founder and Chief Explorer

[CityScout.us](#)

Built MVP for The Urban Adventure App: <http://www.cityscout.us> : Stealth, members-only.

Technologies: PWA (Progressive Web App), Django, AWS, Bootstrap, JQuery, Python, Puppet, Nginx, etc.



3/2011 to 4/2016 (5 yr)

### Senior Engineer

[Facebook](#), 1 Hacker Way, Menlo Park, CA

Large-scale Infrastructure Engineer on data storage service (ZippyDB) and Monitoring systems teams.

- o Data Infrastructure: Worked on ZippyDB, a distributed and replicated key-value store based on Paxos and RocksDB. Helped port the real-time counter service (unique counters based on HyperLogLog) to ZippyDB (with live customers) and migrated it to Folly Futures for our C++ APIs. Implemented the first production version of the shard load balancer, which enabled balancing on several metrics across possibly heterogeneous hardware. (more: [Data@Scale 2015](#))
- o Core Infrastructure: Monitoring Systems: Moved the backend operating metrics and monitoring systems from a MySQL backend to HBase, and worked on Disaster Recovery. Worked on many of the subsystems such as real-time percentile aggregation, prediction modeling, and the real-time computation engine. System processed over 1 Billion time series datapoints per minute at the time I left the team. ([O'Reilly Strata 2012](#))

Heavily involved in debugging many Facebook technical problems (outages / SEVs) and reported at several Incident Reviews, mentored interns, and had a lot of fun at every possible hackathon.

Technologies: C++, Python, PHP/Hack/HipHop; HBase/HDFS/Hadoop, MySQL, RocksDB, Paxos.



4/2003 to 2/2011 (7 yr)

### Senior Engineer

[Yahoo Search Technology](#), 701 First Ave., Sunnyvale, CA

Linux cluster kung-fu. Worked on web search content systems for 7 years, with 4 managers, on 3 teams, during 5 cluster doublings, with thousands of nodes per cluster. Mainly worked on cluster management and monitoring software (in C++) and the operation of the production web crawlers.

- o Content Operations Team: Managed the deployment, operations, and monitoring of production web crawler clusters in distributed data centers. Typical projects included moving 1.8 PB across coasts. Catalog Agility Project: Primary architect on crawler scalability and data storage changes (C++) resulting in a "You Rock Award".
- o Crawler Team: Catalog Master: Built and extended cluster management and monitoring software for managing thousands of nodes (C++). Many projects ranging from investigating problems based on statistical anomalies in freshness, comprehensiveness, and quality (Python) to writing disk burnin software (C) used by all of YST (Yahoo Search Technology). Wrote CTA package (Cluster Trends Analyzer) in Python and CluStats (Cluster Stats and alerts) libraries in C++.
- o Echelon Team: Echelon Master: worked on cluster management software for controlling indexing pipeline and batch processing (C++). Echelon Copier (EchCp in C & C++): cross-coast database distribution software with extensive performance throttling controls for delivery production index data to other groups.

Technologies: Unix Systems Programming; Networking; C++, Python, Perl, Shell, SQL.

6/2001 to 3/2003 (1 yr, 9 mon)

### Engineering Manager

Tapstone Inc., Walnut Creek, CA. (Start-Up, 12 developers, 20 people)

Tapstone's Server Change Center is an IT tool for managing enterprises with hundreds to thousands of machines. Agent software running on each computer detects changes to the file system, the registry, and configuration files (on both Solaris and Windows.) These

agents each communicate a stream of events in real-time to a central appliance (running Linux and Postgres.)

- o Designed the distributed network agent for portability, scalability, and low-impact.
- o Wrote low-level back-end software for Solaris, Linux, and Windows NT in Python and C++. Implemented hardware scanner, control daemon, and various libraries.
- o Reviewed code, unit tests, system tests, and design work as the manager and lead engineer of the agent group (4-6 developers, including myself.)
- o Designed APIs, libraries, and protocols as lead engineer of the architecture team during system design phases. Reviewed system-wide design and implementation choices to support the goals of high performance, fail-safe operation.
- o One of the authors of the provisional patent application: Management and Monitoring System for Distributed Computer Networks.

Technologies: Systems Programming (cross-platform); Python, C++, Java; TCP/IP, XMLRPC, SQL.

#### Work Experience:

---

- Design: Plan and develop company-wide software libraries, APIs, protocols, databases, and other infrastructure.
- Review: Conduct code reviews, system design reviews, performance analysis, technical mentoring.
- Test: Design and build unit tests, system tests, and staging / roll-out environments.
- Lead: Evaluate feasibility and suitability of designs and technologies, given business and resource constraints.
- Run: Maintain and operate online services with e-commerce transactions 24x7.

#### Education:

---

1987-1991: Columbia University, School of Engineering and Applied Science

Bachelors of Science in Electrical Engineering (BSEE).

Masters EE: Worked toward a masters in EE with a focus on VLSI design, while employed by the Computer Science Department at Columbia University. Left to become CEO of a start-up in 1994 without completing degree.

Teaching Assistant: Object-Oriented Design and C++ (John Lakos), Networks (Matt Blaze), C-Programming (Dan Duchamp), Scientific Computation (Joseph Traub).

**Hobbies:** (other than hacking) Gardening, Woodworking, Lean Startups, Tinkering, MTG.