Joseph D. Batchik

josephbatchik@gmail.com - (410) 599-3550

Work Experience

Apple Inc. (iCloud) - San Francisco, CA — 2016 - Present
Software Engineer — http://apple.com/

- Designed key aspects of the backing blob store engine that enables iCloud to scale to hundreds of millions of users each month.
- Architected the new multi-tenant compaction system which provides better throughput guarantees and reliability.
- All of this work provides better resource forcasting for finance teams.
- Engineered a proactive solution to data loss prevention. Lead to the discovery of several undiscovered, subtle bugs in the underlying frameworks and data store.
- Vastly reduced database load by introducing a framework that allows numerous subsystems to share the same pooled resources concurrently.
- Languages / tools used: Scala, Java, Cassandra, MapReduce

Cloudera - San Francisco, CA — Summer 2015
Software Engineer Intern — http://cloudera.com/

- Implemented network performance increase in Apache Spark that reduced traffic by over 90%
- Integrated Apache Avro as a first-class citizen into Spark core for use in RDDs
- Languages / tools used: Scala, Java, Apache Spark

Google - New York, NY — Summer 2014
SRE Engineering Practicum Intern — http://google.com/

- Implemented load testing infrastructure for newly release software, allowing for early detection of bugs and performance defects
- Reduced request latency for 70% for back-end monitoring services
- Languages / tools used: Java, Python Protocol Buffers, Google data stores.

Amazon - Seattle, WA — Spring 2014
Software Developer Engineer Intern — http://amazon.com/

- Overhauled internal search capabilities for the Enterprise Data Warehouse team, allowing for near-real time search capabilities for Amazon's data analytics.
- Designed the new search system to be fault tolerant to preserve data integrity.
- Languages / tools used: Java, various Amazon cloud products, including Cloud-Search and SNS.

John Hopkins University Applied Physics Lab - Laurel, MD — Summer 2013
Engineering Intern — http://jhuapl.edu/

- Working with a team, developed a sensor management system used to control and collect data from multiple telescopes remotely.
- Languages / tools used: Java, Ant, SVN, SQL, Google Protocol Buffers.

Education

Rochester Institute of Technology - Rochester, NY — 2012 - 2016
Major: Computer Science — In-Major: 3.82 GPA, Overall: 3.60 GPA
Graduated *Cum Laude*

Skills & Certifications

Languages Scala, Java, Python, Go, Rust
Tools Git, Spark, Avro, Gradle, PostgreSQL, Cassandra, Protocol Buffers, Thrift, OpenJDK JMH
Certifications Cloudera Certified Developer for Apache Hadoop, 2012
Apache Spark Contributor Developed a solution to allow for Spark to efficiently read / write Apache Avro data formats. Worked on features in the Spark SQL engine.

Self-Directed Projects

Raft Key-Value Store

- Distributed key-value store that provides linearizability guarentees for all operations.
- Out of the box support for leader election, transparent handling of failing nodes, and correctness under network partition.

CRDT Distributed Tally Service

- Distributed backend counting service that is capable of withstanding large amounts of concurrent requests.
- Uses G-Counters as the backing asynchronous replication model.
- Implemented with a combination of lightweight threads for the request handling, Zookeeper as the election system, and Thrift as the shared communication protocol.

Github Language Analysis

- Data analyzer and ingest pipeline using Go to determine the programming language usage across all of GitHub.
- Provides key insights to language usage trends over time and in comparison to each other.
- Implemented a distributed ingest pipeline to increase processing capabilities.