

Arron Norwell

Summary

A results-oriented engineer with proven experience designing and building robust, scalable, and concurrent backend software with millions of users, thousands of requests per second, and petabytes of data. Expertise with distributed computing, SQL and no-SQL database application development, and performance optimization. Experience architecting systems, leading teams, and delivering products successfully and on-time.

Skills

- Software Architecture and Design
- Relational DBs (Postgres, MySQL)
- Distributed Systems
- APIs (HTTP, REST, OpenAPI, Protobuf)
- Cloud (AWS, Google Cloud, Azure)
- NoSQL (DynamoDB, Elasticsearch, Redis)
- Scalability and Performance Optimization
- DevOps (CI/CD, CDK, Containers)

Languages: Java, Python, Rust, Scala, Ruby, Go, Typescript, Javascript

Experience

Oct 2020 - **Senior Software Engineer**, *Amazon Web Services*, Boston, MA

- Present Designed and implemented the Elastic File System (EFS) cross-region replication feature, using Java, Python, and AWS CDK.
- Built simulation environment and property testing system to allow programmable, developer-driven, end-to-end randomized testing of EFS under the presence of faults and concurrent modifications.
- Analyzed and optimized performance of the EFS backup and restore system, resulting in a 4x throughput improvement.
- Supported and mentored a team of 8 engineers with code reviews, design reviews, learning sessions, and interviews.
- Operational and infrastructure development for a globally-deployed distributed system with millions of users.

June 2017 - **Principal Software Engineer**, *Datto*, Boston, MA

- Oct 2020 Initial developer and technical lead on reimplementing of SaaS backup system, using Scala, Elasticsearch, and ZFS. This system went live ahead of schedule with an improved featureset and has now replaced the original system, supporting millions of users and storing dozens of petabytes of data.
- Prototyped new product features that became incorporated into product, including a customer-facing backup details information view, reporting visualization view, and Azure-based Exchange data restore tool.
- Set up application monitoring framework and libraries to consistently collect app events and metrics, providing observability, dashboards, and KPIs for a large distributed system that handles hundreds of millions of events per day.
- Supported three scrum teams with architectural guidance, mentoring, and code reviews.
- Worked with Scrum product owners and engineering leadership to help plan product roadmap and set technical direction.

Nov 2013 - **Senior Software Engineer**, *Datto*, Boston, MA

- June 2017
- Designed and implemented stability, scalability, and feature improvements for a large distributed systems architecture, able to support backups for over a million users daily and built in Ruby on Rails using Redis, PostgreSQL, Cassandra, and Amazon S3.
 - Acted as project technical lead for a team of seven developers on a project to deliver a full Backup and Restore Solution for Office 365. Project was completed successfully and on-time. Presented the project to hundreds of partners at the DattoCon 2015 conference.
 - Designed and was main contributor to open-source projects for an internally-used Apache Cassandra client and schema management tool.
 - Designed and implemented an extendable API client and authentication framework that handled tens of millions of requests per day against many different APIs.
 - Led a team involved in building a permissions management system as part of an effort to convert to a service-oriented architecture.

Apr 2012 - **Software Developer / Team Lead**, *mDialog*, Toronto, ON

- Nov 2013
- Designed and implemented components of a distributed, asynchronous, and highly scalable video streaming platform built in Scala, Java, and C, using Redis, PostgreSQL, Apache Zookeeper, and ZeroMQ.
 - Horizontally and vertically scaled a software platform from zero users to over 250,000 concurrent video streams running on clusters of over fifty machines; designed software components proven to scale vertically to thousands of requests per second.
 - Open-sourced several internally-developed Scala libraries, including an asynchronous Redis client, an asynchronous HTTP client and server, and a ZeroMQ library.
 - Interviewed potential new hires and participated in code reviews.
 - mDialog was acquired by Google in April 2014.

May 2009 - **Software Development Engineer in Test**, *Microsoft*, Mountain View, CA

- Aug 2010
- Led project to Design and implement VOIP test framework for quick and flexible creation and execution of tests, with web UI for test reporting.
 - Created automated graphing tools for visualization of performance data.

Education

Sept 2010 - **M.Sc. - Computer Science**, *University of Toronto*, Toronto, Ontario, Canada

- Feb 2012
- Thesis: *A Threshold for Clusters in Real-World Random Networks*
 - GPA: 4.0

Sept 2006 - **B.Sc. - Combined Honours, Mathematics and Computer Science**, *University of British Columbia*, Vancouver, British Columbia, Canada

- May 2009
- GPA: 92

Selected Projects/Open Source

Claippy (2024 - Project), github.com/ANorwell/claippy

The command line AI power programmer.

QuickBlog (2020 - Project), github.com/ANorwell/quickblog

An easy to use blogging tool that allows anyone to set up a blog in minutes. Powered by Vue.js, Typescript, and Gitlab Actions CI/CD.

SNP Reporter (2020 - Project), github.com/ANorwell/SNPReporter

A Rust-based tool to scrape Single Nucleotide Polymorphism (SNP) genetic data and generate a genetic report based on raw data from 23AndMe and similar genetic or ancestry sites.

anorwell.com (2010-2020 - Project), anorwell.com

My personal website is a minimally-featured blog tool, as well as other web-based experiences. Reimplemented multiple times in PHP, Python/MySQL, Vue/Typescript-powered static SPA.

CassSchema (2015 - Primary Contributor), github.com/datto/cass_schema

A Ruby gem for managing multiple cassandra schemas across multiple clusters.

Pyper (2015 - Primary Contributor), github.com/datto/pyper

A Ruby gem for constructing data storage and retrieval pipelines, with support for data validation and serialization.