Pinglei Guo

Vancouver, BC, Canada

Email: plguo002@gmail.com Website: at15.dev GitHub: https://github.com/at15 LinkedIn: https://linkedin.com/in/at1510086

WORK EXPERIENCE

AWS: Software Engineer - Amazon CloudWatch and X-Ray   Vancouver, Canada   Sep 2020 - Present

- Led a team of 6 engineers to implement W3C trace ID support in AWS X-Ray. Refactored data sharding and indexing logic to remove X-Ray trace ID timestamp dependency, unblocking 20,000 AWS accounts and enabling OpenTelemetry integration.
- Collaborated with Lambda and Application Signals teams to integrate OpenTelemetry in X-Ray ingestion API. Reduced Python Lambda function cold start time by 60% by removing OpenTelemetry collector from the Lambda layer.
- Designed and implemented conversion logic between OpenTelemetry trace spans, X-Ray segments and CloudWatch transaction search log events, reducing customer trace data storage cost by 70% using log storage so they can sample 100% traces.
- Maintainer of open source projects on GitHub: Amazon CloudWatch Agent and Amazon Distro of OpenTelemetry Collector.
- Developed the ECS extension for OpenTelemetry Collector, enabling Prometheus discovery with metadata from ControlPlane.

Google: Software Engineer - Batch job on Kubernetes   Sunnyvale, CA   May 2019 - Nov. 2019

- Eliminated the 5% job lost during controller deployment by optimizing the order of database update and job submission.
- Reduced debug session startup latency from 2 minutes to 40 seconds by loading artifacts on demand from object store.

PayPal: Software Engineer - Multi cluster container orchestration platform in Go   San Jose, CA   May 2018 - May 2019

- Developed a Go-based monitoring tool to correlate logs, metrics, and traces on deployment failures, automatically identifying root causes base on topology. This reduced on-call debug time from 10 minutes to 2 minutes.
- Introduced canary deployment strategy and dependency readiness check, reducing deployment rollback ratio by 20%.

Dongyue Web Studio at SJTU: (Part-time) Full stack web developer & Tech lead   Shanghai, China   Sep. 2013 – Jan. 2016

- Led web and mobile team to redesign the online event booking website tongqu.me, implementing a REST API and single-page application using AngularJS. Reduced codebase size by 60% and improved page load time by 80%.
- Implemented a Redis and MySQL based queue system to handle ticket rushes, preventing website crashes and ticket overselling while supporting 5x more traffic.

PROJECT EXPERIENCE

Distributed database benchmark system   github.com/benchhub   UCSC   Nov. 2017 – March. 2018

- Designed a specification for running database benchmark for RDBMS and TSDB in distributed environment.
- Implemented a continuous integration service that stores benchmark results in databases to detect performance regression.

Distributed Time Series Database   github.com/xephonhq/xephon-k   UCSC   Nov. 2016 – Present

- Implemented a distributed time series database on top of Cassandra in Go. Support both JSON and Protobuf via HTTP/2.
- Designed a columnar storage engine modeled after Parquet and InfluxDB with high compression and less write amplification.
- Created a benchmark suite for OpenTSDB, KariosDB, InfluxDB with a generic interface to support different TSDB.
- Surveyed popular TSDB design and implementation, made an interactive online report awesome-time-series-database.

GPU accelerated in-memory time series processing   github.com/at15/ts-parallel   UCSC   Apr. 2017 – June 2017

- Expanded benchmark suite for different C++ GPU computing framework on CUDA and OpenCL, Thrust, Boost, ArrayFire.
- Implemented OLAP queries like top-K, group by for multi dimensional time series data on both CPU and GPU backends.

EDUCATION

MS. Computer Science	University of California Santa Cruz	Sep. 2016 – Mar. 2018
BS. Materials Science	Shanghai Jiao Tong University	Sep. 2012 – June 2016

SKILLS

Language	Go, Java, JavaScript, TypeScript, Python, Dart, Rust, C++, SQL
Infra	AWS, GCP, Cloudflare, Kubernetes, Docker, CDK, Terraform, Prometheus, Grafana
Database	DynamoDB, Cassandra, MySQL, PostgreSQL, Elasticsearch, Redis
Framework	gRPC, React, Flutter, Spring