



# Andrew Liang

 [github.com/liangdrew](https://github.com/liangdrew)  [liangdrew@gmail.com](mailto:liangdrew@gmail.com)

 [linkedin.com/in/liangdrew](https://linkedin.com/in/liangdrew)  [liangdrew.com](https://liangdrew.com)

## Experience

### Databricks • Software Engineer Intern

Jan - Apr. 2019

San Francisco, CA

- Added features to custom **Spark** runtime to support team verticals in resource management, security, and library management
- Architected mechanism for Databricks control plane to consume arbitrary event messages from highly-distributed **Spark** clusters in the data plane using **Scala**
- Contributed to development of **instance pools**: a VM caching feature used to reduce **Spark** cluster creation time by **90%**

### NVIDIA • Distributed Storage Engineer Intern

May - Aug. 2018

Santa Clara, CA

- Contributed to core development of **AIStore**, a distributed storage system for AI in **Go**
- Invented a **petabyte-scale** object sorting layer within AIStore: [bit.ly/2yeYxOc](https://bit.ly/2yeYxOc)
- Implemented dynamic node **keepalive scheduler** based on TCP's RTO (RFC 6298)
- Optimized for performance bottlenecks at scale using **pprof**, **go-torch**, **tcpdump**

### Taplytics • Backend Engineer Intern

Sept. - Dec. 2017

Toronto, ON

- Created distributed push notification pipeline with **Cloud Pub/Sub** and **Cloud Functions**
- Built real-time webpage element classifier in **JavaScript** for A/B test recommendations
- Added RESTful HTTP APIs in **Node.js** to support data models backed in **MongoDB**

### Wattpad • Platform Engineer Intern

Jan. - Apr. 2017

Toronto, ON

- Built an open-source client library in **Go** for interfacing with third-party ad-serving API
- Implemented search-related HTTP APIs to in **Go** to support new ad requirements
- Leveraged circuit breaker design pattern in **Go** to harden web microservices

### Paytm Labs • Data Engineer Intern

May - Aug. 2016

Toronto, ON

- Patched data pipelines in **Python**, **Scala** to support new data ingestion requirements
- Wrote daily indexing flow to write 30M documents from **Elasticsearch** to **HDFS**

## Skills

### Languages

Go	Python	Java
Scala	C++	SQL

### Technologies

Git	AWS	K8s
Linux	GCP	Bazel
Spark	Docker	Jenkins

## Projects

### go-adzerk

HTTP client library in **Go** for interfacing with Adzerk's API

### wattpoll

Microservice built with **Go** and **MySQL** which supports interactive polls on Wattpad

## Interests

Distributed computing  
Scalable web architecture  
Big data systems  
Computer security

## Education

### University of Waterloo

Software Engineering 2020