Gunjan Kaphle

Experienced Software Engineer working on a regular basis to build scalable and reliable frameworks. Passionate about stream processing, distributed systems, and operational excellence.

Personal Info

Phone
(619)-560-6209

E-mail
gunjankaphle@gmail.com

Website
gunjankaphle.github.io

Github
github.com/gunjankaphle

Skills

Experienced:
Java, Python, Go, AWS, Kafka, Vertica, MySQL, Airflow, Flask, Looker

Familiar:
JavaScript, Ruby, Docker, GCP, Jersey, Ansible, Oracle DB, Redis

Interests/Self-Learning

Apache Spark, React

Presentations

Cerner Developer Conference, 2018
Building smarter data pipelines using Apache Airflow

Data Science Seminar, The RealReal, 2020
Stream Processing with Kafka and KSQL

Work Experience

Software Development Engineer, AWS SageMaker — MAR 2020 – Current
Amazon, Seattle, WA

- Implement new features and improve current architecture for AWS SageMaker Hosting, a platform to deploy ML models and run real-time predictions, using Go, Java, and AWS services
- Design and build a system to rigorously test feature launches and architectural changes that helps identify major issues before production releases
- Build internal and external error handling framework throughout the Hosting services to enhance customer experience and aid platform troubleshooting
- Build an automated system to gather and aggregate weekly operational metrics for all our APIs, saving over 4 hours of weekly repetitive work for On-Call

Data Engineer — SEP 2019 – MAR 2020
The RealReal, San Francisco, CA

- Built an ETL framework using Python, Kafka, and Apache Airflow to move data from different systems to Data Warehouse, BigQuery
- Wrote and maintained Google Cloud Functions to support the ETL architecture and test our ETL pipelines using anomaly detection algorithms
- Designed and built a system to extract and delete data across our services to comply with Data Protection Regulations, such as GDPR and CCPA
- Created CI/CD pipelines using Github Actions. Improved the code base by adding unit tests, integration tests, code coverage checks, and linters

Associate Senior Software Engineer — JAN 2016 – SEP 2019
Cerner, Kansas City, MO

- Extended Apache Airflow and created a Workflow Management System to speed the process of building ETL pipelines and creating batch processing jobs for different teams at Cerner
- Implemented a RESTful API in Java for users to post files that got ingested to our data warehouse, parsed it, and made it consumable for other apps via a REST API. Over 50K files and 100gb of data processed daily
- Collected, aggregated and baselined user experience metrics across the client base for every user, as granular as 1 minute, using Apache Kafka and Vertica
- Built a Python Flask Web Application on top of Ansible to deploy apps to over 70k hosts

Software Engineering Intern — SEP 2014 – DEC 2015
Thomson Reuters on behalf of Maverick Software Consulting, Mankato, MN

Education

Bachelors of Science in Computer Information Systems — DEC 2015
Minnesota State University, Mankato, MN — *GPA 3.90*