



REZA RAJAN

 /in/rezarajan
 rezarajan

Data Engineer

Experience

Mochi - DevOps Engineer

Dec 2022 - Present

Technologies: Bash, Nix, SQLite, InfluxDB, Ansible, Terraform, Kubernetes, Cloud-Init

- Architected a CI/CD pipeline for building, testing and deploying of multiplatform containerized applications, *reducing deployment errors by 90%*.
- Employed automation and Infrastructure-as-Code to provision highly available infrastructure using Proxmox, Terraform, Ansible and K3S, *enabling 98% operational reliability and reducing overall recovery time by up to 40%*.
- *Reduced incident response time by 60%* through deploying comprehensive monitoring solutions using Grafana and Prometheus, resulting in faster issue resolution and increased service availability.

myKaarma - Data Engineer

Jul 2023 - Jan 2024

Technologies: Python, SQL, Flink, Kafka, PySpark, AirFlow, Redshift, Lambda, Iceberg

- Collaborated with cross-functional teams to design and deploy a cloud data lakehouse architecture on AWS which *replaced 3 stored procedures from the legacy MySQL data warehouse*.
- Developed and deployed a ELT data processing pipeline using Python, Airflow, AWS Glue and AWS Redshift, to process terabyte-scale big data, *reducing processing times on large datasets by 70%* compared to SQL stored procedures on the MySQL data warehouse.
- Executed a PoC for a distributed streaming data processing pipeline using Apache Flink AWS S3, AWS Glue and AWS Iceberg, providing real-time reports on message data, *reducing time-to-value by 95%*.
- Automated customer report generation and delivery using AWS State Machines and Lambda functions, with SLA guarantees, *reducing late report delivery incidents by 99%*.

myKaarma - Data Analyst

Feb 2022 - Jul 2023

Technologies: Python, SQL, Power BI, Looker Studio, Airflow, PySpark, S3, Iceberg

- Communicated with major client groups to gather requirements for *five group-level reports*.
- Developed and delivered data products and drill-down reports for *over 500 business clients* based on group-level report requirements.
- Enabled self-service analytics through integration of data sources with Power BI, *reducing ad-hoc report requests by 10%*.
- Executed a PoC for a ETL data processing pipeline, using Airflow, PySpark, S3 and Iceberg which *reduced processing times of large-scale message data by 60%* over the MySQL data warehouse.

Certifications

Data Engineer

Udacity - 2024

Self-Driving Car Engineer

Udacity - 2021

Data Analyst

Udacity - 2020

Education

Bsc. Mechanical Engineering

University of Waterloo
2015- 2020

Awards

University of Waterloo President's Scholarship of Distinction

2015

Languages

English - Native

Technical Skills

Programming Languages:

python, golang, sql, bash, nix

Data: postgresql, mariadb, sqlite, redshift, influxdb, cassandra, mongodb, s3, flink, lambda, dbt, airflow, kafka, spark, elasticsearch, etl/elt

Visualization: grafana, power bi, looker studio, plotly

Data Libraries: pandas, scipy, numpy, tensorflow, pyspark

Infrastructure: linux, docker, kubernetes, aws, proxmox, ansible, terraform, opentofu

GitOps: git, github actions, ci/cd