Frank Zhang-Zheng

Phone: (989) 615-4879 | Email: contact@frankzhangzheng.com | LinkedIn: in/fzhangzheng

Education

Michigan State University East Lansing, MI
Bachelor of Science in Physics 2014 – 2019

Skills

Programming Languages	Python | SQL | TypeScript
Cloud	AWS (S3, Redshift, Step Functions, Lambda)
Streaming	Apache Kafka | AWS (Kinesis, SQS, SNS)
DB	SQL | NoSQL | MongoDB | Snowflake | AWS (DynamoDB, RDS)
Deployment	AWS CDK | Terraform | Docker | Kubernetes

Work Experience

Associated Bank Milwaukee, WI
Data Engineer/Data Platform Engineer Nov. 2021 – Present

Data Contract Framework

- Developed a CI/CD pipeline using GitHub Actions to automate data contract deployment
- Built a REST API with Python/FastAPI in AWS with DynamoDB table to provide data contract access
- Created technical documentation and training materials to facilitate user adoption

Event-Driven Streaming

- Pioneered and implemented an event-driven streaming framework to ingest high-volume Kafka data into Snowflake, creating the first near-real-time data processing capability for the organization
- Engineered a scalable data ingestion pipeline using AWS Lambda, Apache Kafka, AWS S3, and Snowflake Snowpipe, capable of processing hundreds of messages per minute
- Successfully migrated project architecture from batch file processing to a streaming solution under tight deadlines

Manual Workflow Automation

- Automated a manual workflow for forced closed accounts, reducing the time spent on the process by 75%
- Designed and implemented a data pipeline to extract, transform, and loading banking data into Salesforce CRM
- Developed Python processes to use the Salesforce Bulk API for creating new cases and distributing processed data to third-party vendors and regulators

Xpress Technologies Scottsdale, AZ
Data Engineer Nov. 2020 – Nov. 2021

- Engineered and deployed an automated data ingestion system, accommodating various file types (CSV, Parquet, TSV, Excel, plaintext) by dynamically identifying formats via file extensions and magic bytes

- Orchestrated data processing and quality checks using AWS Lambda Step Functions, integrating with an AWS DynamoDB table that stored data quality rules (e.g. null value thresholds, data type consistency)
- Developed and maintained a Node.js/TypeScript service on a CRON schedule to synchronize carrier data from MongoDB to HubSpot CRM, ensuring timely customer status updates for marketing operations
- Engineered an event-driven data pipeline to periodically ingest weather data from an external API (FreightWaves SONAR) into a Kafka topic, enabling real-time consumption by a downstream PostgreSQL database, and access to data needed for machine learning operations
- Used TypeScript, Node.js, and Docker to develop and deploy microservices to AWS to facilitate data transfer across the organization, with a focus on delivering clean and enriched data to the machine learning team to increase accuracy and usage of the core pricing product