Vamsi Krishna

Java Data Engineer,

9390707338, krishraju5002@gmail.com

www.vamsikrishnamraju.com

PROFESSIONAL PROFILE

- With over 4 years of technical experience, I excel as a **Java Data Engineer**. My focus is on **designing**, **implementing**, and **optimizing** data pipelines, ETL processes, and data architectures to ensure robustness and efficiency.
- Utilized Agile methodology across the software development life cycle. Proficient in designing and developing ETL processes within **Azure Data Factory**, facilitating seamless data migration from external sources to destination.
- Skilled in crafting parallel data processing workflows, utilizing advanced techniques including Join, Merge, Lookup, Deduplication, Filtering, Dataset Management, File Set Lookup, Complex Flat File Parsing, Transformation, Aggregation, and XML Parsing. Specialized in optimizing data engineering processes to achieve optimal outcomes.
- Extensive experience designing, developing Reactive Event Driven **Microservices** using Streaming Platform Confluent Enterprise Kafka, **Kafka Connect**, Debezium, Event Sourcing, **CQRS design patterns**.
- **ChatGPT:** Integrated ChatGPT in java(created a command line interface that talks to ChatGPT)
- Hands on with Cloud platforms (**AWS, AZURE**) and container orchestrators (**Kubernetes**), **Docker** and API Gateways like **Apigee**.
- Proficient in utilizing the **Azure Marketplace** for searching, deploying, and procuring a variety of applications and services
- Experienced practitioner of Object-Oriented Analysis/Design and **Heroku's 12 Factor principles** in Cloud-Native microservices development.
- Expertise in achieving high code coverage using Unit tests, Integration tests, Spring Boot Slice based tests, Mockito based tests.
- Solid experience designing API/microservices that employ both SQL Databases (Oracle, SQL Server, DB2, MySQL) and NoSQL databases like (MongoDB, ElasticSearch and Redis).
- Strong knowledge of Java with experience in Core Java.
- Excellent written and verbal communication skills

TECHNICAL SKILLS

Cloud Platform & Technologies

▪	AWS, AZURE, Kubernetes, OpenShift, Elastic Beanstalk, EKS, AWS RDS	
▪	Confluent Kafka Platform 5.x, Kafka Streams, API Gateway Apigee, Okta, PingFederate, OAuth, OIDC, Redis 5.x/4.x, Elasticsearch 7.x/5.x, MongoDB 4.x	

Languages and Frameworks

▪	Spring Framework 5.x (IoC, Web, MVC, AOP, TX, JDBC), Spring Data, Spring Security, JPA, Hibernate, Spring Rest, Restful Services, OpenAPI/Swagger.	
▪	Java 11/8, Kotlin, Python, Pyspark, SQL	
▪	Spring Boot 2.x, Spring Cloud Netflix OSS (Config Server, Eureka, Ribbon Zuul, Hystrix, Spring Cloud Sleuth), Spring Cloud Streams, Spring Kubernetes, Spring Kafka, Apache Avro, Google Protocol Buffer, JSON Schema	

Azure Ecosystem

▪	Azure DataLake, Azure Databricks, Azure Data Factory, Azure SQL, Azure Synapse Analytics.	

DevOps & Other Tools

▪	Git, GitHub, Bitbucket, SVN, Apache Maven, Gradle	
▪	Jenkins, ArgoCD, Terraform, IntelliJ Idea, Eclipse, WebStorm, Visual Studio Code, **ChatGPT**	

Development Methodologies

▪	Agile (Scrum, Test-driven development (TDD), UML (Sequence, Class, Interaction Diagrams).	

PROFESSIONAL EXPERIENCE

TD Insurance **Apr '22 – Feb'24**	**Role: Data Engineer** **Insurance Self-Serve Platform:** **<u>Initiatives</u>**: ● Participated in the Online Platform team and worked on the backend platform for the Quote generation platform for Home and Auto Insurance. ● Designed/Developed "File a Claim" and "Manage your claim" customer journeys for Auto Insurance. ● Utilized Azure Cluster services alongside Azure Data Factory V2 to efficiently ingest a vast and diverse range of data from various source systems into Azure Data Lake Gen2 ● Engineered custom data transformation logic using Azure Data Factory Data Flows expression language, empowering efficient data manipulation and cleansing operations tailored to TD Insurance's specific data requirements.

	● Collaborated closely with business stakeholders, data analysts, and domain experts to understand insurance data domain intricacies and translate business needs into scalable and robust Data Flow solutions. ● Streamed the data from Relational Database (DB2) via CDC (SQ-Data) to ElasticSearch to build a Kibana based dashboard for Actuaries and Underwriters to serve their analytics needs – Claims by region, demographic, age, make/model. Used Kafka Streams (KStream, KTable API) to enrich the data against internal reference data, do transformations and write into ElasticSearch cluster. **<u>Responsibilities</u>**: ● Participated in technology meetings with solution architect to finalize the E2E CI/CD pipeline, DevOps, Source Code Analysis (SAST) tools, Code coverage tools. ● Reviewed SQL scripts and optimized them using Spark SQL to improve performance. ● Built User stories in JIRA for each customer journey during a brain storming session with the team. ● Implemented **Continuous Integration/Continuous Deployment** (CI/CD) pipelines for Azure Data Factory using **Azure DevOps**, enabling automated deployment of data pipeline changes across development, testing, and production environments ● Integrated (using OAuth Client Credentials flow) with 3rd party providers like Canada post for address lookup/validation, loaded vehicle make/model information from mainframe file feeds (using Spring Batch) into Elastic search and exposed via API for lookup. ● Involved in all phases of the Software Development Life Cycle including requirement gathering, designing the application, implementing the design, testing and maintenance support. ● **ChatGPT:** Integrated ChatGPT in java(created a command line interface that talks to ChatGPT) **Technologies Used: Spring Boot 2.x,Azure Data Factory, Azure Cloud, Python, Terraform, Java 11/8, ElasticSearch 6.x, Confluent Kafka 5.x, Swagger, EFK(ElasticSearch, Fluentd, Kibana), Splunk,Jenkins,**

Scotiabank **Mar '20 – Feb'22**	**Role: Data Engineer** **<u>Initiatives</u>**: ● Implemented Data Client layers by creating multiple Rest Api's. Implemented the applications using Spring Boot framework and handled the security using Spring Security. Used Microservice architecture with Spring Boot based services interacting through a combination of **Apache Kafka** message brokers and REST. ● Developed and maintained data processing pipelines using Azure Databricks, leveraging Apache Spark for scalable and high-performance data transformations. ● Support the development of data pipelines using Azure Databricks to process and transform data from various sources. This includes learning to write code in Python, SQL within the Databricks environment to perform data transformations and aggregations. ● Assist in ingesting data from different sources into Azure Databricks and storing it in appropriate formats. Learn to work with Azure storage services such as Azure Data Lake Storage or Azure Blob Storage for data storage and retrieval. **<u>Responsibilities</u>**: ● **<u>DevOps:</u>** Setup Jenkins based CI/CD pipeline, used Maven for project builds, written unit and integration tests to provide complete unit and functional coverage. All APIs are deployed to RedHat Managed OpenShift (Kubernetes) PaaS platform. Configured Auto Scaling Policy using Horizontal Pod Autoscaler. ● **<u>Agile:</u>** Followed agile methodology – daily scrum meetings, standup sessions, creating and sizing user stories in JIRA, quick prototyping and focus on delivering working software than elaborate documentation. As a team, always met targets to provide over 80% test coverage for the running code. ● Support data analysis tasks within **Azure Databricks** by learning to use tools like **Spark SQL** and **DataFrame operations**. ● Collaborate with senior data engineers and data scientists to understand project requirements and contribute to the development of data solutions. Assist in documenting technical designs, workflows, and best practices for future reference. ● **<u>Technologies Used</u>: Java 8, AWS, Azure, Docker, Spring Boot 2.x, Spring REST, Spring Netflix OSS (Config Server, Cloud Sleuth, Zipkin, Hystrix Circuit Breaker), Spring Kafka, Confluent Kafka 5.x,**

	AWS RDS, JPA, Swagger, New Relic, SonarQube, Jenkins, Hashicorp Vault, Azure Databricks, Spark SQL.

Aug '16 – Jul '17	**Role: Software Intern** **Initiatives**: ● Establishment of private cloud infrastructure using **OpenStack** - To build the infrastructure service in the cloud for providing the storage, compute resources. ● Installed the necessary components by configuring network interfaces and name resolutions in both controller and compute nodes. ● installed Hypervisor(KVM), SQL(**MariaDB**) to store information on controller nodes. ● Installed messaging queue(Rabbitmq-server) on controller nodes for coordination among services. ● Configured Authentication and authorization to all OpenStack services by using Openstack Identity Service(Keystone). ● Configured dashboard that provides a web-based interface which is used to perform operations like launching an instance, allocating IP address and setting access control.

EDUCATION

- Master's in Computer Science, **Western University**, Canada.
- Bachelor's in computer science, **Gitam University**, Hyderabad, India