Nuthan Reddy

Big Data Solutions Architect

9474985552 | nuthan.reddy999@gmail.com | linkedin.com/in/nuthanreddy/ | github.com/NuthanReddy

Summary

Senior Software Engineer and Big Data Solutions Architect with 10+ years of experience architecting scalable data systems and leading high-performing engineering teams. Expert in distributed systems, cloud platforms, and big data technologies.

Work Experience

Microsoft   Mar 2022 – Present
Senior Software Engineer   *Hyderabad*

- Directed Privacy Signals Team managing data deletion, export delivery, processing and compliance reporting across Microsoft services.
- Architected and delivered 2-20x performance improvements in Data Scanning by eliminating bottlenecks and non-deterministic code.
- Designed and deployed scalable Data Map solutions with APIs for asset tagging and updates, improving data governance.
- Established deployment, monitoring and alerting frameworks for the Signals Team, ensuring high reliability and observability.
- Led cross-functional teams to deliver distributed data engineering solutions on Azure and Kubernetes.
- Drove cloud migration and compliance initiatives, collaborating with business and engineering stakeholders.

Walmart   Oct 2020 – Mar 2022
Software Engineer III   *Bangalore*

- Managed development of rule-based data processing framework for Enterprise Business Services Teams, focusing on scalable distributed systems.
- Crafted modularized CILL rules engine with config-driven architecture for improved maintainability and compliance.

- Developed Code Freeze App using Go to automate GitHub repo management via Slack, streamlining deployment processes.
- Collaborated with cross-functional teams to deliver cloud-based data solutions on GCP and Kubernetes.
- Implemented monitoring and alerting for data pipelines, ensuring reliability and compliance with business requirements.
- Mentored junior engineers in distributed systems and data engineering best practices.

Sigmoid   Jan 2019 – Oct 2020
Software Development Engineer I   *Bangalore*

- Built planet-scale streaming pipeline supporting 100k req/sec, 1GB/sec throughput presented at Apache Con NA 2020, leveraging distributed systems and cloud technologies.
- Spearheaded implementation of BigQuery-based NitroDB achieving 30% cost reduction while maintaining performance and compliance.
- Designed and deployed multi-tenant search feature with Elasticsearch integration, delivering responsive experience for multiple organizations.
- Contributed critical bug fixes to spark-snowflake as committer, ensuring data correctness at scale and collaborating with global teams.
- Worked closely with business and engineering teams to deliver scalable, reliable data pipelines on GCP and Kubernetes.
- Mentored new hires on distributed data engineering and compliance best practices.

Capgemini   Jun 2015 – Jan 2019
Consultant P2   *Bangalore*

- Delivered enterprise data solutions for clients across multiple industries, focusing on distributed systems and cloud migration.

- Facilitated technical design sessions and translated complex business requirements into scalable architectures, emphasizing compliance and data governance.
- Mentored junior team members on software development and data engineering best practices, including distributed systems and cloud platforms.
- Implemented monitoring, alerting, and compliance frameworks for enterprise data pipelines.
- Collaborated with cross-functional teams to deliver high-impact data engineering projects on AWS and Azure.

PROJECTS

Azure Data Governance Platform | *Python, Azure, Data Architecture*
- Architected and delivered an enterprise-wide data governance solution for privacy and compliance.
- Consulted with internal teams to optimize data scanning performance by up to 20x, removing critical bottlenecks.
- Designed and deployed scalable APIs for data asset management, enabling robust data governance across services.
- Established a production-grade deployment and monitoring framework, ensuring high reliability for a mission-critical system.

Config-Driven Data Processing Framework | *Spark, Scala, GCP, Architecture, CI/CD*
- Designed and delivered a modular, rule-based data processing architecture for enterprise business services.
- Enabled flexible and maintainable data processing by creating a config-driven engine, reducing development overhead.

- Streamlined deployment workflows by developing a CI/CD automation tool for GitHub repository management.

End-to-End Streaming + Batch Data Platform | *Kafka, GCP, DataFlow, Java, Kubernetes, Architecture*
- Delivered an end-to-end design, build, and deployment of a planet-scale streaming platform for a major ad-tech client.

- Architected the solution to support 100k req/sec and 1GB/sec throughput, productionalizing a high-impact customer use case.
- Engineered a resilient, auto-scaling architecture on GKE with a circuit breaker pattern to ensure 99.99% uptime.

Cloud Data Warehouse Optimization | *Apache Spark, BigQuery, Performance Tuning*
- Led a client project to migrate and optimize a large-scale data warehouse to BigQuery, achieving a 30% cost reduction.
- Consulted on performance, implementing advanced algorithms (t-digest) to deliver 10-100x query performance improvements.

Multi-Tenant Search Solution | *Elasticsearch, System Design, Architecture*
- Designed and implemented a secure, multi-tenant search architecture for a client's SaaS platform.
- Delivered a responsive and scalable search solution by integrating Elasticsearch with the core data platform.

Spark-Snowflake Data Integration | *Apache Spark, Snowflake, Python, Data Migration*
- Guided a client's data integration strategy by contributing critical bug fixes to the open-source spark-snowflake connector.
- Designed and built a robust ETL pipeline for a 3rd party data migration from GCS to Snowflake.
- Provided consulting on Snowflake administration for cost management, security, and performance.

EDUCATION

Indian Institute of Technology (IIT) Kharagpur	2011 – 2015
Bachelor of Technology, Industrial & Systems Engineering	
Sri Chaitanya Jr College, Gudavalli	2009 – 2011
XII, Board of Intermediate Education	
Gowtham Concept School, Gudivada	2008 – 2009
X, Board of Secondary Education	

SKILLS

Core Technologies : Python, Scala, SQL, Apache Spark, Spark Internals, Kafka, Kubernetes, Docker
Cloud & Data Platforms : AWS, Azure, GCP, Databricks, Snowflake
Methodologies & Practices : CI/CD, MLOps, Agile, Technical Project Delivery
Key Competencies : Data Engineering, Distributed Systems, System Design & Architecture, Architecture Consulting, Client Management, Data Privacy & Compliance, Leadership & Mentoring, Problem Solving