Ravishek Kumar Ranjan

Bengaluru, India · ravishekr7@gmail.com · +91-6205194383 · LinkedIn

Summary

Experienced **Software Engineer with 3+ years of experience** in designing and delivering scalable, cloud-native data and backend systems. Specialized in building high-availability, distributed data integrity services across **databases, Kafka, Iceberg, and AWS S3**. Proven expertise in **asynchronous API design, event-driven remediation workflows, and large-scale job orchestration using AWS Lambda, SQS, and Step Functions**. Skilled in optimizing **Iceberg metadata, reducing Athena query latency,** and improving cost efficiency. Adept at ensuring data consistency, observability, and resilience in multi-tenant, mission-critical environments. Strong foundation in **data engineering, backend development, and AWS architecture** with a track record of solving complex technical challenges and delivering measurable impact. Passionate about exploring and learning emerging technologies, particularly in the field of Artificial Intelligence, and applying them to build innovative, impactful solutions.

Skills

Programming Languages:	Python, C++, JavaScript, Scala
Frameworks/Tools:	Apache Spark, Apache Kafka, React, Bootstrap, Delta Lake, Presto
Databases:	PostgreSQL, MySQL, MongoDB, DynamoDB, Iceberg
Cloud & Infra:	AWS (S3, EMR, Athena, Lambda, API Gateway, Step Functions, Glue, SQS, SNS, CloudWatch)
Other:	Data Structures & Algorithms, CI/CD, Git, Docker

Work Experience

Guidewire — Bengaluru, India
Software Engineer 2 — Nov 2023 – Present

- Diagnosed and fixed **small file issues in Iceberg** tables (causing metadata bloat), by writing **Spark compaction jobs** using **AWS Glue notebooks** - resulting in lower Athena latency and reduced cost.
- Designed and implemented a cloud-native service to proactively **detect and remediate data gaps** and inconsistencies across large-scale, **mission-critical data pipelines—integrating databases, Kafka, and Iceberg—to safeguard** end-to-end data integrity and reliability across distributed systems.
- Led design of **asynchronous remediation APIs** using API Gateway, Lambda, and Step Functions to decode and republish **lost Kafka messages**, with **observability via Datadog** logging.
- Enhanced an existing orchestration framework to **dynamically manage and dispatch DataGap jobs across 200+ environments** in a region—using SQS and AWS Lambda for event-driven execution with **customizable concurrency limits—optimizing resource utilization, minimizing job overlap**, and ensuring timely hourly validation at scale.
- Engineered and rolled out a **solution to handle data mismatches caused by Point-in-Time Recovery (PITR) in serverless**, distributed systems—where partially processed transactions (LSNs) become invalid post-rollback—by building Athena-driven verification workflows to **detect and discard stale data, propagate invalidation confirmations across all consuming systems**, and prevent false data gap validations, **ensuring post-recovery consistency.**

Tekion Corp — Bengaluru, India
Associate Software Engineer — Jan 2022 – Sept 2023

- Developed **Spark-based batch and streaming ingestion pipelines** to build and maintain **fact, dimension, and aggregate tables** from diverse sources—including **relational databases, Kafka, and Kinesis**—using checkpoint-based incremental loads for **batch jobs and real-time processing for streams**; architected a **Delta Lake data warehouse on AWS S3** to ensure data integrity, high performance, and scalability.
- Partnered with Machine Learning, Analytics, and Reporting teams to deliver curated fact, dimension, and aggregate tables in Delta Lake. Enabled **flexible access via APIs, Athena/Presto queries, and direct S3 integration for EMR** or custom processing—supporting use cases from model training to analytics dashboards and financial reporting.
- Optimized ETL processes by improving join strategies, refining DataFrame transformations, and **implementing efficient S3 partitioning** logic—reducing **processing time by 40%** and cutting **S3/Athena query costs by 30%**.

Projects

PRobot – CLI-based Smart PR Assistant *(Python, OpenAI API, Jira API, Confluence API, GitHub API)*

- **Developed a CLI tool** that integrates with Jira and Confluence APIs to gather development context and design details for code changes.
- Used OpenAI API with custom prompts to generate concise, **professional PR titles and descriptions**, followed by user confirmation before submission.
- Automated GitHub PR creation via API, **reducing PR preparation time by 70%** and improving review quality through richer, standardized documentation.

CHMP – Hospital Management Platform *(Go, ReactJS, PostgreSQL, Docker)*

- Contributed to an **open-source**, **full-stack** web application built for hospital reservation, financial reporting, and operational management.
- Created scalable backend APIs in Go and integrated new frontend features using **React.js**.
- Set up foundational project structure, handled **Docker-based** dev environments, and collaborated with engineers for rollout.

Achievements

- 6 Stars in Problem Solving on Hackerrank.
- Secured 226th rank in Winter of Codes by Coding Ninjas.
- **Go**ogle Code Jam, Facebook HackerCup, Codechef Snackdown - cleared prelims.
- Received Outstanding Internship Award at Protrainy for exceptional web development skills.

Education

Dayananda Sagar College of Engineering — Bengaluru, India
BE, Information Science Engineering — Aug 2018 – Jun 2022