

BHARATH THATI

Senior Data Engineer · AWS · Spark · dbt · Distributed & Cloud-Native Systems
Hyderabad, India | bharath2792@gmail.com | +91-8464942761 | bharaththati.com

PROFESSIONAL SUMMARY

Senior Data Engineer with 11+ years of experience designing and delivering enterprise-scale data platforms, distributed ETL systems and analytics solutions on AWS — across Fintech, Asset Management, Compliance and E-commerce domains. Deep expertise in cloud-native data lake and warehouse architectures, large-scale Spark workloads, and metadata-driven pipeline frameworks. Proven track record building dimensional data models, financial reporting systems integrating 60+ enterprise datasets, and event-driven ingestion platforms processing 100+ feed types. Experienced collaborating with cross-functional stakeholders to translate complex business requirements into scalable, audit-ready solutions. Actively exploring modern lakehouse and streaming platforms such as Iceberg, Hudi, Delta tables, Kafka, and Kinesis, along with Kafka-based processing patterns using Spark Streaming and Flink, along with Databricks and AI agents / agentic workflow systems.

TECHNICAL SKILLS

Cloud & Infra	AWS (S3, EMR, Redshift, Aurora PostgreSQL, Lambda, Step Functions, ECS, Glue, Kinesis, DynamoDB, Airflow, SNS, SQS, API Gateway, CloudFormation, CDK, IAM, VPC, CloudWatch)
Languages	Python (expert), SQL / Advanced SQL, PySpark, TypeScript, Node.js
Data Engineering	Apache Spark, Hive, dbt (200+ models), Airflow, EMR, ETL/ELT pipelines, metadata-driven frameworks, data lakes, dimensional data modeling
Warehousing & DB	Redshift, Aurora PostgreSQL, DynamoDB, MySQL, MPP/distributed systems
Lakehouse / Stream	Apache Iceberg, Apache Hudi, Kinesis, incremental & streaming processing
Architecture	Serverless, event-driven systems, platform engineering, infrastructure automation
Analytics & DevOps	QuickSight, CI/CD, Git, CloudWatch, Jira, on-call / production operations
AI & Agents	Claude Agent SDK, MCP (Model Context Protocol), LangChain, PGVector, agentic workflow design

PROFESSIONAL EXPERIENCE

Coforge (Client: Leading Global Asset Management Firm) · *Senior Data Engineer* Jan 2024 – Jun 2026

- Built and supported event-driven ingestion pipelines processing 100+ financial feed types using AWS Lambda, DynamoDB, ECS, Step Functions and S3.
- Maintained and enhanced a dbt warehouse solution comprising 200+ transformation models supporting financial reporting workloads.
- Developed orchestration workflows using AWS Step Functions and ECS for ingestion and transformation processes.
- Supported production systems through monitoring, troubleshooting and operational support.

Amazon · *Data Engineer* Sep 2017 – Aug 2023

- Architected and delivered a reusable AWS CDK construct library adopted across multiple data engineering teams for standardized infrastructure provisioning and Spark workload orchestration.
- Designed and built scalable ETL and data lake solutions using EMR, Spark, S3, Redshift, Airflow and Python.
- Developed a YAML-driven, metadata-based ETL framework supporting ingestion, transformation and movement of data across multiple enterprise systems and storage platforms.
- Built event-driven ingestion and reporting platforms using SNS, SQS, Lambda, Glue and S3.
- Designed dimensional data models and audit-ready financial and compliance reporting systems integrating 60+ enterprise finance datasets.
- Developed orchestration frameworks supporting Airflow-based and Step Function-based distributed workload execution at scale.
- Built centralized monitoring and operational tooling for tracking execution health, retries, failures and platform metrics across distributed compute environments.

- Worked extensively on Spark and Hive performance optimization, partitioning strategies and large-scale distributed data processing workloads.
- Collaborated with finance, accounting, compliance and business stakeholders to gather requirements and deliver enterprise-scale analytical solutions.
- Participated in architecture reviews, code reviews, deployment activities and platform engineering standardization.

Amazon · *Support Engineer / Lead Generation Analyst* Aug 2014 – Sep 2017

- Developed automation solutions for large-scale operational workflows, significantly reducing manual audit and data correction effort.
- Built internal productivity tools, dashboards and reporting systems improving auditor efficiency and operational visibility.
- Developed ML-assisted workflows and text classification models using Python and AWS SageMaker for automated attribute prediction and data enrichment.
- Worked on large-scale web scraping, deduplication and seller identification workflows for lead generation and onboarding systems.

NOTABLE PROJECTS

Personal Finance & Habit Tracker · bharaththati.com

- Designed and self-hosted a multi-user web application (Node.js / Express / SQLite) running on Oracle Cloud Always Free, served behind a Caddy reverse proxy with auto-TLS — covering habit tracking, expense management and to-do reminders with multi-channel notifications (ntfy.sh, Telegram).
- Applied production-grade security: bcrypt password hashing, constant-time login, per-IP rate limiting, account lockout, per-user data isolation via FK-scoped queries, and HTTPOnly/SameSite session cookies.

AI Agent Engineering — Claude Agent SDK

- Designed and deployed three role-specialized AI coding agents (developer, reviewer, architect) using the Claude Agent SDK, applying least-privilege tool scoping per agent and task-based model-tier selection (Sonnet for execution, Opus for design judgment).
- Integrated reusable agent capabilities via Skills and connected external systems through the Model Context Protocol (MCP), with structured verification and sandbox testing workflows.

FORWARD-LOOKING TECHNOLOGIES

Apache Iceberg · Apache Hudi · Streaming Architectures (Kinesis, Kafka + Spark Streaming/Flink) · Modern Lakehouse Design · AI Agents & Agentic Workflow Systems

EDUCATION

B.Tech, Mechanical Engineering · NIT Surat · 2014