

Eegapuri Lokeshwar Reddy

Data Engineer

9959242262 ◇ elokesh4292@gmail.com ◇ Telangana, India ◇ [LinkedIn](#) ◇ [GitHub](#) ◇ [Coursera](#)

SUMMARY

Experienced Data Engineer with 3 years in ETL operations, data pipelines, and enterprise data warehousing. Expertise in monitoring, troubleshooting, optimizing, and automating ETL workflows to ensure seamless data processing and system performance in production and disaster recovery environments. Practical understanding of dimensional modeling gained through EDW experience. Strong understanding of OLAP, data marts, data warehousing, and workflow orchestration. Passionate about big data technologies and cloud computing. Seeking an opportunity to apply operational expertise and development skills to design, build, and optimize scalable data engineering solutions.

WORK EXPERIENCE

System Engineer

Apr '19 — Feb '22

Tata Consultancy Service (TCS)

Hyderabad, India

Project: STC (Saudi Telecom Company) – Enterprise Data Warehouse (EDW) & Confluent Kafka Streaming

Environment: Ab Initio, Confluent Kafka, Teradata, Red Hat Linux, Jira, Remedy, SQL, Shell Scripting, Python

Key Responsibilities:

Enterprise Data Warehouse (EDW) Operations

- Managed and optimized ETL workflows in OLAP, Datamart, and BI environments to ensure smooth data processing.
- Proactively monitored key EDW and database performance metrics via Teradata Viewpoint, while automating server-level monitoring through shell scripts for user connectivity, system availability, and resource thresholds, enhancing operational efficiency.
- Executed deployments of new ETL jobs, including scheduling, database object deployments, and server-level file system adjustments, ensuring seamless integration of developer-provided components and scalability for future data growth.
- Managed ETL job updates (optimizations, refinements, decommissioning) and ensured seamless integration via thorough connectivity testing.
- Performed pre- and post-deployment validations, ensuring database changes and job scheduling adjustments were aligned with system capacity and business needs.
- Maintained SLA compliance by optimizing job execution order and ensuring seamless database changes without disrupting ongoing operations.
- Developed optimized SQL queries to monitor daily KPI data, ensuring SLA compliance; generated and distributed discrepancy/resolution reports to teams.
- Facilitated knowledge-sharing sessions for new team members, ensuring smooth onboarding and project understanding.

Data Ingestion & Transformation

- Oversaw pull (SFTP, SQL extraction) and push (direct file transfers) ingestion methods from various sources, including CRM, IRB, and CDR systems.
- Implemented data quality checks through automated scripts, resolving data integrity issues (duplicates, missing data, format inconsistencies) for real-time and batch processing.
- Developed automated monitoring jobs to check file arrival status and query database KPIs (Postpaid, Prepaid, Sales, and Landline data), sending real-time email alerts in case of issues.

Incident Management & Issue Resolution

- Implemented proactive monitoring of system performance, job failures, and SLA breaches, leveraging Jira & Remedy ticketing systems for issue tracking and resolution.
- Managed database downtime issues, troubleshooting problems such as locks, CPU spikes, file system delays, and collaborated with teams for rapid recovery. Served as the primary point of contact for incident communication, providing clear and timely updates to clients and internal teams.
- Identified and resolved data rejection issues in ETL jobs caused by transformation errors, working closely with developers to implement fixes and automate data reprocessing, ensuring 100% data completeness.

Disaster Recovery (DR) & Drills

- Assisted in setting up and testing DR environments, ensuring minimal downtime and data integrity during DR drills.
- Created comprehensive documentation of all ETL workflows, including data sources, processing schedules, dependencies, and storage requirements, aiding smooth DR execution.
- Managed DR drill communication, ensuring clear and timely updates to all stakeholders.
- Performed post-DR data integrity validations, ensuring data consistency after failover.

Kafka Streaming Operations

- Managed Kafka streaming operations for 6 months, alongside data warehouse workflows, focusing on monitoring, scheduling, and administration.
- Developed Python scripts using Kafka client libraries to automate health checks, monitor consumer lag, and send alerts.
- Handled Avro and JSON message formats in Kafka topics to ensure reliable serialization and schema compliance across producers and consumers.
- Administered Kafka topics, retention policies, and consumer groups, troubleshooting and resolving performance bottlenecks in streaming pipelines.

EDUCATION

B-Tech in Mechanical Engineering , CVR College of Engineering, Hyderabad (GPA: 9.56)	2015 — 2019 Hyderabad, India
Intermediate in MPC , Narayana Junior College (GPA: 9.7)	2013 — 2015 Hyderabad, India

CERTIFICATIONS

- Confluent Fundamentals Accreditation**, Confluent
- Academy Accreditation - Databricks Fundamentals**, Databricks
- dbt Fundamentals**, DBT
- IBM Data Engineering Professional Certificate**, IBM (In Progress)

ACHIEVEMENTS & AWARDS

On the Spot Award for outstanding performance and exceptional contributions to ETL operations and system optimization. Received **client appreciation** for ensuring seamless system operations and maintaining 100% KPI availability without SLA breaches.

Demonstrated strong coding skills by clearing the **first round of TCS CodeVita Season 7 (2018)**, a global coding competition.

PERSONAL PROJECTS

E-Commerce Data Pipeline | Data Engineering Project [Link](#)

- Tech Stack:** Apache Airflow, MySQL, PostgreSQL, FDW, Python, Pandas, SQL, Ubuntu, LucidChart
- Key Contributions:**
 - Developed an ETL pipeline to automate data ingestion from CSV files into MySQL and PostgreSQL.
 - Used Apache Airflow to orchestrate tasks such as file extraction, database loading, and data transformation.
 - Implemented PostgreSQL FDW (Foreign Data Wrapper) to enable cross-database queries between MySQL & PostgreSQL.
 - Optimized query performance and ensured secure authentication using .my.cnf and .pgpass.
 - Exported processed data as CSV files for further analysis.
- Outcome:**
 - Built a scalable and reusable data pipeline that demonstrates industry-standard ETL workflows.
 - Improved data integration across multiple databases, enhancing data accessibility for analytics.

Telecome Data Pipeline | Data Engineering Project [Link](#)

- Tech Stack:** Apache Airflow, Apache Spark (PySpark), PostgreSQL, Shell Scripting, Python, Ubuntu, LucidChart
- Key Contributions:**
 - Designed & implemented a batch ELT pipeline to process large-scale CDR (Call Detail Records) data for fraud detection.
 - Developed PySpark transformations to detect fraudulent patterns like high-frequency calls, short-duration international calls, and excessive customer service interactions.
 - Orchestrated ETL workflows using Apache Airflow, automating file extraction, transformation, database loading, and logging.
 - Stored raw & processed data in PostgreSQL, ensuring structured storage, auditing, and future analytics compatibility.
 - Implemented logging & error handling mechanisms, tracking processing metadata and failures for debugging.
 - Optimized batch processing performance by staging transformed data in local storage before loading into PostgreSQL.
- Outcome:**
 - Built a scalable & modular fraud detection pipeline, capable of handling millions of CDR records efficiently.
 - Enhanced debugging & monitoring through audit tables & detailed logging mechanisms.
 - Enabled future ELT enhancements, supporting DBT integration for advanced transformations and potential real-time processing using Kafka.

Cubon | Game Development Project [Link](#)

- Tech Stack:** Unity, C#
 - Developed a basic PC game as a self-learning project, applying programming logic, problem-solving, and game development principles.
 - Designed and implemented custom animations, player controls, and interactive game mechanics.
 - Built game logic and UI components, demonstrating coding proficiency and creative problem-solving skills.

3D Animation & Modeling Project [Link](#)

- Tech Stack:** Blender, Unity
 - Created and animated 3D models using Blender, applying texturing and rendering techniques.
 - Integrated models into a game engine for visualization, demonstrating creativity and technical adaptability.

SKILLS

Tools/Platforms: Apache Airflow, Ab Initio, Confluent Kafka(Client Library Integration), Apache Kafka, Hadoop(HDFS, YARN), Apache Spark(SparkSQL, PySpark), Dimensional Modeling, Data Modeling, Entity-Relationship Diagram(ERD), Data Warehousing, JIRA, Remedy, Red Hat Linux, Ubuntu, Databricks, Git, CI/CD, Looker Studio

Programming Languages: Python, SQL, Shell/Bash Scripting

Libraries/Frameworks: Pandas, NumPy, Requests, BeautifulSoup, SQLAlchemy, DB API

Data Repositories: Teradata, PostgreSQL

Cloud Tools & Services: AWS S3, AWS Glue(Learning)

Soft Skills: Problem Solving, Communication, Analytical Thinking, Adaptability, Attention to Detail, Risk Management, Decision-Making, Emotional discipline, Patience, Reliability, Knowledge Sharing

Interests Big Data and Cloud Technologies, Distributed Systems, Devops

Languages English, Hindi, Telugu