GOKUL R   AWS DATA ENGINEER

• +1 (647) 957-5025 • +91 9943316256 • gokul.227.r@gmail.com • www.linkedin.com/in/gokul-07

PROFILE SUMMARY :

• 6+ years of experience in AWS Data Engineering, Cloud Architecture & Data Solutions, specializing in ETL development, data migration, large-scale data processing across Healthcare, Insurance, Retail domains.

• Expertise in designing scalable, cost-effective, high-performance data architectures, ensuring data reliability, security & compliance.

• Strong experience in building, optimizing serverless and containerized data pipelines, handling batch & real-time data ingestion, transformation & orchestration.

• Skilled in data lake implementation, incremental data loading, schema evolution & performance tuning to enhance cloud-based data platforms.

• Proficient in CI/CD automation, infrastructure as code (IaC), and workflow orchestration, streamlining cloud deployments and ensuring fault-tolerant, highly available solutions.

• Strong problem-solver and collaborator, working with cross-functional teams to automate workflows, optimize cloud infrastructure, and drive business insights through data-driven solutions.

KEY SKILLS :

AWS: S3, Athena, DMS, Glue, Data Pipeline, Aurora, RDS, DynamoDB, DocumentDB, Lambda Function, API Gateway, Fargate, Parameter Store, Code Pipeline, CloudFormation, IAM, CloudWatch, Secret Manager, EventBridge, VPC, EC2, EMR, ECR, ECS, EKS, SNS, SQS, SES

PROGRAMMING LANGUAGES: Python, PySpark

DATABASES: MySQL,PostgreSQL, MS SQL Server, Oracle, DB2, MongoDB

ETL TOOLS: DMExpress Syncsort, Informatica IICS.

CERTIFICATIONS :
- AWS Cloud Practitioner
- AWS Data Engineer Associate
- AWS Developer Associate
- AWS Solutions Architect Associate

PROFESSIONAL EXPERIENCE :

Process Manager   **OCT 2024 – PRESENT**
Eclerx, Bangalore, India

Project : Customer Identity and Access Management (CIAM) | **Client :** COX

Description: Migrated CIAM datasets from Oracle, Adobe and Splunk to AWS to building functionalities like user registration, authentication, authorization, and profile management.

Technologies: AWS, PySpark, Python, SQL, Terraform, Jenkins, JIRA.

- Designed and implemented a unified data model for CIAM functionalities.
- Conducted POCs to optimize ETL performance using AWS Glue and EMR.
- Developed PySpark-based Glue ETL pipelines for identity data processing.
- Built incremental load strategy to maintain 13 months of historical data.
- Automated AWS infrastructure deployment using Terraform.
- Integrated SES for job status notifications and error alerts.
- Streamlined CI/CD process using Jenkins for production releases.

Project : Health Scorecard Automation | **Client :** COX
Description: Automated manual Excel-based dashboards by integrating multiple data sources into AWS.
Technologies: AWS, PySpark, Python, SQL, Terraform, Jenkins, JIRA.

- Designed and developed an automated data pipeline for dashboard generation.
- Replaced manual Excel-based calculations with scalable PySpark ETL jobs.
- Implemented Athena-based storage for computed business metrics.
- Developed incremental load processes to refresh monthly data.
- Built Glue workflows to orchestrate ETL, crawler, and reporting jobs.
- Automated infrastructure provisioning using Terraform.
- Deployed scripts via Jenkins CI/CD to maintain development standards.

Solutions Engineer   AUG 2022 – MAR 2023
Aligned Automation, Bangalore, India

Project: GT Connect | **Client:** Lilly
Description: Migrated a legacy healthcare application (AMT) to AWS, rebuilding key functionalities like File Upload, New Transaction, Search Transaction, User Management.
Technologies: AWS, PySpark, Python, SQL, GitHub, JIRA
- Analyzed legacy AMT application and redesigned architecture in AWS.
- Conducted POCs for data ingestion using Glue ETL and ECS Fargate.
- Developed PySpark ETL jobs for file validation and ingestion into Postgres.
- Created APIs for new transaction processing and search functionalities.
- Implemented SES notifications for job completion alerts.
- Built monitoring and logging framework using DynamoDB and CloudWatch.
- Deployed infrastructure using CloudFormation templates and CI/CD pipelines.

Project: LRL Learning & Development – Data Ingestion | **Client:** Lilly
Description: Modernized legacy learning management system by ingesting structured & semi-structured data into AWS.
Technologies: AWS, PySpark, Python, SQL, GitHub, JIRA
- Designed AWS architecture, ER models, and data flow diagrams.
- Developed PySpark scripts for data ingestion and cleansing in Glue ETL jobs.
- Built Lambda functions to trigger Glue jobs via EventBridge based on S3 events.
- Implemented API Gateway to expose conformed data for external applications.
- Automated infrastructure setup using CloudFormation templates.
- Deployed ETL scripts via CodePipeline and ensured production stability.

Associate Business Analyst   OCT 2021 – AUG 2022
Infosys, Bangalore, India

Project: PME Database Archival Solution | **Client:** Nike
Description: Developed an AWS based data archival and purge solution from on-premises databases.
Technologies: AWS, PySpark, Python, GitHub, Terraform, Jenkins
- Designed architecture for archival, purge, reporting, and restoration workflows.
- Conducted POCs on AWS services for optimal archival strategies.
- Developed PySpark-based Glue jobs for data migration and cleansing.
- Automated AWS Glue workflows with triggers, notifications, and logs.
- Built Lambda functions to integrate API Gateway, Glue, and Fargate-based purge framework.
- Deployed Terraform scripts for infrastructure setup across environments.
- Automated CI/CD pipelines in Jenkins for seamless deployment.

Application Developer   DEC 2019 – OCT 2021
Artech Infosystems – Client IBM, Bangalore, India

Project: History Data Migration | **Client:** Blue Cross Blue Shield of Massachusetts (BCBSMA)
Description: Developed an AWS based data archival and purge solution from on-premises databases.
Technologies: Mainframe (JCL, COBOL, DB2), ETL Tool (Syncsort), AWS (S3, Athena, Glue)
- Analyzed Mainframe files, gathered job status, and mapped file structures.
- Developed Syncsort ETL jobs to transform and extract data in CSV format.
- Performed UAT to validate Syncsort outputs against Mainframe files.
- Configured AWS S3 for storage and created Athena tables for query execution.
- Designed migration workflows and participated in client walkthroughs.

Software Engineer Intern   DEC 2018 – DEC 2019
iPrimed Education Solutions, Bangalore, India
- Assisted in software development and participated in various projects.

EDUCATIONAL QUALIFICATION :
- Bachelors of Technology in Information Technology   JUL 2015 – APR 2019

ADDITIONAL INFORMATION : Work Authorization / Visa Sponsorship Required