Prashanth Chinthapoola
Data Engineer
Dublin, Ireland D03 Y7R0 | +353899616493
prashanth.chinthapoola@gmail.com | LinkedIn

PERSONAL SUMMARY

Experienced Data Engineer and AWS Developer with over 5 years of expertise in designing, developing, and maintaining scalable data pipelines and ETL/ELT processes. Proficient in AWS technologies (Glue, Redshift, Lambda, S3) and Azure tools (Data Factory, Synapse, Databricks). Skilled in Python, PySpark, SQL, and Java for data transformation and warehousing. Adept at automating workflows with Terraform, CI/CD tools (GitLab, Jenkins), and cloud-based migration processes. Proven track record of optimizing query performance, troubleshooting complex data issues, and delivering secure, efficient solutions.

CORE SKILLS

- **Programming Languages**: Python, PySpark, SQL, Java, Shell Scripting
- **Cloud Platforms**: Amazon Web Services (AWS), Azure
- **Databases**: Redshift, Oracle, MySQL, Snowflake, PostgreSQL, RDS, Microsoft SQL Server
- **CI/CD Tools**: GitLab, GitHub, Jenkins
- **Workflow Orchestration**: Airflow, Kubernetes, Docker
- **AWS Services**: Lambda, Redshift, S3, Glue, Kinesis Firehose, Athena, API Gateway, CloudWatch, AWS RDS, Cloud9, SNS
- **Azure Services**: Azure Databricks, Azure Data Factory, Azure Data Lake, Azure Blob
- **Data Science & Statistics:** Pandas, NumPy, Scikit-Learn, Exploratory Data Analysis
- **Machine Learning**: Data Processing, Feature Engineering, Model Deployment
- **Data Visualization & Analytics:** Tableau
- **DevOps & Automation**: Terraform, CI/CD (GitLab, Jenkins), Monitoring (CloudWatch)
- **Development & Collaboration Tools**: IntelliJ, Visual Studio Code, Putty, JIRA, ServiceNow, SSIS, ETL Tools, Agile Environment, Linux

---

WORK EXPERIENCE

AWS Developer
Broadridge | Hyderabad
Jan 2022 - Sep 2022

- Implemented a robust AWS infrastructure to streamline operations, designing system architecture based on client requirements.
- Leveraged AWS services like Lambda, Redshift, and S3 for efficient data extraction, transformation, and processing.
- Developed a serverless data pipeline using DynamoDB, Lambda, API Gateway, Elasticsearch, and Redshift, ensuring seamless cross-account data replication and real-time data changes via DynamoDB triggers.
- Implemented data pipeline orchestrations using Airflow, improving efficiency and reliability
- With strong Python skills, optimized data retrieval and SQL query performance.
- Managed infrastructure as code using Terraform, automating deployments through GitLab and Jenkins while enhancing operational efficiency and pipeline performance.

- Also worked as support for Java project for cross-account data transfer.

Data Engineer

OPTUM | Hyderabad

Sep 2019 - Dec 2021

- Developed ADF pipelines to fetch data from various sources and orchestrated data movement across staging, agnostic, and specific layers using Synapse.
- Built data layers with PySpark in Azure Synapse and created views on Azure Data Lake parquet files for downstream applications like Power BI.
- Automated the migration of 5TB of on-premises data using Azure AzCopy and PowerShell, migrating Oracle data to Snowflake.
- Designed technical architecture for large-scale migration, leveraging Python (Pandas, NumPy) for data profiling and Azure Functions for complex file handling.
- Migrated Informatica jobs to PySpark in Azure Databricks and built efficient data pipelines with ADF.
- Backend Development: Developed custom data processing APIs, integrating with multiple AWS services for real-time analytics.
- Decision-making & problem-solving: Led data pipeline troubleshooting and performance tuning, ensuring 99.9% system uptime.

Data Engineer

Amazon | Hyderabad

May 2018 - Sep 2019

- Developed a Python-based extraction app to capture CDC data from sources and load it into S3.
- Designed the Data Warehouse architecture on Redshift and implemented data transformations and complex business logic using PySpark.
- Utilized AWS Glue and Spark libraries to apply business rules on extracted parquet data before transferring it to Redshift.
- Additionally, performed performance tuning by analyzing execution plans, focusing on optimizing factors like garbage collection and shuffle write processes.
- Utilized AWS CloudWatch, SNS, and Step Functions to monitor, automate, and orchestrate complex workflows, enhancing system reliability and operational efficiency.

EDUCATION

Dublin Business School

Master | Data Analytics

Sep 2022 - Jan 2024

Anurag University

Bachelor of Electronics & Communication | Engineering

Oct 2014 - Dec 2018

REFERENCES

Available upon request