|  |  |
| --- | --- |
| M Suresh  GCP Big Data Engineer |  |

| PH: +91-9182705808| suresh.machavarapu91@gmail.com | <https://www.linkedin.com/in/suresh-m-042469230/>

# Summary

Results-driven GCP Data Engineer with 3 years and 5 months of hands-on experience in designing, building, and optimizing scalable data pipelines and workflows for advanced analytics. Proven expertise in leveraging Google Cloud Platform services such as GCS, Dataproc, BigQuery, and Juniper to develop robust, data-driven solutions. Adept at implementing cost optimization strategies, ensuring efficient resource utilization. Skilled in integrating diverse cloud services to maintain data integrity, accessibility, and performance across analytics platforms. Committed to delivering high-quality, scalable solutions that drive business insights and decision-making.

# Experience

## *GCP DATA Engineer | Capgemini Technology Services Ltd | MAR 2022 – Present*

**Project: OPS ANALYTICS (HSBC CLIENT)**

**Role: GCP Data Engineer (MAR 2022 – May 2024)**

* Created and optimized BigQuery views, PL/SQL scripts, implementing partitioning and clustering strategies to enhance query performance.
* Developed PySpark jobs to efficiently transfer and transform data from GCS (for backup) to BigQuery for analytical processing.

**Role: DevOps Engineer (June 2024 – Present]**

*Pipeline1: On-Prem to GCP Big Query Migration:*

* Scheduled and orchestrated on-premise data migration jobs using **Control-M**, transferring data from Windows NAS to Juniper NAS.
* Designed and developed **Juniper data pipelines** to ingest data from on-prem sources (MSSQL, Oracle, Juniper NAS) into Google Cloud Storage (GCS) in Avro format for backup and subsequently published to **BigQuery** for analytics.
* Automated the **BigQuery procedures** using **Juniper ChronoX**.

Win NAS

Juniper NAS

GCS (backup) BigQuery

*Pipline2: GCS to GCP Big Query Migration:*

* Enabled secure data ingestion by configuring service account-based authentication for receiving Parquet/CSV files from source teams into Google Cloud Storage (GCS).
* Orchestrated PySpark jobs and BigQuery procedures using Cloud Composer, ensuring reliable and automated data workflows.

Azure/AWS

GCS (Backup)

BigQuery

* Exported and analyzed **log data in BigQuery** to support **cost optimization and usage analysis**.
* Created **shell script** to automate the service account key auto rotation and store it in secret manager.
* Developed a **generic Python script** to automate the scheduling of **BigQuery procedures** using **Juniper ChronoX**.
* Created **change orders** to release production deployments.
* Documented processes using **confluence** and also created **jira templates** to raise work requests.
* Setup **Virtual environments** on control-m server for python scripts.
* Created **Python scripts** as per requirements like getting gcs bucket details, to list bq datasets, tables and columns across the GCP project.

# Education

## PH.D. in TRANSPORTATION | JaNUARY 2021 | NIT AGARTALA, TRIPURA

## M-Tech in TRANSPORTATION | SEPTEMBER 2015 | NIT Hamirpur, HIMACHAL PRADESH

# Skills & Abilities

|  |  |
| --- | --- |
| * PL/SQL, Shell scripting, Pyspark, Spark SQL, Python, Linux commands * Big Query Data Warehouse for Analytics * Data Lake (GCS) * Dataproc for spark jobs * Logging, BQ Sinks, DLP API, Git, GitHub, WinSCP * Interact with GCP Services via UI, CLI, Python client lib, API * I have knowledge on Composer, Vertex AI, looker studio, Data Fusion ETL, Cloud Functions |  |

# Certifications

* GCP Associate Cloud Engineer Certification.
* GCP Professional Data Engineer Certification.