

Machavarapu Suresh

GCP Data Engineer

✉ suresh.machavarapu91@gmail.com ☎ +919182705808 🇮🇳 India 🌐 [Suresh M](#)



WORK EXPERIENCE

Capgemini Technology Services Ltd

MAR 2022 – Present

- Results-driven GCP Data Engineer with 4 years of hands-on experience in designing, building, and optimizing scalable data pipelines and workflows for advanced analytics.
- Adept at implementing cost optimization strategies, ensuring efficient resource utilization.
- Committed to delivering high-quality, scalable solutions that drive business insights and decision-making.
- Developed and optimized BigQuery PL/SQL scripts, views.
- Built PySpark jobs for efficient data migration from Google Cloud Storage (GCS) to BigQuery.
- Created Python scripts to interact with GCP services.
- Scheduled and orchestrated on-premises data migration jobs using Control-M software.
- Designed and Automated Juniper data pipelines to seamless data ingestion and analytical processes.
- Setup GCP linux VM for GCP recommendations app and interact with Gemini GEN AI model for responses.
- Developed GCP looker dashboard for strategic data feed (SDF)

EDUCATION

M-Tech in TRANSPORTATION | SEPTEMBER 2015 | NIT Hamirpur, HIMACHAL PRADESH | 8.16

Ph.D. in TRANSPORTATION | JANUARY 2021 | NIT Agartala, Tripura | 9.31

PROJECT

MAR 2022 – Present

OPS ANALYTICS - HSBC CLIENT

- Developed and optimized **big query PL/SQL scripts, views**, and complex queries to improve performance and reduce cost.
- Developed **PySpark jobs** to migrate data from GCS to big query (Analytics)
- Created **Python scripts** as per requirements like getting GCS bucket details, to list BQ datasets, tables and columns across the GCP project, etc.
- Designed and orchestrated scalable ETL workflows using Google Cloud Composer (Apache Airflow), improving data pipeline reliability and automation.

DBS REFINARY ANALYTICS - HSBC CLIENT

- **Designed Control-M Jobs** to schedule on-premises data migration jobs, transferring data from Windows NAS to Juniper NAS/SQL DB/Oracle DB.
- Created **Juniper data pipelines** to ingest data from on-prem sources (MSSQL, Oracle, Juniper NAS) into GCS and subsequently published to BigQuery for analytics.
- Gathered business requirements and Developed **BQ Scripts** as per the requirements.
- Automated the **BigQuery Scripts** using Juniper ChronoX that triggers scripts using a generic Python Script.
- Implemented **Cloud Build CI/CD pipeline** to deploy the BQ Scripts & resources/service account key auto rotation and store it in secret manager.
- Setup **GCP Linux VM** for GCP recommendations app and interact with **Gemini GEN AI model** for responses.
- Implemented **change orders** for production deployments.
- Documented processes using **confluence** and created **Jira templates** to raise work requests.

SKILLS

Languages: PL/SQL, PySpark, Linux commands

Databases: BigQuery, Cloud SQL

GCP Services: GCS, Dataproc, Composer (Airflow), Vertex AI (Gemini Model), looker studio, ETL, Logging, BQ Sinks, IAM.

Others: GitHub, WinSCP, Confluence, Jira

Achievement

GATE Rank, Certifications: Associate Cloud Engineer and GCP Professional Data Engineer.