

Mahesh Thummalapalli

Senior Consultant

Profile

Mahesh is an experienced Data Engineer with 11+ years of demonstrated experience working in the information technology and services industry. Strong technical skill in both AWS and GCP, coupled with a proactive ability to overcome technical challenges and excellent communication and interpersonal skills allow Mahesh to effectively solve complex business problems.

Mahesh is a Certified GCP Google Cloud Professional Data Engineer and Cloud Architect who holds experience across a broad range of industries including Retail, Financial Services, Media, Automobile and Education. Mahesh has proven expertise in Data Engineering, Data Migration and Solution Design.

Mahesh stands out as an innovative and creative thinker, possessing impressive problem-solving competencies and incisive analytical abilities. He has a solid work ethic and exceptional interpersonal skills, plus the capacity to tackle the most intricate software challenges.

Key Skills & Technologies

- AWS
- Google Cloud Platform
- Data Engineering
- Data Wrangling
- Data Modelling
- Data Visualisation
- Azure
- Strategic Planning
- System Integration
- Machine Learning
- Deep Learning
- Stakeholder Management
- Redshift
- DMS
- Cloud Watch
- SAP
- Cloud Composer
- Python
- SQL
- Kubernetes
- Pubsub
- Terraform
- Linux
- Unix
- TensorFlow
- Apache

Consultant Profile

Selected Project Experience

Major Retail Company – GCP Data Engineer

Company migrated existing Hadoop warehouse systems to Google BigQuery. ETL flows that were running on premise PySpark was replaced with DAGs and Pipelines in Airflow to read incremental data from Siebel,CPRV to BigQuery.

Roles and Responsibilities included:

- Designed the framework that can be used to extract data from Siebel server and load to GCP bucket
- Designed logical layers in BigQuery
- Designed partitioning strategy for large tables
- Designed and developed data pipelines to read data from Siebel server, extract as file, load to staging tables in BigQuery
- Built SQL scripts to load data from stage to target tables in BigQuery.
- Created a semantic layer consisting of views on top of datamart layer to expose to Looker

Leading US Media Giant – GCP Data Engineer

Company migrated existing Teradata warehouse systems to Google BigQuery. ETL flows that were running on premise Informatica were replaced with DAGs and Pipelines in Airflow to read incremental data from SQL server to BigQuery. Roles and Responsibilities included:

- Developed Kubernetes image in the framework used to extract data from SQL server and load to GCP bucket
- Used Kube Secrets to securely access on premise SQL servers and extract data for incremental loads
- Designed partitioning strategy for large tables
- Designed and developed data pipelines to read data from SQL server extract as file, load to staging tables in BigQuery.
- Built SQL scripts to load data data from stage to target tables in BigQuery.
- Scheduled jobs in cloud composer using Airflow using various DAG operators and Python scripts
- Implemented file sensors, to handle intra environment dependency with Autosys.

Leading US Educational Provider – GCP Data Engineer

- Requirements gathering and organising meetings with teams
- Built data pipelines to read data from files and load to staging tables in BigQuery as External tables.
- Built SQL scripts to load data to target tables and run them in pipeline
- Developed code to read data from streaming systems like Pub Sub and load to BigQuery
- Scheduled jobs in cloud composer using



Industry Experience

Retail

Financial Services

Media

Automobile

Education

Certification

Google Cloud Certified |
Professional Data
Engineer

Google Cloud Certified |
Professional Cloud
Architect

Microsoft Azure Certified |
Azure Data Associate

Bachelor – Technology,
Information Technology |
Pondicherry University,
India

