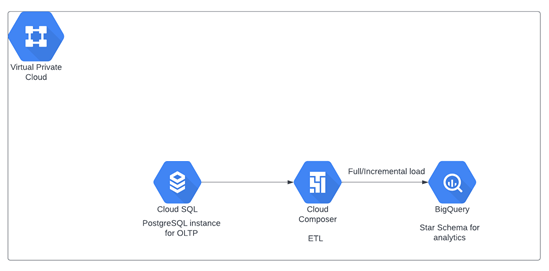
**Capstone Project**



**Process Flow:**

* Created a new Postgres instance
* Created OLTP tables in Postgres
* Created python script using faker library to generate and load initial data into OLTP tables (*OLTP\_Initial\_Load* script)
* Created staging and target tables into BigQuery
* Create an ETL Log table to record the batch run time everyday
* Created below 2 ETL scripts:

1. Created python script to Extract data from Postgres tables based on LastRunDate from ETL log table, transform as per star schema and Load into BigQuery Staging tables. (script name =LoadStaging.py)
2. Created python script to Load data from staging tables to Target BigQuery tables

(script name =LoadTarget.py)

* Create a DAG in composer to execute above 2 ETL scripts
* Created a new python script to generate incremental data of 5000 new orders (*IncrementalLoad* script)
* Ran the DAG again to load incremental data
* Prepared analytics queries