

## Self

Hi, I'm Mohith. I have 4+ years of IT experience, currently working in CGI. I have completed my masters (MSc) from MS Ramaiah University of Applied Sciences.

I have 3.5 years of experience in Snowflake development with DBT + Mule soft across

I have Worked on projects across Insurance and e-commerce domains. I'm proficient in building scalable data pipelines in Snowflakes, and handling large datasets in formats like CSV and JSON.

I've worked with Snowpipes, Streams, Tasks, time travel and COPY commands as well as developing automated controlled transformations using DBT.

I have worked with both waterfall and Agile methodologies.

---

## Project Overview:

I'm currently working on an insurance project to build a scalable cloud data platform using Snowflake that enables real-time analysis of insurance claims. Here we manage and process large volumes of operational data for analytics and reporting.

### Source to Destination Flow:

The data originates from multiple transactional systems—primarily MySQL and Oracle databases. We use Talend as our ETL tool to extract, transform, and load the data into AWS S3, which acts as our data lake.

### Snowflake Integration:

From AWS S3, we ingest data into Snowflake using Snow pipe.

We define external stages and file formats to handle various file types like CSV, JSON, and Parquet.

Snow pipe allows us to automate and streamline the data loading process with minimal latency.

## Here we follow Medallion Architecture Implementation:



### Bronze Layer – Raw/Staging Data:

"This layer stores raw data directly ingested from S3.

We use Snow pipe and External Tables to maintain the original structure for reprocessing."



### Silver Layer – Filtered and Cleansed Data:

"Here, we apply some transformations using DBT, Streams, Tasks, and Dynamic Tables.

This includes standardization, de-duplication, and enrichment to prepare data for analytics."



### Gold Layer – Business-Ready Data:

"This layer contains final datasets optimized for reporting and decision-making."

We use Views and Materialized Views to serve KPIs and dashboards to business users and analysts." this is pretty much about myself and work experience.

=====

-----My  
Responsibilities

I have been involved in

- 1 Team Collaboration (Agile)  
I actively take part in team meetings like daily check-ins, planning sessions, reviews, and feedback discussions to keep projects on track.
- 2 Automating Data Flow  
Setting up stages, file formats, and Snowpipe for automated data ingestion.
- 3 Designing Efficient Data Structures  
I create and improve database designs in Snowflake so they run fast and can handle large amounts of data.
- 4 I use tools like dbt, Streams, and Tasks to turn messy raw data into clean, useful information.
- 5 Building Reports  
I create views and dashboards that help others easily understand and use the data.
- 6 Tracking Work and Sharing Progress  
I manage tasks in Jira, make sure we meet deadlines, and show stakeholders how our data systems and dashboards work.

----- user stories

- 1.Claims Data Ingestion: claims data from MySQL to be ingested into Snowflake daily using COPY command,
- 2.Incremental Loads for Members: incremental member updates captured through Streams + Tasks, so that only changed records are processed, reducing cost and improving pipeline efficiency.
- 3.Historical Tracking (SCD Type-2): build streams & track changes
- 4.Data Quality Checks: dbt tests applied on staging and curated tables (unique, not\_null, accepted\_values), so that I can trust the accuracy of data used for analytics.
- 5.Environment Management with Cloning: perform zero-copy clone of the production claims dataset to test without impacting live data or incurring extra storage costs.

-----facts and  
dimensions

## Fact Tables

Fact\_Claims

Fact\_Claim\_Line

Fact\_Eligibility

Fact\_Encounters

Fact\_Billing

## ◆ Dimension Tables

Dim\_Member

Dim\_Provider

Dim\_Payer

Dim\_Diagnosis

Dim\_Procedure

Dim\_Facility

## Challenges:

- 1) Snowpipe is failing due to error records.
- 2) some queries takes long time for executing.
- 3) frequently input file columns are changed due to that snowpipe failed
- 4) while migrating data oracle to snowflake datatype issues