Tiger Analytics enabled a large US Healthcare provider to create a personnel 360-degree view with a powerful Cloud Ecosystem

Tiger Analytics developed a fully centralized, robust, standardized cloud ecosystem using AWS Data Lake and Snowflake Cloud Warehouse. It enabled the client to access advanced analytics, such as a holistic view of personnel, and in the process - making it easier to improve data quality and ingestion. The solution also proved extremely useful in reducing the data load on the cloud ecosystem.
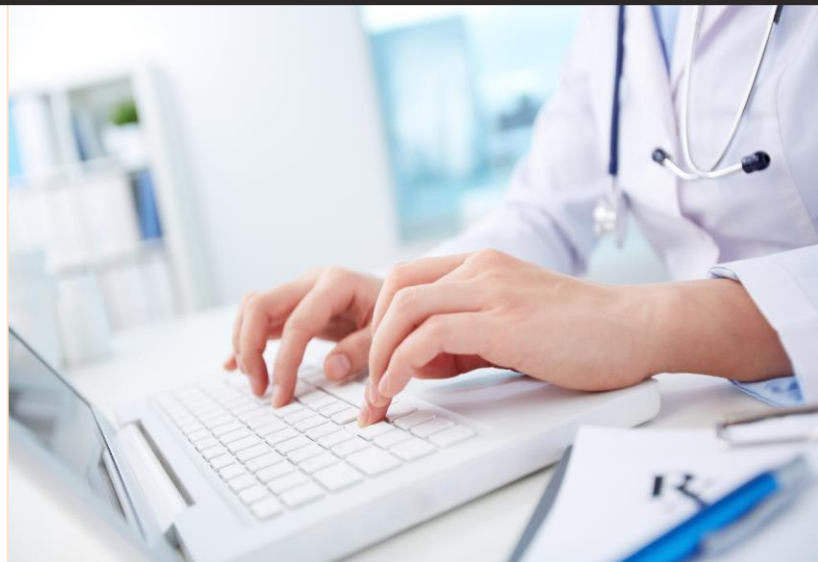
The Background

Our client is a trusted Healthcare provider based in the US. They wanted to develop a 360-view of personnel by leveraging AWS Data Lake and Snowflake Cloud Warehouse. Hence, their expectations included migrating existing on-premises Oracle and Hive environments to the AWS Data Lake and Snowflake Data Warehouse and implementing the data pipeline to extract data from different on-premises sources. The client also wanted to create multiple fact and dimension tables or one de-normalized table in the target.

Key Challenges

\ **Compromised data integrity:** Poor data quality led to unreliable outputs and potentially skewed analyses.

\ **Limitations of data ingestion:** Real-time change data capture and data ingestion from the source were unavailable, creating information processing lags.

\ **Overwhelming data load:** Data consumption was heavy, straining the system resources and slowing down business efficiency.

Our Solution

Tiger Analytics worked on developing a solution to help the client enhance data infrastructure, improve data quality, and establish robust data governance. The first step was physically separating each pipeline and associated data to offer clear demarcation, reduced interference, and better individual pipeline efficiency.

Next, the team implemented end-to-end orchestration using Airflow to streamline processes and automate workflows. Then, enterprise data governance was established using Collibra.

The team also conducted data quality checks and classification using Great Expectations and leveraged Oracle GoldenGate for real-time data replication.

Other solution highlights included:

- Glue data catalog
- Integration of dbtvault and transformation ETL in Snowflake
- Deploy dbt CI/CD and dbt pipeline automation using Terraform.

Tech Stack

/ **AWS**

/ **Snowflake Lakehouse**

/ **Terraform** (CI/CD)

/ **AWS Glue** (Batch Ingestion)

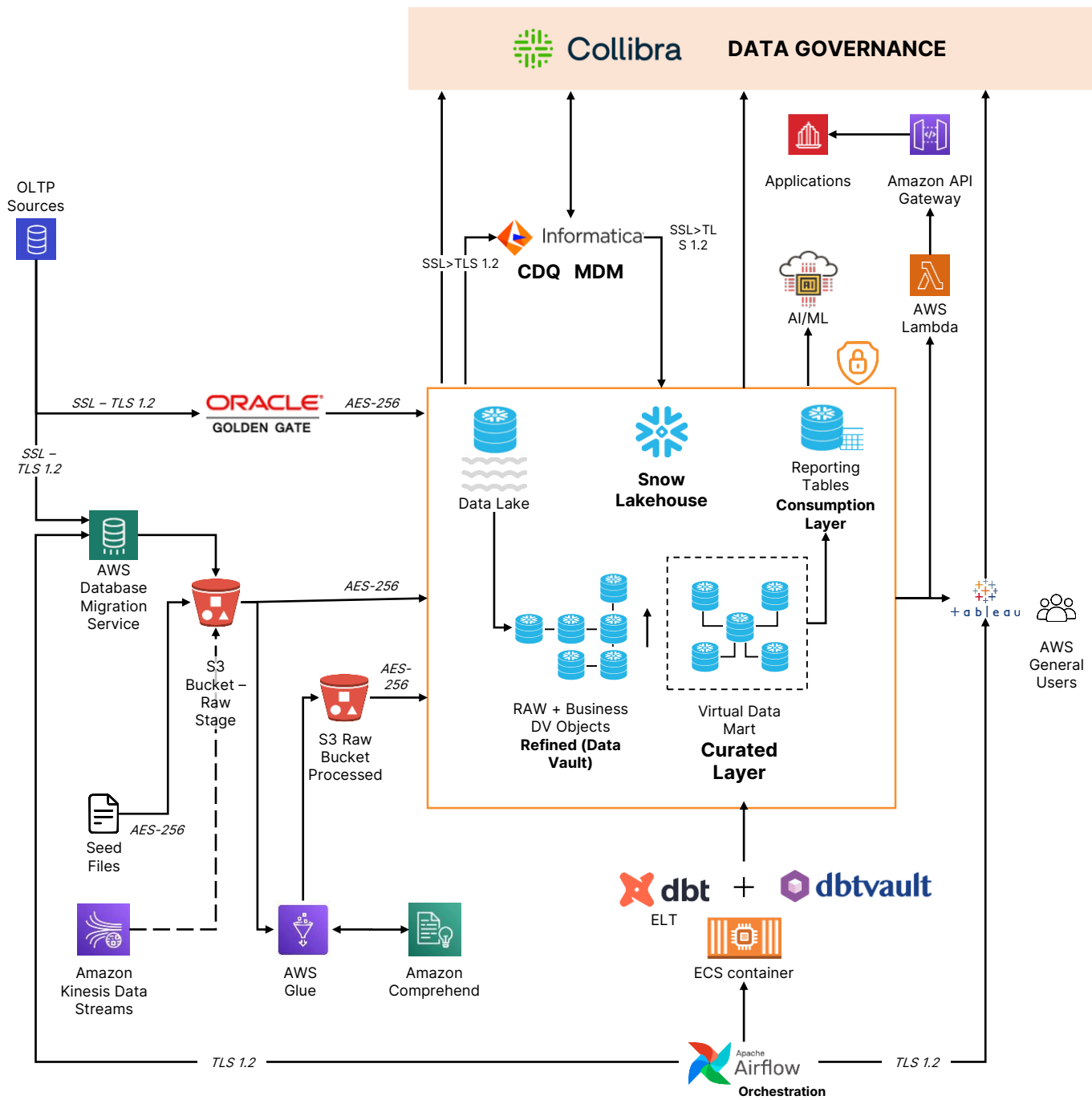/ **Oracle GoldenGate** (Real-Time Ingestion)

/ **dbt + dbt Vault**

/ **Airflow** (Orchestration)

/ **Great Expectations**

/ **Collibra Data Governance**

Solution Architecture

Case Study | Tiger Analytics©

Value Delivered

A unified single source of truth was created by extracting data from varied sources like ECHO and BOCA while also transitioning on-prem Oracle and Hive environments to the AWS Data lake and SnowFlake Data warehouse.

A holistic personnel view was built by integrating multiple datasets such as Personnel SF, Personnel Cerner, and Personnel Echo to create a 360-degree perspective of personnel data.

Data consumption was enhanced through Snowflake by establishing a curated layer containing transformed tables, integrating the dbt vault, transforming ETL logic like CERNER and SF, and integrating Tableau connectivity with Snowflake for seamless analytics.

About Tiger Analytics

Tiger Analytics is a global leader in AI and analytics, helping Fortune 1000 companies solve their toughest challenges. We offer full-stack AI and analytics services & solutions to help businesses achieve real outcomes and value at scale. We are on a mission to push the boundaries of what AI and analytics can do to help enterprises navigate uncertainty and move forward decisively. Our purpose is to ***provide certainty to shape a better tomorrow.***

Being a recipient of multiple industry awards and recognitions, we have 4000+ technologists and consultants, working from multiple cities in 5 continents.

www.tigeranalytics.com

US | UK | Canada | India | Singapore | Australia