

The Power of a Modern Data Warehouse

Highlighting scalability, performance and analytics value



The Client

- Godrej & Boyce Mfg. Co. Ltd. is the flagship company of the Godrej Group. The Company engages in construction, engineering, manufacturing consumer durables, aerospace, defense, logistics, technology, health care, tooling, process equipment, material handling, locking and storage solutions
- The solution delivers a modern, scalable, and reliable data infrastructure that ingests, processes, and stores data for reporting — while laying a strong foundation for future AI/ML workloads through reusable, well-governed data pipelines.

Overview

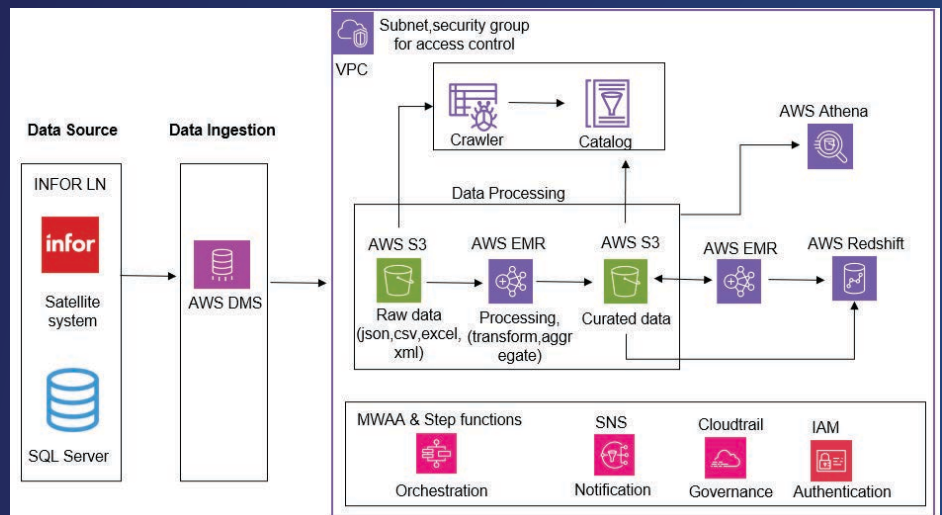
The client aimed to modernize their data infrastructure by leveraging a public cloud platform to improve agility, scalability, and operational efficiency. With over 40 disparate data sources across ERP and non-ERP systems, they required a robust, unified solution to centralize data ingestion, enable reusable pipelines, and support reliable reporting and advanced analytics — all within a secure, governed, and future-ready architecture.

Challenges

- Migration of 7 data cubes (~10TB, 5,000+ tables) to AWS S3 Data Lake
- Disconnected data sources spread across multiple systems
- High data volume requiring efficient migration (~10TB)
- Need for real-time data synchronization (CDC)

Solution

- Implemented ETL pipeline using AWS DMS to replicate data from MS SQL Server to Amazon S3 (raw layer)
- Data transformation and CDC handled via Apache Hudi scripts on Amazon EMR for scalable processing
- Curated data loaded into Amazon Redshift using Glue jobs to support KPI generation and analytics
- AWS QuickSight integrated with Redshift for seamless dashboarding and business insights



Benefits

- Increased the availability of customer application by moving to AWS
- 25% Reduction in overheads and cost