

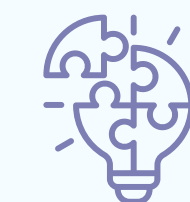
Modernizing Legacy Platform for a CMT Customer

Business Problem



- Increased CAPEX due to capacity bottlenecks on legacy on-prem Cloudera Hadoop slowing down queries and product launches
- Rising TCO from a fragmented tech stack and inefficient data movement
- Increased vendor pressure for license renewal, which led to shorter deadlines for migration
- Huge volume of data to migrate and validate: ~2 PB, ~27K tables, ~2.4K reports/queries

Solution



- Defined target architecture with EA: AWS + Snowflake Data Cloud, with AIOps/operational enablement
- Delivered migration via factory-led MigFac with reconciliation checks (file-size + partition counts)
- Repointed Ab Initio pipelines to AWS/Snowflake with minimal disruption
- Repointed BI dashboards/reports and stabilized batch + near real-time workloads
- Set up secure, reliable on-prem → AWS → Snowflake data movement (custom network pattern)

Value Delivered



- Modernized platform:** migrated from on-prem data platform to AWS-hosted Snowflake
- Performance impact:** reduced query time from ~10 hours to ~5 minutes
- Scale delivered:** successfully transitioned ~2PB of data
- Cost/risk optimization:** retired Cloudera licenses and associated infrastructure, reducing end-of-life cost exposure
- ETL migration:** migrated 3,500+ ETL scripts, including 700+ converted in under 7 weeks
- Delivery assurance:** seamless migration with minimal business disruption, delivered on time

50K TPS

Data Successfully Migrated to Snowflake on AWS

3500+

ETL Scripts Migrated, Including 700+ Converted in Under 7 Weeks

~10 Hrs → 5 Min

Query Performance Improvement After Platform Modernization