

# Enterprise-Scale AWS Cloud Migration

Cloud migration and re-architecting using a custom automation toolset



## Client

Fortune 50 energy company

## Background

Our client, a Fortune 50 energy company needed to migrate all their existing data and analytics platforms to Amazon Web Services and redesign the existing architecture to leverage AWS-native technologies. This was an initiative to optimize performance and reduce long-term operating costs by taking advantage of recent advancements with AWS cloud-native solutions.

## Challenges

- Large and complex data sets
- Parallel migration and re-architecting requirement
- Need for high migrated data quality

## Business Value

- Fast time-to-value
- Lower total cost of ownership using end-to-end AWS services
- Increased agility and scalability

## Technologies

- VmWare Greenplum
- Talend ETL
- HVR Software
- Tableau

## Challenge

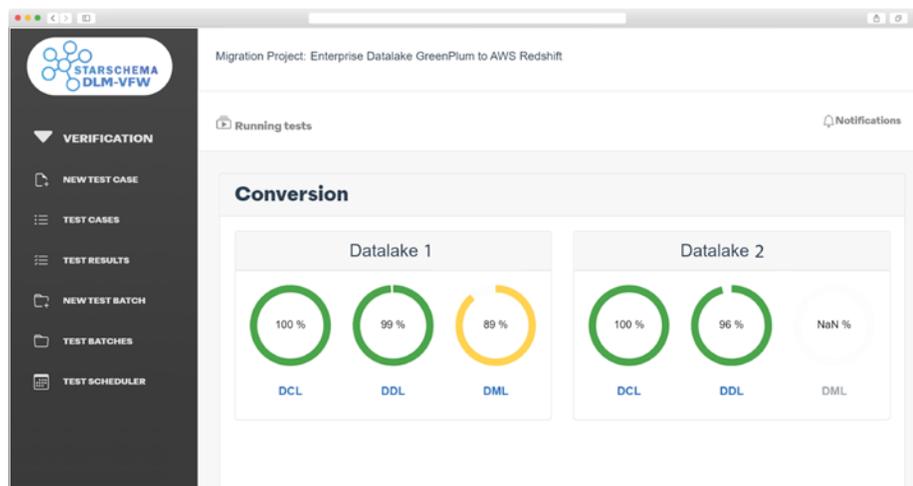
The project involved lift-and-shifting over three thousand Talend ETL jobs, thousands of stored procedures and hundreds of Tableau reports from on-premises databases while altering relevant architecture to take advantage of AWS-native technologies.

The migration had to ensure fast time-to-value, which necessitated retaining the quality of dashboards during the migration process to minimize the maintenance required after go-live.

## Solution

Starschema conducted a preliminary assessment of critical factors such as workloads, peak times and potential issues. This reduced the associated burdens on the client during the migration, and we used the results to optimize the solution to ensure effectiveness.

The Starschema team's senior AWS architects designed the new architecture and led a "bottom-up" team focusing on the platforms, data integrity and ETL/UDF conversion and a "top-down" team handling report conversion and outlying data quality issues.



Starschema's proprietary Data Lake Migration Verification Framework (DLM-VFW) enabled the client to monitor and test the results of the ongoing migration in a browser-based user interface.

Through analysis of the client's on-premises data lake we created a project plan that enabled continuous report delivery for each business unit. Our team developed a web-based application that allowed the client to monitor and test migration results, including object availability, data integration latency, data type mismatches, null columns and row counts. We were also able to rely on our close working relationship with AWS to recommend new features and introduce them on-the-fly during the migration.

The solution's central component was a custom-developed toolset for automating and optimizing tedious and repeatable tasks to increase productivity when re-architecting the client's on-premises data lake and its components in the AWS cloud, using AWS-native services where possible.

Using this toolset, we lift-and-shifted the client's large and complex data sets from Oracle and Greenplum to AWS while simultaneously re-architecting them to fully leverage cloud-native technologies such as S3, Spectrum, Glue and Redshift, greatly reducing the manual labor involved.

Our team paid special attention to making the maintenance and future optimization of the new architecture easier. The metadata inventory and version control system we employed for productivity during the migration would later serve this secondary purpose, and we used recorded video conferencing and extensive documentation to transfer knowledge to the client.

After the go-live, Starschema provided hypercare for 30 days, using our state-of-the-art in-house operations monitoring systems to enable predictive maintenance, reduce the reliance on manual labor and future-proof the client's new, cloud-based environment.

## Outcome

The client realized excellent time-to-value by using automation, saving hundreds of man hours during the migration process, and was able to rely on the new architecture during the end-of-year closing shortly after go-live. This was also possible because the migrated dashboards required very little maintenance thanks to the high data quality retention during the automated processes.

The client reduced the total cost of ownership by using end-to-end AWS services to achieve savings on hardware, license and maintenance costs, and the eventual fine-tuning of the architecture will help realize further cost optimization.

Our custom automation toolset includes tooling and frameworks for UDF conversion, DDL conversion and data type optimization and allows for the automated migration of databases, ETL pipelines, ingestion and visualization platforms. It supports Greenplum, Redshift, Snowflake, Tableau, Talend and HVR alongside other technologies and enables exceptional data quality to be retained during the migration.

In total, we migrated over 3800 Talend ETL jobs, 3150 PLSQL stored procedures and 630 workbooks from 210 published data sources.

## Go Further

Dig deeper into the proprietary framework and expertise that made this migration possible and explore Starschema's AWS Cloud Migration solution.

Start your journey:  
[starsche.ma/aws](https://starsche.ma/aws)

## ASK THE EXPERT

---



### **Anjan Banerjee** Lead Architect at Starschema

Anjan Banerjee, lead architect at Starschema, has been working with cloud technologies and data warehouses for ten years. He has extensive experience in building data orchestration pipelines, designing multiple cloud native solutions and solving business critical problems for many multinational companies. He believes in the concept of infrastructure as code as a means to increase the speed, consistency, and accuracy of cloud deployments. Anjan holds frequent trainings on data munging and various self-service analytics and transformation tools like Talend and Alteryx.

**Contact Anjan:** [anjanb@starschema.com](mailto:anjanb@starschema.com)

## COMPANY

---

At Starschema we believe that data has the power to change the world and data-driven organizations are leading the way. We help organizations use data to make better business decisions, build smarter products, and deliver more value for their customers, employees and investors. We dig into our customers toughest business problems, design solutions and build the technology needed to compete and profit in a data-driven world.

## GET IN TOUCH

■ [info@starschema.com](mailto:info@starschema.com)  
● [starschema.com](https://starschema.com)