WORKLOAD TRANSFORMATION SOLUTION

Teradata Cost Optimization by transforming Batch and Ad hoc Workloads to Hadoop

A Multinational Retail Chain Saves Millions

Client Overview

A U.S.-based multinational retail chain wanted to optimize the current Teradata data warehouse and ensure that its analytical applications are using the most appropriate data platforms. To achieve this, the client wanted to transform batch applications and ad-hoc queries that consume expensive Teradata cycles to the Hadoop/ Hive environment.

Requirements

The client wanted to identify and transform the batch and ad-hoc workload from Teradata to Hadoop.

- Batch transformation includes ingestion of dependent datasets from the source data store, building batch jobs, upload transformed data to target systems, perform data validation, warranty support and production hand-off.

- Ad-hoc transformation includes establishing access patterns, ingestion of dependent datasets from the source data store, query conversion including conversion of Teradata specific operators to the target systems and user-training assistance.

BUSINESS CHALLENGE

Optimize and reduce the current cost of the legacy data warehouse and build a data lake for faster ETL processing and use of advanced analytics while reducing time to analytics.

BENEFITS

- Cost savings
- Transformation to modern technology
- Improved user experience
- Business risk mitigation

Workload Transformation Solution

The Impetus solution involved two key transformations:

One time migration

- Historical data (from Teradata or DB2)
- Scripts and queries used in batch and ad-hoc fashion to access Teradata
- Movement of analytical post-processed data from Hadoop to Teradata

Incremental migration

- Daily, weekly, and monthly increments of data from the current source to Hadoop
- Movement of analytical post-processed incremental data results from Hadoop to Teradata

The Impetus Workload Transformation Solution was implemented for historical data Transformation to identify and offload data and workloads from the enterprise data warehouse to Hadoop.

The solution features an automated utility that converts BTEQ and SQL transformation scripts into equivalent Spark QL/ HiveQL and executes them on the Hadoop/ Hive environment. It also allows users to run a set of data validation checks. Finally, the post-processed analytical data can be loaded back to the source enterprise data warehouse for reporting and access.

Further, a combination of the Impetus Workload transformation product and IBM DataStage were used for incremental data transformation from the source data store to Hadoop. Processed results can also be updated back to Teradata or other source systems.

Solution Architecture
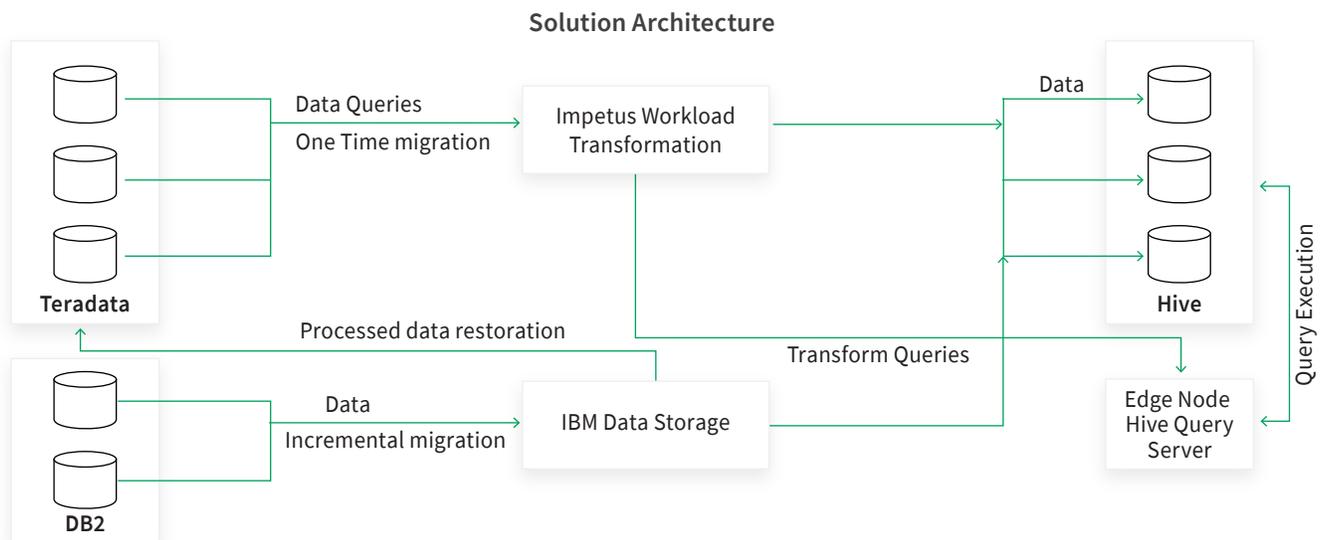
Client's Data Fabric Balancing Solution Architecture

Save more than 80% of your next transformation project cost.
Contact inquiry@impetus.com for details.

IMPETUS

Impetus is focused on creating big business impact through Big Data Solutions for Fortune 1000 enterprises across multiple verticals. The company brings together a unique mix of software products, consulting services, Data Science capabilities, and technology expertise. It offers full life-cycle services for Big Data implementations and real-time streaming analytics, including technology strategy, solution architecture, proof of concept, production implementation and on-going support to its clients. To learn more, visit **www.impetus.com** or write to us at **inquiry@impetus.com**.