

# A large US bank accelerates the implementation of a hybrid cloud-based data lake architecture

---

Automated conversion of Teradata BTEQ scripts in record time

**CASE STUDY**

## Client Overview

A U.S.-based Fortune 500 bank wanted to offload their Teradata-based Marketing Campaign Process (MCP) to Hadoop. They wanted to execute the end-to-end package on CDH environment and produce an output that matched the Teradata output. To achieve this, the client wanted to transform BTEQ queries that consumed expensive Teradata cycles to the Hadoop/big data environment.

## Business Challenge

As one of the largest banks in the U.S., the client had vast amounts of historical data that were used for validating transactions in real-time. However, while analyzing the data, they faced multiple business and technical challenges:



Environment issues that blocked the migration multiple times



High validation time for baseline data issues, data quality, and other Teradata query gaps



Complex BTEQ scripts and procedures

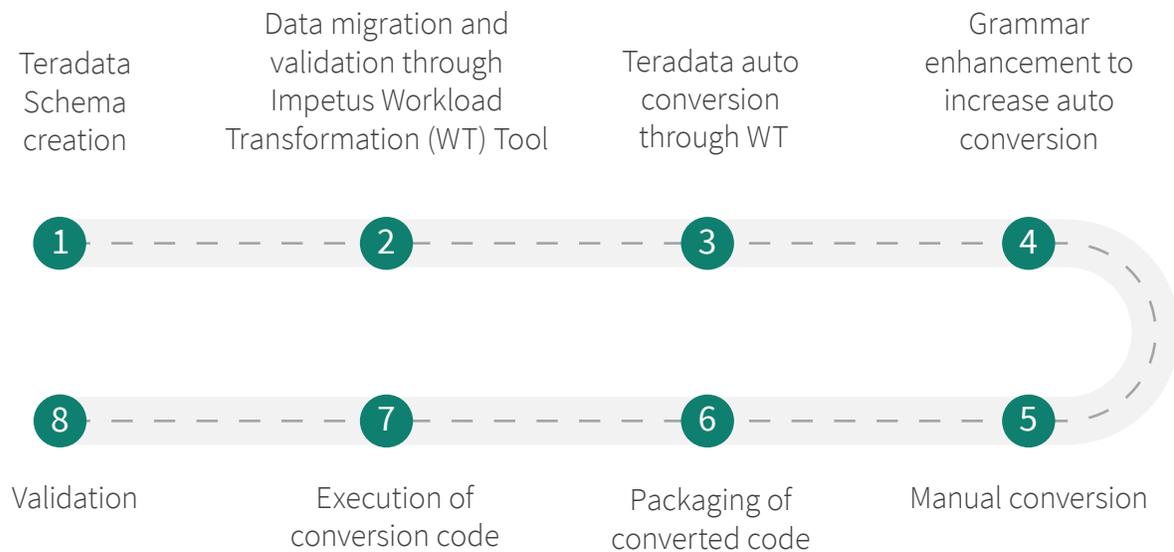


High cost of ownership

# Approach

For an end-to-end transformation of workloads from Teradata to a modern Hadoop-based data lake architecture, the client needed to do the following:

- Define solution architecture and design for campaign management application on CDH
- Migrate schema and data from Teradata to CDH
- Convert code and business logic to equivalent HiveQL/Spark SQL automatically so that risks can be avoided and time to market can be minimized
- Convert code to target-compatible logic achieving 100% transformation objectives



## Solution

Impetus designed a hybrid big data-based cloud solution to identify and offload data and workloads from the enterprise data warehouse to Hadoop. The solution features an automated utility that converts Teradata SQL queries into equivalent Spark QL/Hive QL and executed them on the Hadoop/Hive environment. It also converted the business code into equivalent Python code and allowed users to run a set of data validation checks. The transformed queries were validated both syntactically and semantically to ensure that the transformed codes are 100% precise. Finally, the post-processed analytical data was loaded back to the source enterprise data warehouse for reporting, access, analytics, and providing recommendations.

The key features of the Impetus Teradata Workload Transformation Solution are as follows:

- CDH environment compatible converted code to produce similar data
- Output as original queries on the Teradata environment
- Consumption-ready data served from Hadoop replacing Teradata

## Benefits

- More than 50% faster than the manual approach
- 'As is' application architecture and coding standards in the converted code
- Reusable UDFs to enable faster future conversions and enabling
- Provided performance recommendations

## Value Delivered

Our comprehensive solution aligned best practices with our client's enterprise data lake goals.

- Auto transformed 1203 out of 1374 BTEQ queries with 87.55% success
- Converted Hive schema with 104 tables and 300 GB of compressed data
- Generated campaign data and reports on Hive to match Teradata output
- Packaged and executed the core process for CDH production deployment
- Recommended performance tuning and cluster sizing for the production environment
- Performed semantic validation for all the transformed queries based on an indigenous query validation engine. It generated diverse sample datasets and compared data on source and target catering to different use cases.

To learn more about our Teradata transformation solutions contact us at **migration@impetus.com**

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](http://www.impetus.com).