

IMPETUS

# Automated Netezza Transformation to Big Data Open Source

CASE STUDY

## Client Overview

Our client is one of the largest cable companies in the world\*, offering a wide range of services including basic cable, digital cable, DVR, digital phone, and Internet services.

## Business Challenge

Their data warehouse, built on a Netezza platform, was expensive and offered limited scalability. Massive volumes of data were choking their data storage and processing capacity resulting in the following challenges:

*\*The client depicted in this case study is an actual client of Impetus and reserves the right to remain anonymous.*



High storage & processing costs



Constrained EDW capacity



Limited analytical access to data



Inability to process semi/unstructured data

# Approach

To transform workloads from the Netezza platform to Hadoop we followed a three-step approach:

1

## Netezza assessment

Conducted a comprehensive assessment of current Netezza workloads using our cognitive workload assessment engine. We defined a transformation strategy and a complete roadmap to our value proposition.

2

## Created a pilot using data & logic

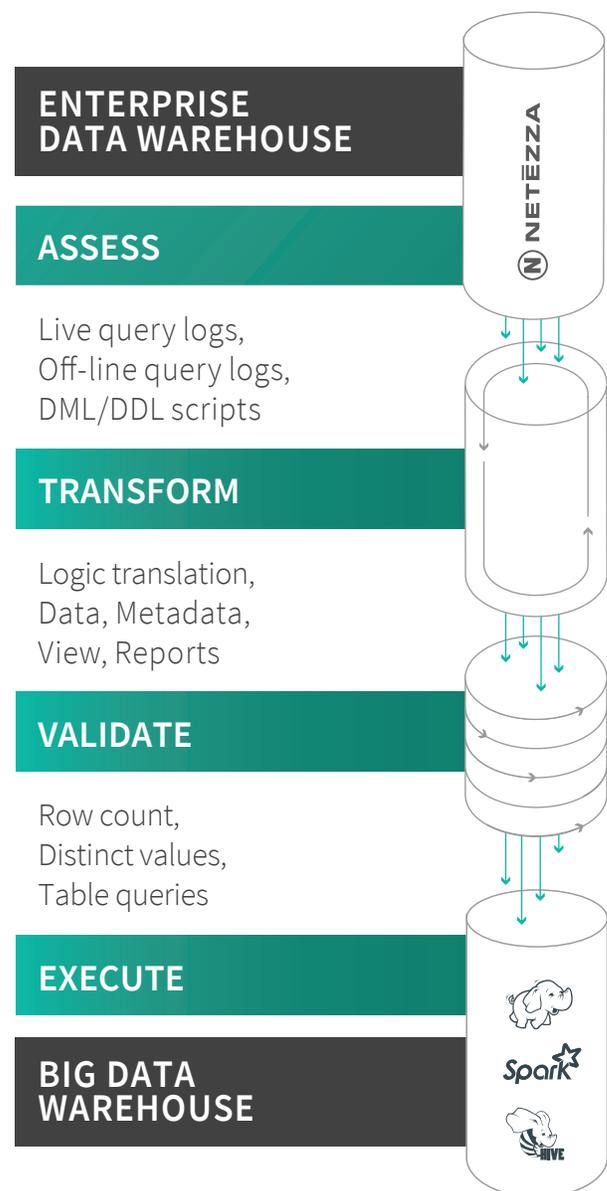
This phase enabled the client to quickly and confidently confirm the expected ROI. With minimal human intervention, our automated platform demonstrated just how simple, secure, and rapid our Workload Transformation can be.

3

## Executed our migration process

Using our proprietary Workload Transformation toolset, we transformed all the remaining workloads.

As illustrated below, our automated, end-to-end transformation process includes assessment, transformation, validation, and execution:



---

## Solution

We designed and created a Hadoop-based storage and retrieval solution that completely offloads the Netezza workloads and achieves the following:

- Creates a Hadoop-based database from Netezza DDL scripts
- Replaces the Netezza command line shell scripts to Hive/Spark beeline scripts
- Identifies and translates the underlying procedural and SQL queries in the Netezza shell scripts into Spark SQL/Hive QL
- Replicates the Netezza-based ETL process on Hadoop
- Efficiently processes a 1TB database with 50 GB of incremental data coming in daily
- Validates the transformed data, metadata, and workloads
- Migrates historical data (one-time) and performs incremental updates

# Value Delivered

Our solution accomplished three things:

**A Scalable & Elastic Architecture**

**Significant Cost Reduction**

**A Proven, Automated, Risk-free approach**

- Transformed the legacy Netezza workloads to an hadoop-based architecture
- Precisely positioned the Netezza workloads in the Hadoop space
- Automagically translated Netezza logic using Impetus automated translation engine technology
- Realized significant time, effort, and conversion cost reductions
- Eliminated Netezza infrastructure costs
- Provided on-site services for manual conversion of any remaining Netezza logic required to achieve 100% execution in the Hadoop environment

To learn more about our Netezza transformation solutions contact us at [\*\*inquiry@impetus.com\*\*](mailto:inquiry@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).