



IMPETUS

Automated Netezza to Cloud Transformation

CASE STUDY

Client Overview

Our client is a government-sponsored enterprise* that provides financial products and services to increase the availability and affordability of housing for all income groups in the U.S.

Business Challenge

As a leading source of financing for mortgage lenders they were expanding into secondary markets and would require a modern data lake architecture. However, they were faced with numerous business and technical challenges:

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



High cost of ownership



Consumption SLA were not met due to constrained EDW capacity



Unbalanced positioning of workloads on Netezza



Inability to process semi/unstructured data

Approach

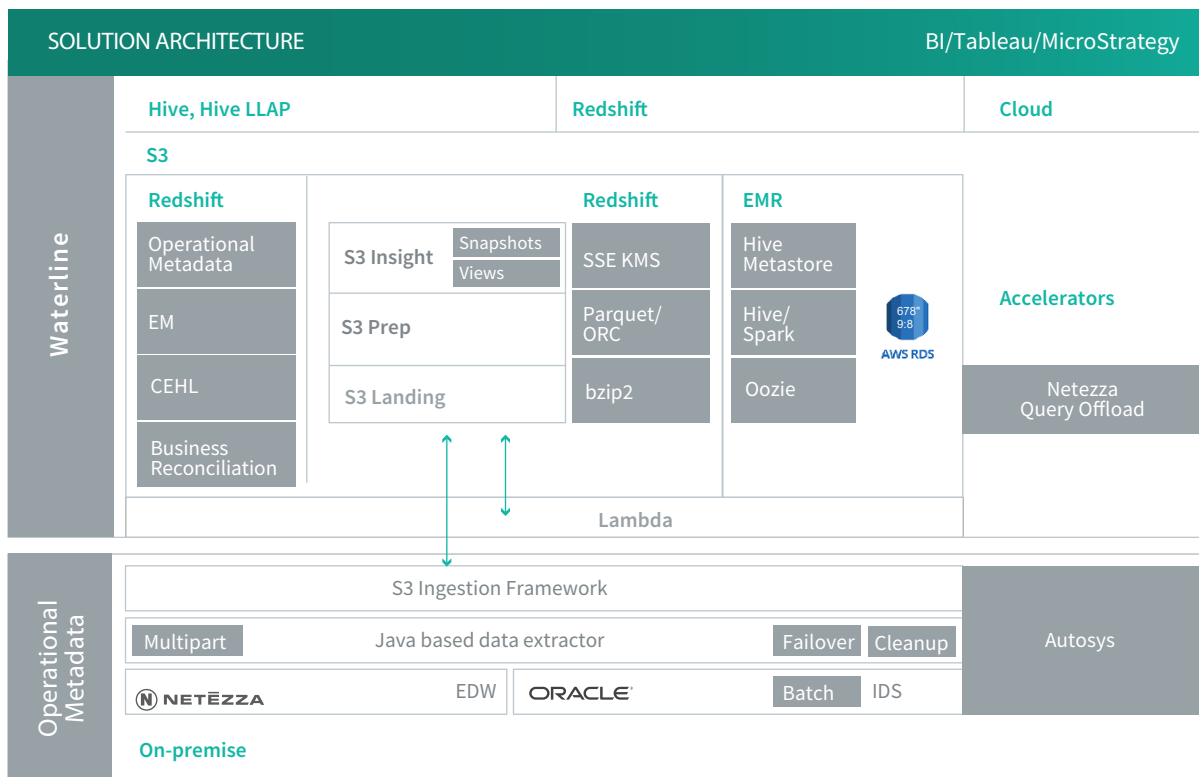
For an end-to-end transformation of workloads from Netezza to a modern Hadoop-based data lake architecture, the client needed to do the following:

- Assess and identify expensive/poor performing workloads and users
- Transform historical data (one-time) and perform incremental updates
- Automatically translate Netezza queries into Spark SQL/Hive QL
- Balance the workloads effectively in the Hadoop space
- Recommend optimal storage formats for the data in the cloud
- Perform SCD Type 1 and Type 2 for slowly changing dimensions
- Refresh Netezza snapshots on AWS on a daily basis
- Recommend best practices for metadata, data lineage, and audits
- Perform a comprehensive business reconciliation for every iteration
- Reload data back to the EDW for reporting/analytical consumption
- Moved reporting/analytical consumption loads to scalable AWS vending layer

Solution

Impetus designed a hybrid big data based cloud solution while accomplishing the following:

- Migrated historical data and developed a reliable mechanism for incremental updates
- Automatically translated consumption queries (i.e., SQL, Stored Procedures, Shell, Scheduler) into Spark SQL/Hive QL
- Ingested data from S3 to Redshift for reporting/analytics
- Replicated Netezza snapshotting process onto AWS
- Used Autosys and Oozie as on-premise/cloud orchestrators
- Assessed the existing workloads and provided recommendations to position them in Hadoop
- Used AWS Lambda as an event-driven notification engine
- Used Amazon EMR to process vast amounts of data on cloud across scalable Amazon EC2 instances
- Used Amazon RDS to store metadata and Workflow info
- Evaluated and performed tool benchmarking for consumption queries in the cloud



Value Delivered

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

- Helped the client to sunset Netezza
- Contained expenditures on aging technologies by freeing up capacity in Netezza
- Used Impetus accelerators to automagically translate Netezza logic
- Realized significant time and financial savings by reducing manual conversion efforts
- Provided on-site support and services to ensure 100% success
- Drove the innovation agenda by improving data availability across the enterprise

To learn more about our Netezza transformation solutions contact us at 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.