Case Study

RETAIL DEPARTMENT STORE

Infoworks

Company Background

This American retail department store chain was founded in the late 1800s and currently boasts almost 600 retail locations

The Challenge

Like many other brick and mortar retailers, this department store chain has been facing an existential threat from the emergence of online shopping retailers like Amazon. As a result, many years ago, they invested in online initiatives and their online business now accounts for over $1B in revenue.

In 2018, they announced 5 strategic initiatives, the first of which was to expand and improve the value of their customer loyalty program. According to Fortune.com in late 2017, "(This retailer) gets about 50% of its annual sales from the top 10% of its customers, those who spend at least $1,200 a year at the chain. It's about ten times more expensive to win a new customer than to entice a current one to stay, according to experts."

While delivering an improved loyalty program requires the creation of an enticing program offer, it also requires the creation of back end analytics systems that support customer loyalty data modeling and analysis. This means collecting and analyzing decades of customer purchase history information, near real time customer loyalty data from stores, inventory transaction analytics from stores and click stream data analysis from their online store.

They built their original big data loyalty analytics solution on an on-premises foundation based on the Hortonworks Hadoop implementation. While Hortonworks provided a solid foundation, it still required a significant amount of hand coding for each new analytics use case. Over time, they also began to invest in moving components of their solution and their data into the cloud. This resulted in a complex deployment that consisted of loyalty data residing in Microsoft Azure, inventory, loyalty and transaction data coming from on premise systems and the final analytics being performed on Google Big Query.

CASE STUDY: RETAIL DEPARTMENT STORE

Challenges

- Internal hand-coded automation framework was too brittle
- Strong data engineering team, but could not grow the team fast enough to match the pace of requests from the business
- Multi-cloud and hybrid deployment environment

Results

- Developed initial data pipelines and workflows in just 20 business days

Business Benefits

- Data engineering team members are now more available to develop data pipelines for new analytics use cases
- Able to build data pipelines once and run them in multiple environments without recoding

info@infoworks.io

(650) 391-9306

Infoworks

The Challenge (continued)

In order to enable faster time to deployment for continuously evolving analytics demands from the business, the retailer invested in building their own automation layer which was built to deal with challenges around change data capture, data ingestion and data transformation as well as moving the data efficiently across their hybrid data environment.

Over time however, they found that "we spent more time troubleshooting our hand-coded framework than we did actually using it." While the automation framework was more efficient for building repeatable use cases that resembled the existing use cases, it was rigid and brittle and would break when new use cases or datasets were introduced that didn't closely resemble the previous use cases.

> "Infoworks made our data engineers significantly more productive almost immediately after we started using it.
>
> *Lead Enterprise Architect*

The Solution

The retailer invested in Infoworks DataFoundry, an Enterprise Data Operations and Orchestration (EDO2) system, to replace their existing hand-coded framework and give them a single development and deployment platform that worked across all of their environments without requiring them to recode. They used DataFoundry to first perform high speed replication between their Azure environment, which held 70TB of historical transaction data, and their on-premise system. They used it to ingest loyalty data from Oracle, with implementation of change data capture (CDC) for managing constant updates as new data arrived, as well as handle inventory data streaming into the system via Kafka. Ultimately, data is moved to the Google Cloud Platform where loyalty program analytics views are being created on top of Google BigQuery.

The Results

The new implementation based on Infoworks DataFoundry provides significantly greater agility while accelerating the creation of manageable end-to-end data pipelines. This is critical as they continue the digital transformation of their business and refine the analytics around their loyalty program, which is expected to continue to evolve as one of their strategic initiatives going forward.

Additionally, the Infoworks-based data pipelines built to run on-premise on Hortonworks can also run on the Google Cloud Platform without recoding. This provides the retailer a much greater level of deployment flexibility and the ability to take advantage of new cloud-based capabilities as cloud vendors continue to evolve their environments.

About Infoworks

Infoworks provides the first Enterprise Data Operations and Orchestration (EDO2) software system to automate the development and operationalization of data pipelines from source to consumption in support of business intelligence (BI), machine learning (ML) and artificial intelligence (AI) applications. Infoworks' code-free development environment allows organizations to develop and manage end-to-end data workflows, or migrate existing data and workflows, without requiring an army of big data experts. Infoworks delivers capabilities to automate and simplify development of data ingestion, data preparation, query acceleration and ongoing operationalization of production data pipelines at scale. Infoworks supports cloud, multi-cloud, and on premise deployments, enabling customers to deploy projects to production within days, dramatically increasing business agility and accelerating time to value.