

CASE STUDY

Dremio Drives Increased Revenue by Improving Supply Chain Analytics for Global Consumer Products Company



CUSTOMER

Global Consumer Products Company



OBJECTIVES

Improve supply chain analytics—increasing the efficiency of data access and accelerating time to insights about supply chain and consumer demand by transitioning from a traditional RDBMS architecture to a modern data lake analytics platform.

CHALLENGES

- Data analysis resulted in conflicting reports that impaired business decision making
- Duplication of data and business logic caused inefficiencies
- Cube implementations were expensive, time-consuming and delivered minimal value

SOLUTION/DATA ENVIRONMENTS

The platform consists of several services:

- Compute: Dremio, Azure Databricks
- Ingestion: Azure Data Factory, Apache NiFi
- Storage: ADLS, HDInsight, SQL Server, Oracle
- Clients: Power BI, MicroStrategy, Qlik (among others)

RESULTS

- Accelerated adoption of the company's modern cloud data platform
- Sub-second response times dramatically accelerate time to results
- Significantly reduced data duplication and increased efficiency
- Increased focus on analytics rather than infrastructure
- Easier data access helps projects move forward faster
- Reduced development cycles minimize the number of data engineers required to surface data to users

Summary

A global consumer products company is using Dremio to improve supply chain analytics—increasing the efficiency of data access, analytics and insights about supply chain and consumer demand. They are transitioning from a traditional RDBMS-only architecture to a more modern and flexible data lake analytics platform that leverages Dremio, Azure Databricks, ADLS and various relational databases.

The Business

This multinational consumer products company produces more than 240 million products every year at more than 150 manufacturing sites in 30 countries. The company sells its 200 brands in more than 180 countries.

The Challenge

It was challenging for the company to obtain accurate insights into its supply chain and consumer demand across global operations. Data analysis was decentralized, siloed and inefficient, resulting in duplication, wasted resources and conflicting reports.

The company operates in over 30 markets where products are sold through different types of channels. As a result, it receives a large volume and variety of datasets which the business relies on for its KPIs. There was no way to centralize and facilitate the reuse of these KPIs and the business logic needed to generate them. Hundreds of analysts across the company were using different BI tools and repeating the same calculations and recreating the same business logic and KPIs in their various BI tools.

To remedy this, the company launched a strategic initiative to help the business grow through better analytics and insights into their supply chain and consumer demand. Their goal was to integrate analytics across the company and its many business units and brands by centralizing data access and analysis using a data lake.

The company started its journey by creating a data lake using Microsoft Azure coupled with HDInsight and Power BI. However, the company's business users did not feel comfortable working with the new architecture

as there was no easy way for users to access data on the data lake with the business tools they were comfortable using. For example, brand managers could not easily get the data for the brands they were interested in and would have to call IT to extract datasets, which could take up to 12 weeks to complete because of the backlog and the complex ETL process involved. Many BI users and data analysts couldn't take advantage of the data lake, so they exported data from the lake into SQL Server data marts. The result was delayed time to value, low platform adoption and limited insights.

The Solution

The company deployed a hybrid cloud spanning Azure and the company's network. The primary components running on Azure include Azure Data Lake Storage (ADLS), Azure Data Factory, Azure Databricks, HDInsight and Dremio. Many of the data ingestion processes store files in ADLS as Parquet and CSV files. Dremio serves as a query acceleration and semantic layer platform, providing fast, centralized semantic access to the data in the cloud as well as to a variety of relational databases the company has yet to migrate. Dremio's dynamic security controls allow only the right data to be exposed to the right people.

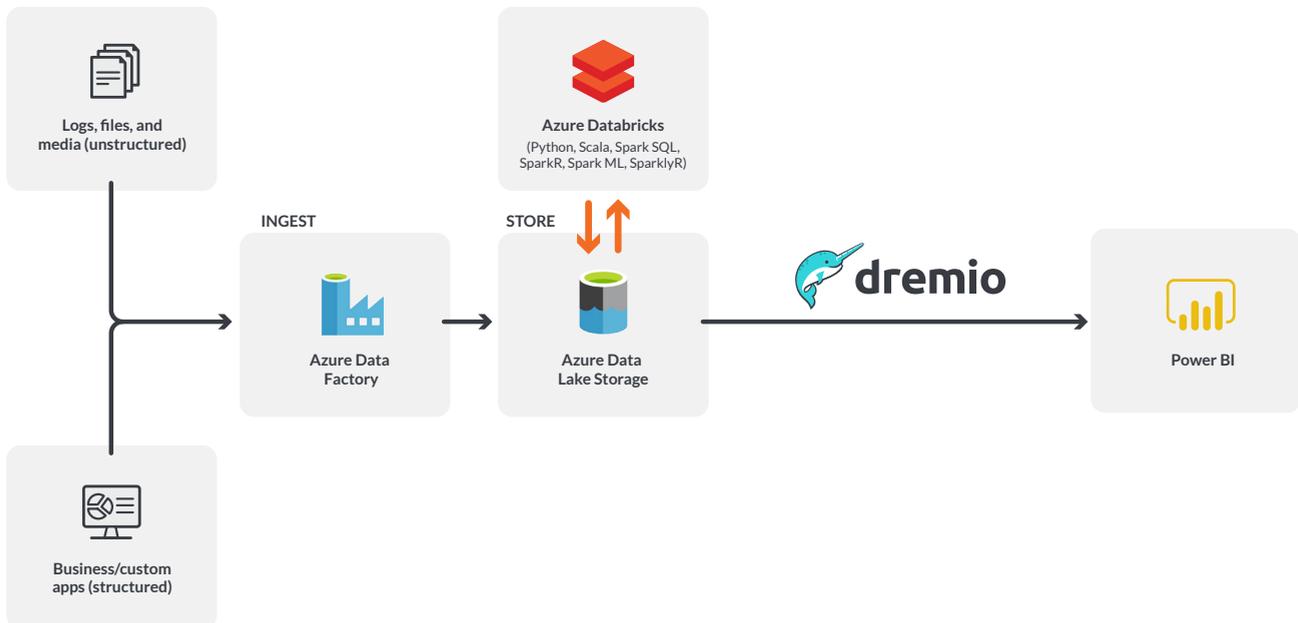
The company uses two different ETL frameworks. For their global data products that have common data models across multiple markets, they use Kylo NiFi to ingest data into the ADLS storage. They add the files to ADLS, where Dremio reads them and enables analysts to access them for live, interactive analysis. For their more regional and point solutions, they use Azure Data Factory with Databricks.

Results

WIDER VISIBILITY INTO THE BUSINESS ENABLES BETTER DECISION MAKING

Prior to moving to the data lake with Dremio, the company's analytics capabilities were extremely limited—they could only monitor which items left the manufacturing plant. With Dremio, they can store and easily analyze huge volumes of data from sensors and third parties who provide information such as weather patterns and traffic information. The company's supply chain workflows are now significantly more efficient, and analytics can be done at the point of sale to allow for more targeted revenue-driven projects.

Azure Dremio Architecture



ACCELERATED ADOPTION OF MODERN DATA PLATFORM

Many employees who work with traditional RDBMSs like SQL Server or Oracle initially had reservations about migrating to the new platform. Dremio not only made the platform accessible but also allowed users to continue to leverage data in their RDBMS by joining them with datasets in the data lake storage. Dremio served as a change accelerator to get people on board and bought into the vision of a modern cloud-based data platform.

ELIMINATED DUPLICATION AND INCREASED EFFICIENCY, REDUCING COSTS

Dremio enables the company to eliminate duplication of data and business logic as well as the associated cost and complexity. Other solutions that it considered required creating a copy of data from ADLS to put into another persistence layer, adding significant overhead. With Dremio, the company can avoid duplicating data from ADLS while centralizing business logic and KPIs in Dremio's semantic layer, allowing them to keep their BI layer thin.

Prior to Dremio, the business analysts and developers in each market would choose their own BI solution and create KPIs and caching in that BI tool. Since 90 percent of the KPIs are identical across markets, this resulted in conflicting reports. It also added additional expense, as the company had to hire specialized developers for each of the different technologies (BI tools and OLAP cubes) that were being used.

The company is now leveraging Dremio to centralize both KPI calculations and caching, which no longer need to be repeated inside dozens of different BI deployments and regional markets. They have also been able to significantly reduce the use of OLAP cubes, some of which took more than 24 hours to refresh, increasing the versatility and agility of some key markets.

ACCESS TO DATA ALLOWS PROJECTS TO MOVE FORWARD FASTER

Dremio has been a huge enabler that allows many teams to access data they wouldn't have been able to otherwise, helping teams accelerate progress on their projects. The central data engineering team used to get a lot of requests for one-off data sets that were time-consuming to fulfill. The company can now make data available from the data lake and get analysts quickly started on projects that would have previously taken months.

CLOUD DATA LAKE ALLOWS THE COMPANY TO FOCUS ON ANALYSIS, NOT INFRASTRUCTURE

The combination of Dremio and ADLS, along with Dremio's ability to join data in data lake storage and on-premise data sources at interactive speed, powers the company's cloud data lake. This approach allows them to focus on driving business value instead of the management headaches associated with managing the infrastructure.



ABOUT DREMIO

Dremio is a fundamentally new approach to data.

We started Dremio to shatter a 30-year-old paradigm that holds virtually every company back. Removing barriers, accelerating time to insight, putting control in the hands of the user. That's Dremio.

Deploy Dremio →

[Learn more at dremio.com](https://dremio.com)

CONTACT SALES

contact@dremio.com

All third party brands, product names, logos or trademarks referenced are the property of and are used to identify the products or services of their respective owners. © Copyright Dremio 2020. All Rights Reserved.