



# FactSet Modernizes Applications with Dremio, Accelerating Data Access and Eliminating Complexity

Challenges	Solution	Results
Need faster access to constant flow of financial data, better way to cache incoming data flows	Dremio Reflections layered on top of existing technologies such as AWS data lake, lays foundation for FactSet's modern data architecture.	Reduced time to access data from 10 minutes to under 30 seconds. Agility without ripping and replacing existing infrastructure.
Need to deliver instant access to financial data	Dremio improves client experiences by speeding data access across multiple sources.	Tens of thousands of investment professionals are equipped with financial data and analytics to make more informed financial decisions.
Need to manage multiple silos of complex data sets and applications, and track data sources	Dremio enables FactSet to create a single pane of glass for data access, with a graph lineage of data sources.	Simplify development, reduce complexity of data engineering pipelines, and modernize applications, allowing engineers to focus on more strategic work.

## CUSTOMER

# FACTSET

<https://www.factset.com>

## GEO

Global

## INDUSTRY

Financial Services

## OBJECTIVES

Modernize and simplify applications to provide financial data from multiple sources to clients as a single pane of glass.

## DATA ENVIRONMENT

- **Data Sources:** Data lake on AWS with Amazon S3, SQL Server, Elasticsearch
- **Storage:** Amazon S3
- **Compute:** Dremio
- **Applications:** BI: Custom web application

## Summary

Financial data and software company FactSet is using Dremio to modernize its applications, accelerating access to crucial financial data by 20x to help clients make better investment decisions.

## The Business

FactSet Research Systems Inc. provides flexible, open data and software solutions to over 170,000 investment professionals around the world. FactSet solutions provide instant access to financial data and analytics that investors use to make crucial decisions. FactSet helps clients stay ahead of global market trends, access extensive company and industry intelligence, monitor portfolio risk and performance, and execute trades.

## The Challenge

In today's increasingly competitive financial environment, investors need instant access to financial data to make decisions. FactSet is an industry leader in acquiring, integrating, and managing financial data, offering clients access to 30+ datasets, and 850 independent data providers. FactSet customers can monitor global markets, research public and private companies, and gain industry-level insight with comprehensive reports that include financials, estimates, debt, ownership information, and more. FactSet's comprehensive data feeds, desktop analytics, web and mobile applications, and services help clients discover, decide, and act on opportunities.

With the explosive growth of financial data from multiple sources stored in different locations, FactSet wanted to provide a single view with fast data access for its clients. FactSet aimed to modernize and simplify the development of its applications to enable new workloads across multiple complex content sets. One of their challenges was caching the constant flow of data from multiple sources. They wanted to replace an internal caching system with a solution that would not require mobilizing too many internal engineering resources.

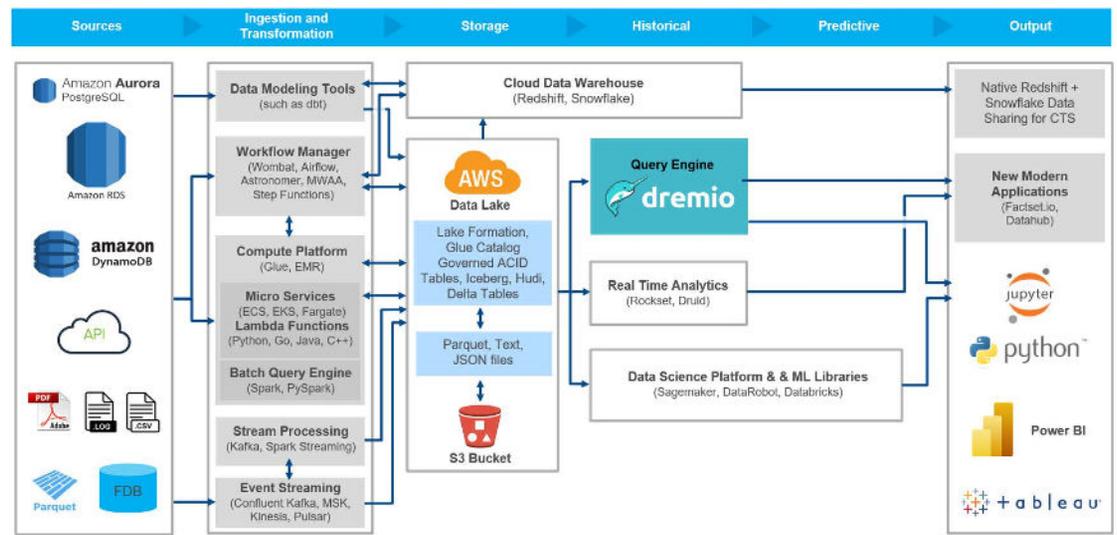


## The Solution

FactSet uses Dremio as a data lake engine across multiple business units over their Amazon S3 data lake. It is a central technology to organize content and enable its products to access data. Dremio powers a custom web application for financial deal transactions that provides financial data and key company metrics about multiple facets of a company, financial performance, acquisition history, risk metrics and more. Each transaction has multiple entities and relationships such as the deal, its buyers, financial information about the company from third parties and investors. They wanted to provide this information to all the entities involved in the transaction so that the analysts can use this information to make better investment decisions.

Due to the complexity of the relationships, FactSet chose Dremio because the performance was better across the board. Instead of every team having to write their own API, they built a GraphQL API on top of the Dremio Arrow Flight interface, so individual products didn't need their own hosted web infrastructure to use the data and don't need to write their own SQL queries.

### FactSet data architecture:



**“Our new Dremio-powered data screening service is giving us an average of over 90% performance improvement over our previous screening method. In one scenario, a query that took 30-50 seconds is now completing in less than 2 seconds.”**

**Matt Topol**  
Vice President, Principal Software Architect, Factset



***“FactSet enables tens of thousands of investment professionals around the world with the data and analytics they need to make crucial decisions. Dremio helped us modernize our existing data infrastructure, achieving dramatic SQL query acceleration, with its reflections capability on top of AWS data lake.”***

**Wilson Tsai**  
VP, Director – Data Platform Engineering,  
FactSet

## The Results

### Modernize and simplify applications with “off the shelf” functionality

FactSet is using Dremio to modernize its data stack without having to rip and replace it so they can provide faster and better services to their clients with minimal interruption. They built their solution on top of their existing tech stack, with existing data sources including the data lake. They could simplify development across multiple content sets. “We basically took Dremio off the shelf and layered services on top of it,” says Wilson Tsai, Director of Data Platform Engineering

### Develop new data screening services that provide huge time savings

They wanted to modernize their data screening capabilities across 30 data domains. It took only five to seven months to be up and running using Dremio for two new data screening services for regulatory and financial data. The new screening median workflow time was reduced by about 25%. Other workloads reported 68-74% time reduction compared to the old screening method. Dremio is enabling significantly faster response times for the product’s analytics/BI capabilities. This also did not require any additional hosting of web infrastructure.

### Eliminate complexity by transforming data silos into a single pane of glass

Dremio helps FactSet transcend data silos, eliminate complexity and provide access to data through a single pane of glass. Users can access data from any source, e.g. DynamoDB database, Redshift, or the data lake. They can make queries about all aspects of a financial transaction and instantly retrieve the data from one central resource.

### Accelerate performance through Dremio Reflections technology

Dremio Reflections technology helped reduce the time to access data from 10 minutes to under 30 seconds. The company looked at Dremio Reflections to replace an internal caching system and found that when they put Dremio technology on top of other existing technologies, it actually gave them more performance than the SQL databases in the underlying data warehouses. Reflections provided exceptional performance across the board, unlike a specialized graph database. So FactSet’s engineers no longer need to spend time and resources on caching and the movement of data and can instead focus on data models higher in the stack than caching. This took the costs out of their solution and resulted in a better utilization of FactSet’s engineering resources.

### Easily trace data sources with graph lineage

With multiple continuous data sources, it is important for FactSet and its users to be able to refer back to the original source of financial data. FactSet uses Dremio’s graph lineage technologies to easily trace back the origin of the data source, increasing the reliability of the data.

## ABOUT DREMIO

Dremio is the easy and open lakehouse platform. Data teams use Dremio to deliver self-service analytics, while enjoying the flexibility to use Dremio's lightning-fast SQL query service and any other processing engine on the same data.

Dremio and the Narwhal logo are registered trademarks or trademarks of Dremio, Inc. in the United States and other countries. Other brand names mentioned herein are for identification purposes only and may be trademarks of their respective holder(s). © 2022 Dremio, Inc. All rights reserved.



## Deploy Dremio

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

## CONTACT SALES

[contact@dremio.com](mailto:contact@dremio.com)