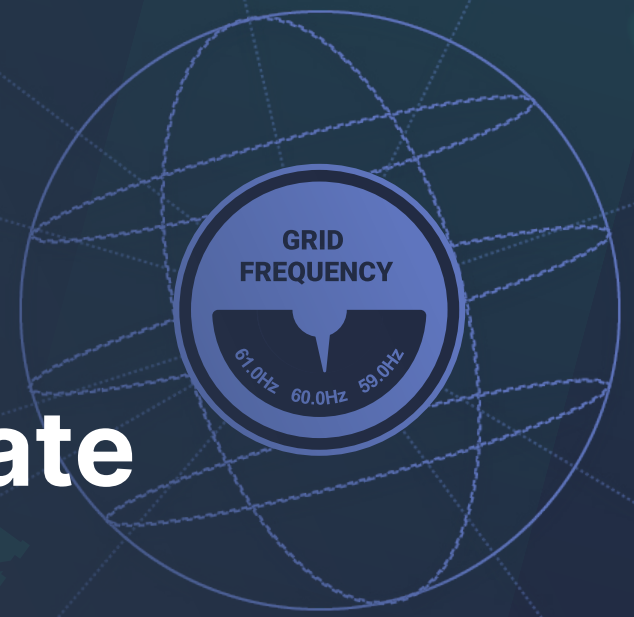




# Dremio Helps Leap Boost Productivity and Accelerate Growth with Self-Service Access to Energy Data



Challenges	Solution	Results
Large, rapidly growing datasets. No self-service access to data.	Leap selected Dremio to provide simple, self-service data access to their team of roughly 15 analysts and data scientists to address these challenges.	Self-service data access for analysts delivering a 30%+ productivity gain.
Difficult and slow to answer business-level questions. Custom scripts hard to build and maintain.	While other tools limited the download of data extracts to between 5,000 and 10,000 rows, Dremio could easily extract a million rows.	Improved data governance. Reduced data engineering workload.
Challenges organizing and maintaining datasets. Need to contain cloud spending	After just a few hours, data engineers realized they could access Parquet data on S3 with Dremio and make it available to analysts through a standard SQL interface.	More efficient use of cloud infrastructure. Faster, higher-quality decisions for improved competitiveness.

## CUSTOMER



<https://leap.energy>

## GEO

USA, Netherlands

## INDUSTRY

Energy

## OBJECTIVES

Build a self-service data management platform to improve analyst productivity, decision quality, and customer service. Position Leap for future expansion with better data access, organization, and governance.

## DATA ENVIRONMENT

- **Ingestion:** Custom PySpark pipelines
- **Storage:** AWS S3
- **Compute:** Dremio
- **BI Client:** Looker
- **Analytics:** Weights & Biases
- **Metadata:** DataHub

## Summary

Leap selected Dremio to provide self-service access to collected data from thousands of energy meters, dramatically improving analyst productivity, reducing data engineering workload, and improving the quality and timeliness of business decisions.

## The Business

While timely access to accurate data is important for any business, at Leap, data is the business. Leap facilitates a dynamic energy exchange where households, businesses, and utilities all interact, responding to real-time market pricing signals representing the needs of the grid. In essence, Leap enables virtual power plants by collaborating with partners and providing flexible load to wholesale markets.

Using data collected from smart meters, individual and commercial users with connected energy devices and behind-the-meter battery systems can generate revenue by reducing load during busy periods and offering excess capacity back to the grid. This helps utility operators and energy partners unlock new value streams, increase customer engagement, achieve sustainability goals, and realize more stable and predictable demand. Leap works with multiple load types from smart residential thermostats to fleet-operated EV chargers to commercial battery storage.

## The Challenge

Effective management of big data is critical to Leap. Readings are collected from tens of thousands of smart energy meters using Spark-based data ingest pipelines and are landed in Parquet format in S3 storage. These vast data sets need to be stored and combined with partner and customer data residing in PostgreSQL for analysis. For Leap's vast and growing datasets, ingesting data into a traditional database was not an option. Leap needed a scalable and cost-efficient solution to query and analyze data in place.

Early in its evolution, Leap had no easy way to expose and visualize the vast amounts of collected data to internal analysts. Data Engineers relied primarily on custom Python scripts to extract datasets in CSV format from raw meter readings and made them available in S3 buckets for analysts to download. The extraction process was slow, and custom scripts were tedious to develop and difficult to maintain. Also, the proliferation of disconnected data files and extracts made it challenging to organize and govern data effectively. While open source and commercial tools could query S3 data directly, solutions were slow, difficult to operate, and typically required the 7x24 deployment of expensive fixed-sized clusters.



***“While accurate and timely information is important to any business, data is our lifeblood. Everything we do is based on data, so results must be 100% correct.***

***We had underestimated how much data productivity would take off. Analysts are now able to easily derive and share new data views in Dremio’s semantic layer, organize data into folders, and easily answer business-level questions without the need for custom scripts or coding.”***

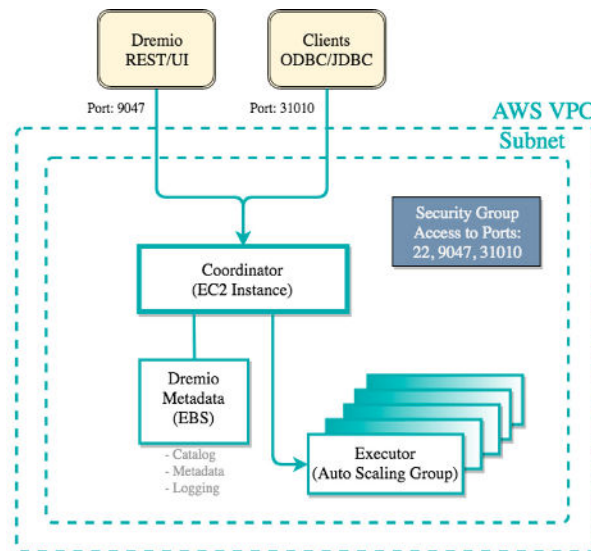
**Marco Rietveld**  
Lead Data Engineer, Leap Energy

## The Solution

Leap selected Dremio to provide simple, self-service data access to their team of roughly 15 analysts and data scientists to address these challenges. Data engineers were heavily invested in Spark-based tools and learned of Dremio through its association with the open-source Apache Arrow project.

Dremio was selected based on a successful hackathon. After just a few hours, data engineers realized they could access Parquet data on S3 with Dremio and make it available to analysts through a standard SQL interface. Dremio rapidly emerged as a preferred solution because it was easy to deploy and use, delivered outstanding query performance, and enabled analysts to extract large datasets easily. While other tools limited the download of data extracts to between 5,000 and 10,000 rows, Dremio could easily extract a million rows. Support for large data extracts was essential because analysts often needed years’ worth of data to analyze annual and seasonal usage patterns.

A critical capability was seamlessly joining tables residing in Parquet files with partner data residing in PostgreSQL tables. With Dremio, data extracts and joins previously requiring custom Python scripts could be accomplished with a simple SQL query. Because Dremio exposes data via standard SQL interfaces, Leap could use Looker, a BI tool for dashboarding and visualization. Exposing data through a BI tool further improved productivity, enabling new visualizations and helping Leap spot trends quickly and gain new insights from collected data.





***“It’s hard to anticipate problems that you don’t know about until you run into them. It was nice that Dremio had already thought about these problems and baked data governance functionality into the product.”***

**Marco Rietveld**  
Lead Data Engineer, Leap Energy

## The Results

With Dremio, anyone in the company can now access energy data and calculations across tens of thousands of meters in a matter of seconds. Dremio has helped unleash analyst productivity and creativity. With Dremio’s semantic layer, analysts can easily combine data in new and interesting ways to gain business insights without requiring the assistance of data engineers.

More importantly, Dremio is helping build a data foundation that can scale and serve as a platform for future innovation. By helping address a variety of technical and operational challenges related to data management, Dremio is helping position the company to expand into new markets and pursue additional opportunities.

### Self-Service Access For Analysts

With Dremio, analyst productivity has exploded. Leap estimates an overall productivity gain of more than 30%. Terabyte-scale queries that previously took minutes now return in seconds - a game-changer for busy analysts that spend their entire day responding to partner inquiries and developing predictive models.

“We had underestimated how much data productivity would take off,” says Marco Rietveld, lead data engineer at Leap. “Analysts are now able to easily derive and share new data views in Dremio’s semantic layer, organize data into folders, and easily answer business-level questions without the need for custom scripts or coding.”

### Improved Data Governance And Regulatory Compliance

Within weeks of deploying Dremio for self-service data access, usage grew rapidly, and analysts began developing and sharing derived data views. Dremio gave analysts the tools they need to document and logically organize datasets in shared spaces with appropriate security controls. “It’s hard to anticipate problems that you don’t know about until you run into them,” says Marco Rietveld. “It was nice that Dremio had already thought about these problems and baked data governance functionality into the product.”

The energy sector is highly regulated. Data needs to be 100% correct, and regulators may challenge providers at any time to demonstrate how they arrived at decisions. Compliance requires easy access to prior datasets and calculation reproducibility. Dremio provides the capabilities needed to help ensure compliance, including access controls, auditability, and data lineage tracking.



***“Dremio gets rid of all the technical barriers to accessing data. What now takes 20 seconds might have previously required a day of work. It was amazing how quickly productivity improved when analysts could quickly and easily query and organize data themselves.”***

**Marco Rietveld**  
Lead Data Engineer, Leap Energy

### Reduced Data Engineering Workload

Before deploying Dremio, analysts relied on data engineers to develop custom PySpark scripts to extract datasets in CSV format from raw meter data in S3. Whenever analysts required access to new data, data engineers needed to develop, test and deploy new PySpark pipelines leading to a backlog of requests. By exposing data with Dremio, analysts can directly query Parquet files on S3, PostgreSQL, and other sources. They can also join data from multiple sources using simple SQL queries.

Not only does this improve analyst productivity, but it dramatically reduces data engineering workload. Reducing the data engineering backlog increases business agility and helps Leap respond faster to changing market conditions and competitive pressures.

### More Efficient Use Of Cloud Infrastructure

Unlike some query tools that require the continuous deployment of expensive cloud infrastructure, Dremio supports Elastic engines enabling organizations to right-size cloud resources for each distinct workload. Rather than relying on a “one-size fits all” model, Dremio can schedule execution engines to run independently at different times and automatically start and stop engines based on workload requirements at runtime. This allows organizations to isolate workloads and maintain service levels while dramatically reducing cloud infrastructure spending.

### Faster, Higher-Quality Decisions For Improved Competitiveness

Dremio frees analysts and data scientists to focus on the business. It provides a way for anyone in the company to access energy data and calculations on tens of thousands of meters within seconds. With easy access to data, analysts can respond faster to customer inquiries and quickly detect and resolve anomalies in data. With Dremio, productivity has been unleashed. Analysts can now spend more of their day analyzing and understanding energy usage patterns and building improved predictive models – activities that are critical to optimizing operations and boosting revenue and profitability.

***“The traditional thinking was that to get adequate performance, we would need to deploy a large cluster right upfront. With Dremio’s elastic query engines, we get excellent performance but only pay for resources when we need them.”***

**Marco Rietveld**  
Lead Data Engineer, Leap Energy

## ABOUT DREMIO

Dremio reimagines the cloud data lake to deliver faster time to analytics by eliminating the need for expensive proprietary systems and providing data warehouse functionality on data lake storage. Customers can run mission-critical BI workloads directly on the data lake, without needing to copy and move data into proprietary data warehouses or create cubes/aggregation tables/BI extracts. In addition, Dremio's semantic layer provides easy, self-service access for data consumers, and flexibility and control for data architects. Dremio delivers the world's first no-copy architecture, drastically simplifying the data architecture and enabling data democratization.

## Deploy Dremio

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

## CONTACT SALES

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

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.