

CASE STUDY

Dremio Data Lake Engine Accelerates Insights and Drives Efficiency in Henkel's Supply Chain



CUSTOMER



<https://www.henkel.com>



GEO

HQ in Germany, with worldwide operations



INDUSTRY

Chemical and consumer goods



OBJECTIVES

Henkel wanted to improve data analytics by joining data silos, accelerating business insights and improving productivity in their global supply chain.

CHALLENGES

- Data silos between supply chain functions hampered productivity
- No transparency or visibility into data across the supply chain
- Extremely slow data query speeds
- Huge, unstructured datasets difficult to access
- Expensive technology maintenance costs

SOLUTION/DATA ENVIRONMENTS

The elements of the solution include:

- Ingestion: ADLS Gen 2, Synapse
- Storage: ADLS Gen 2
- Compute: Dremio, Databricks
- Clients: Tableau, Power BI, Excel

RESULTS

- Increases efficiency (OEE) >10% by sharing data across silos
- Cost savings through supply chain improvements
- Reduces query speed by 30x, from 3-4 minutes to 8 seconds
- Accelerates time to business insights

Summary

Henkel uses the Dremio data lake engine to join data silos, accelerate business insights and save millions through productivity improvements in their global supply chain.

The Business

Henkel is an international chemical and consumer goods company, headquartered in Germany, that operates worldwide, offering leading brands and technologies in three business areas: Adhesive Technologies, Beauty Care, and Laundry and Home Care. The company employs more than 52,000 people worldwide, and North America accounts for 26 percent of the company's global sales.

The Challenge

Henkel's Laundry & Home Care division sells laundry detergents and household cleaning products and accounts for nearly 6.6 billion euros in sales and one-third of the company's core business. The division generates massive datasets in its supply chain for demand planning/forecasting, supply network planning, production scheduling, manufacturing and logistics for 33 production plants, 70 contract manufacturers and 60 warehouses around the world.

In 2016, Henkel struggled to connect the data silos across its supply chain to get insights to better manage the business. They did not have visibility into the data across the silos. They relied on Microsoft Excel reports that were generated weekly, monthly or quarterly. When combining Excel reports with more than 1.4 million rows, it was challenging to integrate data across different silos and functions. As a result, they were only able to look at one aspect of the supply chain at a time or had to turn to external consultants and vendors to do the cross-functional analysis.

Henkel's Laundry & Home division started the process of improving their data analytics by implementing Cloudera and Apache Spark™ in 2017. While this platform was a step in the right direction, there were still a lot of challenges. Cloudera required a large team to maintain the platform,

it did not scale well and query times were slow. To address these limitations the team responsible for the Henkel Data Foundation, a highly integrated data management, processing and analytics platform based on fully elastic cloud technologies, recommended that the Laundry & Home division evaluate Dremio. “We recognized the challenges with our platform when we were introduced to Dremio by the Henkel Data Foundation and saw how powerful it was,” explains Tarun Rana, Corporate Senior Manager Digital Transformation at Henkel.

Henkel wanted to build a data lake solution that would increase the agility of advanced analytics, be able to support unstructured data and create data labs for their data scientists. “With its strong query performance and semantic layer capabilities, Dremio is the perfect backbone for our Henkel data lake,” says Thomas Zeutschler, Director of Data and Application Foundation at Henkel.

The Solution

Over the past year, Henkel has migrated from Cloudera and a data warehouse architecture to a new solution platform based on Microsoft Azure Data Lake Storage (ADLS), Dremio, Databricks and Tableau. They can now natively analyze data stored in ADLS and leverage Dremio for data joins, filtering and transformations. Using Dremio to curate data as virtual datasets also enables Henkel to significantly

increase the speed of critical supply chain dashboards in Tableau. “Dremio is an essential element of our solution that takes away a lot of pain in maintenance and cost from our front-end tool Tableau,” says Rana.

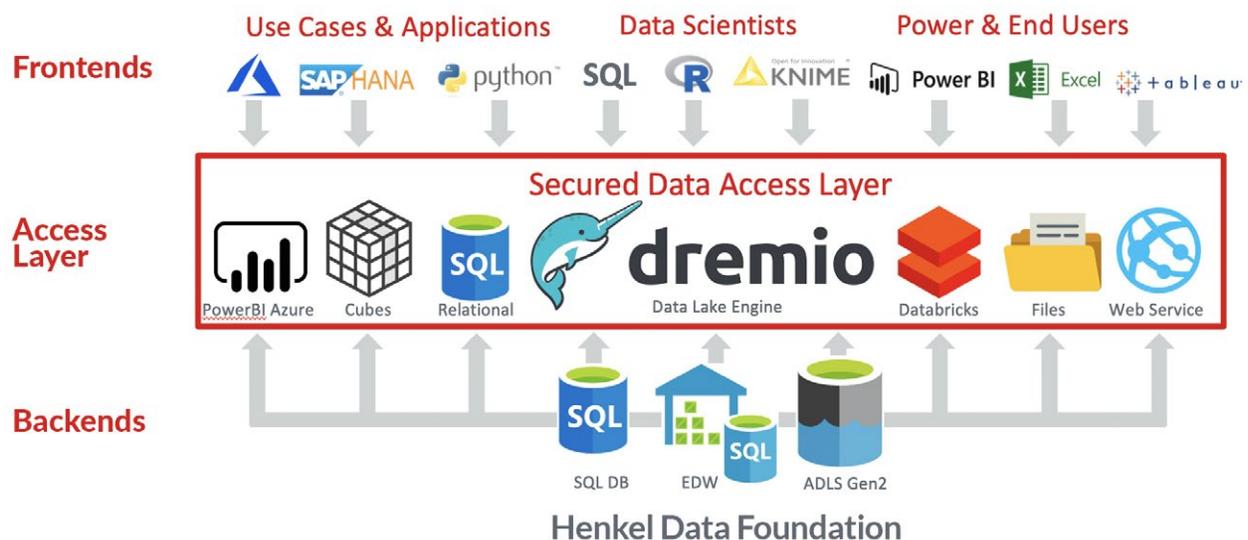
Both Dremio and Databricks play a crucial role in Henkel’s secured data access layer. Henkel uses Databricks mostly for the automatic pre-processing of mass data from the Henkel Data Foundation. Dremio enables the easy implementation of use cases for IT and end users.

The Results

INCREASES VISIBILITY AND TRANSPARENCY

By integrating data silos, Henkel has dramatically increased data visibility and transparency, providing one source of truth for their business data. “Back in 2016, we would come to meetings and everyone would bring their own charts. We would spend the first 15 minutes of the meeting to bring the different charts together in a meaningful way. It took a lot of time,” Rana recalls. “Today our data quality is much higher, and the facts are clearer. We can leverage data across different functions, which we were not able to do before. When you have this capability, you can generate valuable insights to bridge the gap between supply chain planning and production.”

Henkel Data Foundation Architecture Overview



COST SAVINGS THROUGH DATA-DRIVEN PRODUCTIVITY IMPROVEMENTS

In the past, when production planners were asked how many bottles of laundry detergent the line could produce per minute, they would typically make a conservative estimate of the line speed/capacity, not wanting to overpromise.

With significantly increased visibility in their production capacity, Henkel can base their planning on reality. "Once we had dynamic data planning tools, we started to compare plan speed vs. actual speed and saw a huge gap in our Overall Equipment Effectiveness (OEE), a key metric that measures how efficiently our production lines are running," Rana says.

"We now have live, dynamic parameter settings in our planning tool, which has resulted in a >10% OEE increase since the introduction of the system. This is huge. We have 33 production plants with around 400 lines, with 250 connected in real time. It's a huge productivity increase and advantage, and just one example of how data drives cost savings," says Wolfgang Weber, Head of Digital Transformation Laundry and Home Care.

REDUCES COMPUTING AND MAINTENANCE COSTS

Previously, Henkel had been extracting data from multiple sources to Tableau because Tableau could not connect natively to multiple data sources. The scale of data was exploding at such a massive pace that Henkel would need to invest in hardware every six months to continually expand computing power and maintenance. It was becoming hard to justify these numbers that were not hitting any supply chain KPI.

"If you want a stable environment, you need something in between that connects to the data sources, and Tableau just

connects to one. Dremio solved that problem, which has been huge in terms of maintenance and cost," Rana says. "Now that loads are managed by Dremio, we no longer need to keep expanding our computing hardware. This had a big impact in terms of cost savings."

REDUCES QUERY TIMES BY 30X

The old system only had one specific dashboard with a live connection to a 3.5 billion-row dataset. If you wanted to do a query with a specific filter in Tableau, for example looking at a forecast for one brand, it would take 3-4 minutes to execute a query, which was an extremely long time to wait for an answer.

Using Dremio, Henkel reduced its query time 30x, from 3-4 minutes to 8 seconds. "End users immediately noticed the increase in query speed. It was so much faster that some people thought we must be dropping some data or searching through less data," Rana says.

Dremio also enabled Henkel to clean up and improve the quality of their data. "There were big structured datasets with a lot of noise in the data," he says. "We used Dremio to optimize the data with reflections, exploration, aggregation and partitions. That capability is only available in Dremio."

The Laundry & Home Care division now has >500 Tableau dashboards where business users can get fast answers to practically any supply chain query about demand and supply planning, production or inventory and use those insights to make rapid and informed business decisions.

Going forward, Dremio will be used as the backbone for a wide range of upcoming projects. "We are seeing a fast-growing Dremio community within Henkel, and are convinced of the value of the tool," Zeutschler says.



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.