

CASE STUDY

Maersk Scales Analytics from Zero to 1.6 Million Daily Queries with Dremio

At a Glance

The Customer



Challenge

Maersk struggled with unreliable data infrastructure that experienced hours of downtime weekly, preventing user adoption and limiting their ability to extract value from complex shipping and logistics data across their global operations.

Solution

Implemented Dremio's open data lakehouse platform with a robust, scalable architecture using Iceberg format, enabling self-service analytics across their trucking, shipping, warehousing, and cargo airline operations.

Results

- Achieved 99.97% uptime
- Scaled to 1.6 million daily queries serving 3,000 direct users and 5,000 downstream users
- Plans to reach 25,000 daily users by 2024, all maintained by just two semi-dedicated team members.

The Customer

Maersk is a global transportation and logistics company operating far beyond the shipping containers most people associate with their brand. The company manages a comprehensive logistics ecosystem including trucking operations, warehousing facilities, cargo airlines, and port operations worldwide. With such diverse operations spanning multiple continents, Maersk handles extraordinarily complex data scenarios where individual shipping containers might contain goods from multiple customers, each requiring detailed customs documentation, import paperwork, and tracking across their journey through the global supply chain.

The Challenge

Despite having access to vast amounts of operational data from their warehouses, aircraft, ships, trucks, internal systems, and port operations, Maersk faced a critical adoption problem with their initial data infrastructure. After following best practices and building what appeared to be a well-architected system, they discovered that users simply weren't engaging with the platform. For six months, the system experienced literally zero queries on some days, leaving the team questioning what had gone wrong.

The root cause became clear: reliability was fundamentally broken. The system suffered from hours of downtime every week, creating a user experience that drove people away from the platform entirely. Users would quickly abandon any system that couldn't deliver consistent performance, regardless of its other capabilities. This reliability crisis was compounded by the complexity of their data, where shipping containers split across multiple customers created intricate data relationships that required robust processing capabilities. The team realized they needed to completely rebuild their infrastructure to meet user expectations for reliability and performance.

The Solution

Maersk made the strategic decision to rebuild their entire data infrastructure using Dremio as their query engine foundation. They implemented a comprehensive three-layer architecture using bronze, silver, and gold data layers that could handle their complex logistics data at scale. The solution leveraged Dremio's integration capabilities to connect their diverse data sources while running within a Kubernetes environment that could scale up and down to optimize costs.

A critical component of their solution was standardizing on Apache Iceberg format through Dremio, which prevented vendor lock-in and eliminated the need to pay additional fees to access their own data. This open format approach aligned with Maersk's principle that organizations shouldn't have to pay proprietary vendors to access data they own. The team also implemented Apache Superset for dashboarding to reduce their Microsoft PowerBI costs while maintaining high performance.

The rebuilt architecture supported both direct SQL users and downstream applications, with seamless integration to PowerBI that eliminated the need for expensive Azure Analytics Services. By allowing Dremio to handle the heavy lifting of query processing and data management, Maersk could focus on delivering value to their users rather than maintaining complex infrastructure.

Results

The transformation results were remarkable. Maersk achieved 99.97% uptime, experiencing only 3 hours of downtime in 10 months compared to several hours weekly in their previous system. This reliability improvement drove massive user adoption, scaling from zero users to 3,000 direct users running SQL queries and reports, plus 5,000 downstream users accessing data through PowerBI and other applications.

The platform now processes 1.6 million queries daily, with projections to reach 25,000 daily users by 2024 as they migrate SAP reporting workloads to Dremio. Most impressively, this entire operation is maintained by just two semi-dedicated team members, demonstrating exceptional operational efficiency and low total cost of ownership.

The mixed workload performance has been outstanding, with interactive users receiving quick responses while batch processes run efficiently in the background. This has enabled Maersk to replace expensive proprietary solutions while maintaining superior performance, proving that the right architecture and technology choices can deliver both cost savings and improved user experience at enterprise scale.

ABOUT DREMIO

Dremio is the intelligent lakehouse platform for the business, serving hundreds of global enterprises, including Maersk, Amazon, Regeneron, NetApp, and S&P Global. Based on open-source technologies like Apache Iceberg and Apache Arrow, Dremio provides an open lakehouse architecture enabling the fastest time to insight and platform flexibility at a fraction of the cost.

Learn more at www.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). © 2025 Dremio, Inc. All rights reserved.