dremio

CASE STUDY

NCR Uses Dremio to Deliver Business Insights at a Faster Clip

CUSTOMER

NCR

https://www.ncr.com/company

GEO
United States

INDUSTRY
Information Technology

OBJECTIVES
NCR wanted to accelerate analytics time to value. The first high-priority use case was to develop quote-to-cash metrics to identify bottlenecks and opportunities for improvement.

CHALLENGES
- **Slow analytics development due to functional silos created among experts in different data repositories**
- **Required data from many functional parties**
- **The business units were not satisfied with rigidity or the long turnaround time for data requests**

SOLUTION/DATA ENVIRONMENTS
- **Ingestion: Sqoop from legacy platforms to Hadoop**
- **Storage: Hadoop**
- **Compute: Dremio**
- **Client: Tableau**

RESULTS
- **Dremio accelerates the transition from a traditional data warehouse architecture to a modern, open data platform**
- **Dremio shortens development time and reduces development costs—NCR engineers can now re-create data pipelines in days, a task that previously took more than two months**
- **Dremio accelerate queries 10x to 30x. Typical complex queries are delivered in less than 30 seconds**
- **Significant performance improvement increases efficiencies and reduces "revenue leakage" from time wasted waiting for analyses**
- **Faster insights improve NCR's ability to provide exceptional service to its customers and offer tailored services**

Summary

To meet the demand for faster data insights, increase analytic capacity and contain costs, NCR chose Dremio to help modernize its data analytics infrastructure.

The Business

NCR Corporation is a leading software- and services-led enterprise provider in the financial, retail, hospitality, telecom and technology industries. Established in 1884 and headquartered in Atlanta, the company has 34,000 employees and does business in 180 countries.

NCR is currently the leading global point-of-sale (POS) software provider for retail and hospitality, and the top provider of multi-vendor ATM software. NCR creates software, hardware and services including self-service kiosks, POS terminals, automated teller machines, check processing systems, barcode scanners and business consumables. The company creates software, hardware and services that run the enterprise from the back office to the front end and everything in between.

The Challenge

NCR takes a proactive approach to its data management and analysis practices. The company's enterprise data team partners with business leaders to identify new opportunities for business innovation by leveraging the company's massive and growing datasets.

To that end, NCR's business and IT teams took on a daunting challenge: map the entire customer experience at NCR for each customer by plotting data points from the beginning when the company first communicates with a new or potential customer and generates a quote, to when an invoice is issued and a customer completes payment for an order. By mapping this data, analysts can identify new opportunities for better customer service. For example, they can identify bottlenecks in order processing, and evaluate the company's ability to consistently meets delivery commitments.

The teams dubbed the project the Quote-to-Cash (QTC) Analytics initiative. Deployment of the QTC project required the alignment of many different data sources from partners and different data repositories throughout the enterprise, and the normalization of disparate types of data to build a complex data visualization tool.

NCR generates massive amounts of data every day—sales, device service data, IoT status, transactional data, shipping and supply chain data, and much more. In addition to the QTC project, the company needed to modernize and expand its data analytics infrastructure to provide faster and easier access to insights in order to deliver a better, more tailored service to their broad customer base and discover new opportunities for improvement and growth. "This type of capability we were aiming for in our analytics programs requires sophisticated and scalable data platforms," says Ivan Alvarez, IT vice president, big data and analytics, NCR Corporation.

For many years, NCR stored a majority of their datasets in a data warehouse system. Beginning in 2019, the company's IT executives and data stakeholders set a goal to accelerate time to value for analytics requirements. The move was intended to provide several important benefits: long-term cost-effectiveness, improved scalability and increased capacity. It also gave the company an opportunity to perform a complete audit of its legacy data storage in order to eliminate unnecessary data duplication and establish a single source of truth among the company's datasets.

The transition to the new platform was challenging. Moving their data pipelines, project by project, to the data analytics infrastructure took two to three months for each of the company's many critical and large datasets. Further, the migration required several outside consultants to map legacy schemas and produce the required dashboards.
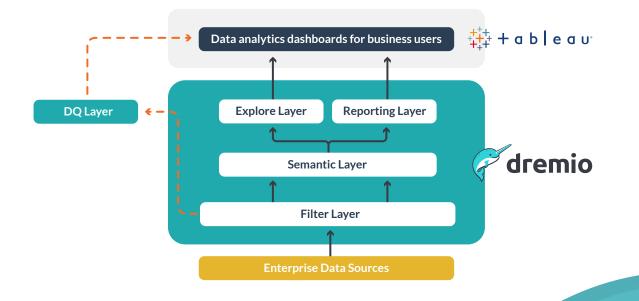
NCR also needed to operate across two different data repositories during the transition—the legacy system and the new hybrid infrastructure. To give its users a common language and interface between the two repositories, the company needed a method to build a semantic layer on top of the two datasets, providing a single source of truth and unified definitions for its attributes. This strategy enables business users to be more self-sufficient with their data, pulling their own queries from trusted data sources without having to engage and wait for involvement from IT partners. The QTC Analytics initiative was one of the first, and the largest, explorations of the semantic layer in production at NCR.

The Solution

After a two-week POC, NCR selected Dremio as a key element in their modernized data analytics infrastructure. The POC proved Dremio's impact on performance and demonstrated how NCR could re-create data pipelines in days rather than weeks.

NCR deployed Dremio between the data stored in the data lake storage and its client analysis tools. Dremio was configured with two coordinators in an active-standby configuration and 69 executor nodes across the YARN queue on Hadoop in a 1-to-1 node configuration. From there, NCR used Dremio's web UI to build the semantic layer and expose reports for analysts. As a result, analysts can now get live, interactive performance for their Tableau dashboards, with their queries running directly on the data lake.

The massively parallel processing nature of Dremio enables NCR to scale their query processing across multiple executors in parallel for data warehouse-like performance against the raw data in the data lake. The raw and physical data reflections feature further accelerates query performance.

The Results

SOLVES INTEGRATION CHALLENGES AND DELIVERS FASTER INSIGHTS

Before Dremio, the transformation and migration had been slow, expensive and yielded poor performance results. With Dremio, NCR is able to directly query data from the data lake with BI tools that are effective and timely. By allowing users to access the data directly through a well-formed semantic layer on top of the data lake, development time for reports and dashboards has been reduced from months to days.

This success has accelerated NCR's move to a more modern approach to making federated information available to multiple functions using a common language. ETL processes are now processing data in the data lake, which saves time and money by eliminating the additional loading and transporting of data between systems and software platforms. Dremio bridges the data warehouse and the data lake, enabling NCR to derive more value between the two data sources. "This integration allows NCR to cross-pollinate data engineering knowledge among platforms and, most importantly, to deliver faster data insights to our internal and external customers," says Alvarez.

QUOTE-TO-CASH ANALYTICS INITIATIVE IS A SUCCESS

The QTC Analytics initiative, the largest NCR project to benefit from the Dremio deployment, proved to be a demonstrable success for the company. NCR's data engineers were able to align many different types of data from a variety of sources into a single tool for its analysts and executive leadership to analyze. The project successfully analyzes data from quotes, delivery dates, invoicing, escalations, purchase revisions and much more—all within a single tool and made possible through the semantic layer powered by Dremio.

ACCELERATES DEVELOPMENT TIME AND DECREASES COSTS

Dremio enables NCR to dramatically reduce what it spends on consultants, a significant expense each month. NCR is now able to re-create data pipelines and associated queries, reporting and dashboards in days instead of months. NCR data engineers with no experience using Dremio are able to reproduce complex table schemas via Dremio's virtual datasets within days. Before the Dremio integration, this had previously taken more than two months to complete.

IMPROVED QUERY PERFORMANCE BENEFITS THE BOTTOM LINE

Dremio accelerates queries significantly compared to NCR's legacy data repository model, driving agility for NCR and enabling faster action on insights. In the past, complex queries could take minutes to execute. Now, the built-in acceleration capabilities of Dremio enables NCR to accelerate queries 10x to 30x, depending on the complexity of the query, volume and concurrency. A typical complex query comes back in less than 30 seconds.

One of the reasons NCR has achieved these extreme efficiencies is through Dremio's data reflections feature—the ability to maintain physically optimized representations of source data transparently to BI users and tools to accelerate queries on very large datasets. Essentially, some of the data can be stored in precalculated forms and used automatically by Dremio's cost-based query optimizer, saving massive amounts of time instead of doing the same calculations on the fly for each query.

ACCELERATES MIGRATION TO THE CLOUD

As a result of these successes, NCR has accelerated its plans to retire its legacy platforms while simultaneously moving to cloud-based solutions. The technique for processing data and reporting using Dremio translates directly to cloud-based computing as well. This increase in performance will have a big impact on the company's bottom line.

dremio