CASE STUDY

Leading Regional Bank Achieves 10x Performance Improvement with Dremio and Apache Iceberg

At a Glance

Challenge

This large regional bank's executive dashboards for their Customer Contact Center had become mission-critical tools despite beginning as an unofficial initiative. Their architecture required complex data movement from their data lake to Amazon Redshift, causing poor performance with queries taking over a minute to run. They needed to simplify their architecture while improving performance and maintaining advanced analytical capabilities like time travel and snapshot analysis.

Solution

The bank leveraged Dremio and Apache Iceberg to eliminate Redshift from their workflow. They transitioned from storing data in Parquet format to Iceberg format directly in their Amazon S3 data lake, with Dremio reading and querying Iceberg tables directly. This implementation took only about a week and a half from architecture approval to a working proof of concept.

Results

- 10x faster query performance (from 80 seconds to 10-12 seconds)
- Eliminated Amazon Redshift licensing and infrastructure costs
- Simplified data pipeline with fewer moving parts
- Maintained all required advanced analytics capabilities
- Enabled executives to be more responsive to customer service trends

About the Customer

The customer is a large regional bank in the United States with a significant retail banking operation. Their Customer Contact Center handles thousands of customer inquiries daily, making it a crucial touchpoint for customer satisfaction and loyalty. The quality and speed of service provided by this center directly impacts customer retention and the bank's reputation in a highly competitive market.

The Challenge

A large regional bank in the United States was running executive dashboards for their Customer Contact Center, providing near real-time analytics on critical metrics such as issue resolution rates, commonality of issues, time to resolution, and talk time. These metrics were essential for monitoring and improving their customer satisfaction scores.

Though the dashboards began as an unofficial initiative created by the data team, they quickly gained popularity with executives and became mission-critical tools for daily decision making.

However, their existing architecture presented several challenges that were limiting their ability to deliver insights efficiently. The bank had implemented a complex data pipeline that required movement from their data lake to Amazon Redshift, adding latency and management overhead to their analytics process. Performance was becoming a significant issue, with dashboard queries taking over a minute to run, creating frustration for executives who needed quick access to information for time-sensitive decisions.

The data team was also struggling to provide advanced analytical capabilities like time travel and snapshot analysis while maintaining performance. As their analytics needs grew, they faced increasing costs from maintaining multiple data platforms and significant management overhead from synchronizing data between systems, creating both financial and operational strain.

The bank needed to simplify their architecture while improving performance and maintaining all the advanced analytical capabilities their executives relied on for effective customer service management.

Why Dremio?

The bank was already using Dremio as their primary platform for data lake analytics, so they were familiar with its capabilities. When evaluating options to improve their Contact Center analytics, they had to make a strategic choice about their data platform direction.

They considered continuing with the existing data lake to Redshift pipeline, which would maintain the status quo but not address the fundamental challenges they were facing. Another option was migrating to Snowflake, which was being implemented for other use cases within the bank and could potentially consolidate their analytics platforms. However, the most promising approach was modernizing their data lake with Apache Iceberg and Dremio to create a more streamlined architecture.

The data team recognized that with Apache Iceberg, they could achieve the functionality they previously relied on Redshift for—directly on their data lake. This approach would allow them to eliminate an entire step in their data pipeline, significantly reducing complexity and potential points of failure. They could also reduce infrastructure and maintenance costs by consolidating on fewer platforms. Most importantly, there was potential to substantially improve performance while maintaining the advanced analytical capabilities like time travel that were essential for their executive dashboards.

The Solution

After receiving approval from their Architecture Board, the bank implemented a solution that leveraged Dremio and Apache Iceberg to simplify their architecture.

Their previous architecture involved several layers and data movements: data was first ingested into an Amazon S3 data lake in Parquet format, then moved to Amazon Redshift for advanced analytics. Dremio was used to accelerate Redshift views, and finally, Power BI dashboards consumed data through Dremio. This multi-step process created latency, synchronization challenges, and higher costs.

The new architecture dramatically simplified this workflow. Data was now ingested directly into their Amazon S3 data lake using the Iceberg table format instead of Parquet. Dremio read and queried these Iceberg tables directly, eliminating the need for Redshift in the workflow. Power BI dashboards continued to consume data through Dremio, maintaining a consistent experience for end users despite the significant backend changes.

The implementation process was remarkably efficient, taking only about a week and a half from architecture approval to a working internal proof of concept. This rapid deployment allowed the bank to quickly validate the approach and begin realizing benefits without a lengthy migration process.

The Impact

The new architecture delivered significant improvements across multiple dimensions, transforming how the bank's executives accessed and utilized customer service data.

Performance improvements were dramatic, with dashboard queries running 10 times faster than before. Queries that previously took 1 minute and 20 seconds now completed in just 10-12 seconds, allowing executives to interact with the data more fluidly and make decisions more rapidly. Notably, these improvements were achieved without even implementing Dremio reflections, suggesting there was substantial potential for further optimization in the future.

Cost savings were realized immediately through the elimination of Amazon Redshift licensing and infrastructure expenses. The bank didn't need to add any additional Dremio resources to support this use case, making the migration even more economically attractive. Operational overhead was also reduced significantly as the team now had fewer systems to manage and monitor.

The operational benefits extended beyond just cost and performance. The data pipeline was now much simpler with fewer moving parts and integration points, reducing the potential for errors and outages. Data synchronization issues that had previously required troubleshooting were eliminated entirely. The team maintained all the required advanced analytics capabilities that executives depended on, ensuring no loss of functionality despite the architectural changes.

Perhaps most importantly, this project enabled the beginning of the Redshift decommissioning process, aligning with the bank's strategic goal of simplifying their technology stack. The executives who relied on these dashboards for daily decision-making now received insights much faster, allowing them to be more responsive to customer service issues and trends, ultimately improving the quality of service provided to customers.

Looking Ahead

The bank plans to continue enhancing their data platform with Dremio and Apache Iceberg as part of their ongoing data strategy. They're exploring Dremio's autonomous reflections feature to further improve performance, potentially reducing query times even more for their most important dashboards.

The team is planning upgrades to newer Dremio versions to take advantage of additional features and performance improvements that have been released since their initial implementation. They're also expanding the use of Apache Iceberg across more of their data lake, standardizing on this format to bring the same benefits to other analytics workloads beyond just the Customer Contact Center.

The data architecture team continues to identify opportunities to simplify their overall data architecture, looking for other areas where similar consolidation approaches could deliver operational and performance benefits. While the bank is currently behind on version upgrades, the success of this project has created momentum for accelerating their adoption of new capabilities, demonstrating the value of modernization to stakeholders throughout the organization.

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.