TELECOMMUNICATIONS GIANT ESTABLISHES UNIVERSAL DATA ACCESS WITH STARBURST ENTERPRISE PRESTO

FORTUNE 100
Telecommunications Company

How Starburst Enterprise Presto is helping a telecommunications company query federated data sources while transitioning to a cost-effective hybrid cloud data architecture.

As a telecommunications giant with more than 20 million subscribers, this customer retains and ingests tremendous volumes of data across various platforms. The company collects daily viewing data in Hadoop, while subscriber and account information and other data are stored in Teradata, Oracle, and various systems. Extracting intelligence from all this data is critical, as data-driven insights help retain existing customers, offer them new packages tuned to their preferences, and more.

Querying all this data would be simpler if it were stored in one platform, but the customer's data is distributed across the above-mentioned platforms and, more recently, Amazon Web Services. "We weren't born in the cloud," said our customer at a recent event. "We've been around for a long time. We have a lot of customers, and a lot of data we're collecting across a lot of different systems."

This case study details why the company chose to deploy Starburst Presto, and how the platform is moving them closer to their ultimate vision of universal data access.
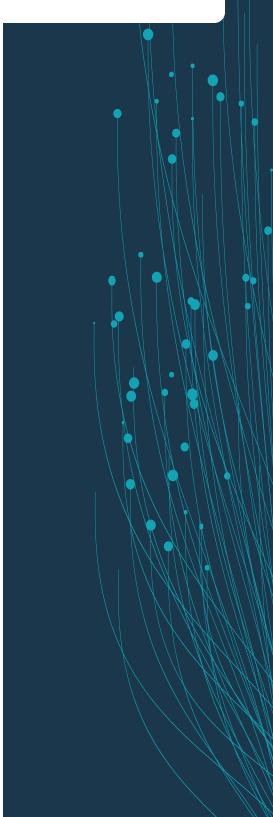
Achieving Universal Data Access

In 2016, the big data team was looking to move data out of Teradata. The system was expensive and underperforming, as it was overloaded with thousands and thousands of daily reports—analytics were the last in line. Yet the company's Enterprise Business Intelligence group still depended on this data and accessed it regularly.

Meanwhile, viewing and event data were streaming into their Hadoop platform, and they often had to join this information with the data in Oracle and Teradata.

"If you wanted to do a big data science project you had to build very large ETL jobs and move data between those platforms," explains our customer. "We didn't have cross-platform queries."

At one point, the company was moving everything into Hadoop, and querying Hadoop directly. Our customer also experimented with Teradata's QueryGrid, which would connect to Hadoop's event data. Ultimately, though, neither of these approaches brought them closer to its desired end state: fast, cross-platform queries, universal data access, and freedom from any one proprietary data warehouse.

The World's Fastest Distributed SQL Query Engine

By deploying Starburst Enterprise Presto as an abstraction layer that operates between the data sources and business intelligence tools, our customer is now able to query data where it resides, and move the data they want to move—all without their end users noticing a thing. "Presto was our secret sauce," said our customer.

> "Presto has connectors to all these different sources. So people didn't have to run ETL. They could just leave the data where it is and just use Presto to connect to these sources."
>
> *Fortune 100*
> *Telecommunications Company*

The number of Presto queries at this customer has scaled rapidly:

- **2,000** queries per day in 2016
- **20,000** queries per day in 2017
- **200,000** queries per day in 2018
- **350,000** queries per day in 2019

"More connections are getting added but the end user experience is not changing," our customer said. "End users are able to access data no matter where it is. This made a huge impact."

Performance at Scale

Previously, analytics were last in line, and users might have to wait days for results. Starburst Enterprise Presto is built to accommodate many users hitting the system at the same time. Our customer uses multiple Starburst-to-Teradata connectors running in parallel to maximize throughput for analytical workloads. Multiple Teradata servers can talk to multiple clusters and hundreds of nodes on the Presto side, ensuring the performance is just as good as querying Teradata directly.

Federated Access

The big data team wanted to empower end users by giving them access to the data no matter where it lives – Teradata, Oracle, Hadoop, or AWS. Similarly, they wanted to ensure that virtual users could access on-prem resources as easily as on-prem users, and vice versa. Starburst Presto makes this possible while simplifying the process for those Business Intelligence users, allowing them to achieve a single point of access to all of this distributed data. They don't know—and don't have to know—where the data lives.

Cost Savings

One of the drivers behind moving away from proprietary data warehouses was cost. Our customer wanted to avoid vendor lock-in and be able to move data to less expensive platforms without disrupting end users. Starburst Enterprise Presto facilitates this in two ways. First, it is agnostic to the data source, and designed to query data wherever it lies. Second, it facilitates ETL, and allows organizations to continue accessing and analyzing data during ETL, so the company can accomplish its larger goals without disrupting the critical daily work of end users. The customer said that Starburst Enterprise Presto also allows them to minimize its labor costs related to these processes.

Toward Universal Data Access & Hybrid Cloud Architecture

Ultimately, Starburst Enterprise Presto is serving two groups inside the customer's organization simultaneously. Presto is giving Enterprise Business Intelligence improved access to more data across more platforms, and helping them extract better insights faster. But Starburst Enterprise Presto is also allowing the company to move data out of those proprietary data warehouses and into the cost-effective & flexible cloud storage, supporting the goals of the big data team.

"When end users are going into on-prem or cloud environments, they will be presented with all the data sets they have access to, irrespective of where the data is located. This offered huge value to our end users."

Fortune 100
Telecommunications Company

![Starburst logo] Starburst
ANALYTICS ANYWHERE

STARBURSTDATA.COM