Starburst
ANALYTICS ANYWHERE

CASE STUDY

Optum Accelerates Queries by 10X with Starburst Enterprise

Information technology service provider Optum is dedicated to shaping a healthcare system that gives patients a complete view of their health, providing them with personalized insights that lead to improved outcomes. A subsidiary of the UnitedHealth Group, Optum uses technology to connect the brightest people, places and ideas across the healthcare ecosystem.

The company's mission depends in part on providing its analysts with fast, secure access to data. Initially, Optum's data lake architecture couldn't support its needs at scale. Tired of poor query performance and inefficient resource utilization, Optum's Advanced Research & Analytics group deployed Starburst Enterprise to improve data access, accelerate time to insight, maintain strong security, and reduce costs.

Ad Hoc Queries in Seconds, Not Minutes

The majority of Optum's data resides in a PB-scale Hadoop data lake, while the rest is siloed in SAS, Microsoft SQL Server, Teradata, and Postgres databases. More than 10,000 users need fast access to this distributed data. Previously, Optum relied on Hive and Spark SQL for analytics. "Our data lake backbone was on a traditional Hadoop infrastructure," explains Optum Principal IO Engineer Mike Prior. "While that approach had its day, it's not flexible. We needed to scale out and separate our compute from our storage without moving the data."

Querying two different databases required copying data from one to another or engaging in an expensive ETL operation, and queries were taking far too long. "Our goal was to have faster access to data," Prior notes. "If an analyst wants to run an ad hoc query, they want the response in seconds, not minutes."

Finally, as a healthcare technology solutions provider, Optum needs to ensure that it appropriately limits access to sensitive Personal Health Information (PHI), and the company wanted a solution that could support these requirements at scale.

OPTUM

> "If an analyst wants to run an ad hoc query, they want the response in seconds, not minutes.
>
> Mike Prior
> Optum Principal IO Engineer

Starburst Enterprise

After evaluating various solutions, Optum chose Starburst Enterprise. Starburst can be deployed on-prem or in the cloud, and Optum ultimately chose the former, running in Kubernetes. Today, the company depends on Starburst and Trino (formerly PrestoSQL) as a high-performance, distributed query engine that gives its users a single point of secure access to all of its data, allowing Optum to query data where it resides.

Starburst, built on open-source project Trino, also removes the complexity of joining different data sources and simplifies and standardizes queries by effectively uplifting everything to ANSI SQL-92. Business Intelligence Analysts and Data Scientists don't need to learn new dialects or techniques, and they have enhanced access to the data they need to generate or discover insights. At Optum, the benefits of the platform include:

Accelerated Ad Hoc Queries

Improved performance was essential for Optum, and ad hoc queries with Starburst are 10X faster than Hive, and 2X to 3X faster than Spark. Not only that, but the platform is consistently fast across different jobs. One user attested that queries which would have taken upwards of five minutes in the company's previous Hive-based environment now finish in under 10 seconds with Trino.

30% Drop in Infrastructure Resources

In addition to improving query performance by 10X, Starburst reduced resource utilization by 30%. One of the primary advantages of Starburst Enterprise is that it separates storage from compute, allowing companies to dial up compute as the need arises, and not pay for resources they aren't actively using. With Starburst, Optum's data largely remains in its data lake, but Trino is deployed in Kubernetes, with different clusters tuned to the needs of different groups. "We're able to spin up and spin down workers as needed," explains Prior, "and we use autoscaling to cover peak demand."

Improved Consistency & Increased Utilization

Prior and his team have found that Trino is a more reliable query engine when faced with larger workloads. Like many other Starburst Enterprise customers, Optum has seen more and more analysts adopting the platform over time. There's no mystery here—when a solution works quickly and effectively, users come to depend on it.

Plus, it's simpler to work through Starburst Enterprise. "Providing users with one endpoint is so much easier," Prior says. "They can use the same familiar tools, but everything is happening faster."

Global Security Management

At Optum, certain users or business groups may need access to PHI, while others should not be able to see patient data. Managing permissions and access policies is essential to Optum's business, and Starburst Enterprise makes this process painless and seamless for Prior and his colleagues. "We just want one place to configure security for all data access and Starburst Trino allows us to do that," he notes.

No More ETL or Data Duplication

Extracting data from one source, transforming this data to make it compatible, and then loading it into another warehouse or data lake is an expensive and time-consuming operation. With Trino, if an analyst needs to query data residing in two siloed databases, no ETL is necessary. Starburst allows analysts to quickly and easily query data where it lies. This reduces data duplication as well, since nothing needs to be copied from one place to another in the first place.

Savings, Success, and Flexibility

Although Starburst Enterprise is now deployed on-prem, Prior appreciates the fact that the platform can be run in the cloud, and that they can make this transition without any disruptive changes. He's open to deploying in the cloud at some point in the future, and possibly establishing a Trino-to-Trino, on-prem-to-cloud connection.

Overall, the advantages highlighted above, including accelerated queries and enhanced resource utilization, are part of a larger impact on the business as a whole. Optum has seen improved customer retention and satisfaction, and the company is anticipating $8M in savings over the long run thanks to analytics insights uncovered with the help of Starburst Enterprise.

> "Providing users with one endpoint is so much easier. They can use the same familiar tools, but everything is happening faster.

Mike Prior
Optum Principal IO Engineer

Starburst
ANALYTICS ANYWHERE