Rakuten accelerates query performance and modernizes analytics program with AtScale

Rakuten

Industry	Use case	Benefits	Key Product Components
RETAIL, E-COMMERCE	SEMANTIC LAYER, ANALYTICS MODERNIZATION, DATA MIGRATION	FASTER TIME TO INSIGHTS,SURGE CAPACITY DURING PEAK, DATA USAGE, OPTIMIZED QUERY RESPONSE TIMES	SNOWFLAKE, TABLEAU

Rakuten, a shopping rewards company, leverages data on shopper behavior, pricing, and commissions to create compelling offers for their customers, who receive cash-back incentives from Rakuten. With over 13 million e-commerce customers and partnerships with 70+ businesses, Rakuten depends on sophisticated analytics and data management to maintain a differentiated offering in the highly competitive e-commerce industry.

While Rakuten originally consolidated data from multiple siloed systems to a single, on-premises data lake built on Hadoop in 2016, they still faced challenges related to maintaining the environment. The sheer electrical costs of hosting their own internal server farm as well as the expensive hardware required presented obstacles for this fast-growing operation. Problematic workloads also caused painful business disruptions as Rakuten began to experience a resource crunch.

AtScale helped the Rakuten team migrate their on-premises Hadoop cluster to a Snowflake cloud data warehouse on Amazon Web Services (AWS). Additionally, AtScale now provides Rakuten with query performance optimization and a single, virtualized view of their data. This freed up both budget and time, with the team working more efficiently in the cloud.

CHALLENGE

Finding an efficient way to process data without resource competition

While Rakuten had moved from their initial SQL database in 2014 to an AtScale-powered Hadoop solution in 2018, this wasn't sufficient and they soon began to experience a resource crunch based on the sheer size of their database. Rakuten s existing architecture meant that business users didn't have the computing resources necessary to work with large datasets. This meant that users disk and processor usage reached critical levels, leading to competition between business units for hard disk access, memory, and CPU time.

The internal team was frustrated with the competition for resources, and the operational overhead and associated hardware and electricity costs also meant the solution was no longer cost-efficient. That, coupled with the continuous processing demands on storage infrastructure, forced Rakuten to consider new solutions for their data needs. They knew they needed more processing capability and flexibility to continue serving their customers effectively.

SOLUTION

Migrating to a cloud-based solution to control the complexity and cost of analytics

AtScale helped Rakuten transition their analytics to the cloud while still retaining all the analytical capabilities they had built, enabling them to deliver consistency for their team. AtScale insulated their Tableau-based reports and dashboards from changes in underlying raw data. While location and schemas changed, the AtScale model was untouched, allowing them to preserve their investment. Through AtScale's semantic modeling capabilities, Rakuten is able to flexibly add new metrics and data definitions that provide consistency across consumption tools. Data virtualization also makes it easier to onboard new data quickly so that it can be queried from BI tools almost immediately.

Once they had moved their data to Snowflake, AtScale was able to help Rakuten better optimize their costs by right-sizing cloud resource consumption based on real-time usage. AtScale also helped smooth out and mitigate user concurrency challenges without requiring additional compute resources and leveraged intelligent aggregates to accelerate query performance while keeping cloud costs down

RESULTS

A consistent, flexible data platform that adapts to their needs

By moving from an on-premises Hadoop data warehouse to Snowflake's elastic, scalable resource model, Rakuten gained additional computing power. This enabled them to maintain responsiveness to queries during peak demand periods, while only paying for what they used. Now Rakuten can maintain flexibility to compensate for crunches, without paying excessively for operational or server overhead.

Because AtScale allowed Rakuten to move from Hadoop to Snowflake seamlessly, the team can now access their data faster and with more consistency. As high-demand queries were shifted to Snowflake, there was an immediate 30% drop in load on the on-premises computing cluster. And, with AtScale's help, Rakuten managed to solve their concurrency issue as well: Rakuten ran a test of 10 terabytes of data with and without AtScale and discovered that AtScale ran the process and at 4x less cost.

And, since AtScale acts as a semantic layer, the switch from where the data was stored did not lead to any interruptions in data access for Rakuten's users. With the same data analytics tools in place before and after the migration, Rakuten didn't need to retrain their team to adopt any new tools or working methods. The only thing they had to adapt to was receiving faster reports than before.