Industry: Media

Gaia Boosts Subscriber Engagement and Data-Driven Decision-Making with Qubole

Business Problem Overview

Gaia's business model depends on engaging viewers with unique and innovative streaming content. So, recommending the right content, to the right viewers, is critical. But the company's legacy SQL rule-based recommendation engine "was very slow, very tedious, and not very accurate," says senior data engineer Patty Vonick. It did little to drive viewer engagement, Vonick added.

The company also had another problem: its technology architecture—which centered around an on-premises single-instance server and a Postgres data warehouse—was not sufficiently flexible or robust. Specifically, it couldn't provide users access to data from different sources and could not handle the data workloads. In addition, modeling processes took too long, and frequent outages led to hours of delays and debugging. For example, some jobs had to be scheduled to process overnight, but these often ran long. So,when employees generated queries and reports the next day, they would compete with the overnight jobs for compute resources. As a result, the system would overload, adding to the backlog and even resulting in failures. And when jobs failed, it could take hours or even the whole day to troubleshoot them.

Gaia needed to solve these problems. It also needed to adopt more state of the art practices like data analytics and machine learning. The company turned to Qubole for assistance implementing a data lake platform and migrating from the outdated data warehouse infrastructure.

About Gaia

Gaia is a member-supported streaming video subscription service available in 185 countries around the world. Using a powerful combination of modern technology and ancient traditions, Gaia produces and curates transformational video content that includes guided yoga and meditation instruction, as well as series and films covering a wide variety of topics, from health and longevity to human transformation and science, all of which aim to empower the evolution of consciousness.

Using Machine Learning to Generate Data-Driven Content Recommendations

With its new architecture in place that leverages Qubole on AWS, job one at Gaia was to replace the legacy Postgres SQL rules-based recommendation engine with one that was quicker and easier to use, and that returned more relevant results. This new machine learning (ML) recommendation engine—based on Apache Spark and XGBoost models on Qubole—generates data-driven content suggestions to help subscribers decide which videos to watch next.

In the eight months since the new engine went live, the results have been impressive. "We've seen a 50 percent lift in average minutes watched"—a critical viewer-engagement metric—says Gaia product data analyst Patrick Lawlor. "We would not have been able to do that before Qubole." In addition, subscriber engagement has significantly improved..

> *Qubole enabled us to use machine learning to provide much better recommendations than the legacy Postgres data warehouse SQL rule-based engine we used to have."*
>
> **Patrick Lawlor**
> Product Data Analyst, Gaia

Improved Analysis and Reporting for Data-Driven Decision-Making

Before Gaia partnered with Qubole, its reports from available company data were incomplete and lacked the business insights needed for decision-making. This was due in part to a technology infrastructure that couldn't handle the workloads, as well as to a data architecture that was inadequate for drawing data from multiple sources.

Qubole enables Gaia to easily query data from a variety of sources—including AWS data repositories and email, financial, and customer-service platforms—to surface critical business insights. So, "It's possible to dig not only one layer down, but three, four or five layers to see why our numbers are what they are," says Andrew Koblitz, senior manager of financial planning and analysis.

There's not just more and different data at the company's disposal—more than 66 terabytes of it reside in the company's new data lake. There's better data. This is because Qubole facilitates the validation of data before it's used for reporting and analysis purposes. So, company leaders can make data-driven business decisions with greater confidence than ever before.
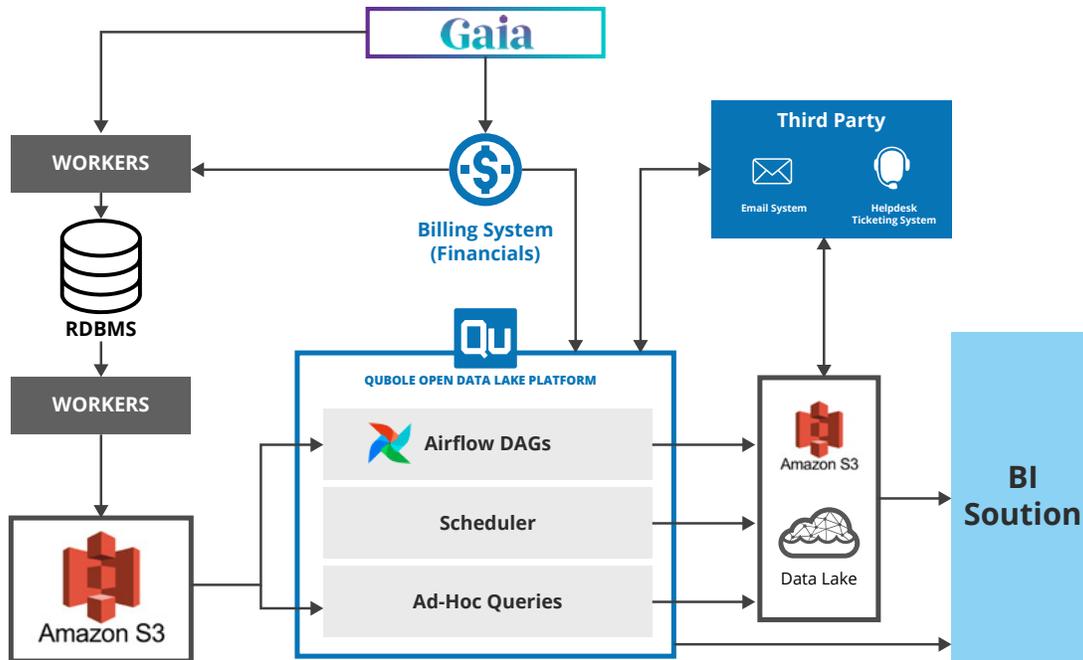
Resilience, Reliability, and Availability

Due to years of band-aids and workarounds, Gaia's legacy technology architecture was complex and fragile. Outages were common—and time consuming. "If an overnight process failed, fixing it was what you did for the rest of the day," recalls data engineer Alex Mendoza. Even when overnight processes didn't fail, they sometimes ran long, extending into working hours—a product of limited compute power. This often resulted in a logjam effect that prevented users from accessing critical data.

Since Gaia implemented Qubole, it's a different story. Now, the system automatically scales up to complete processes, ensuring users always have access to the resources they need. And the system is stable and reliable, meaning major problems have largely become a thing of the past. "I can't remember the last time we all spent swarming a fire," says Alex Mendoza. As for those rare occasions when problems do occur, improved data-validation practices— implemented in Qubole—make it easier to identify the root cause of the issue and to resolve it quickly.

> " We've reduced the amount of time our engineers spend troubleshooting by at least a factor of three."
>
> **Patrick Lawlor**
> Product Data Analyst, Gaia

In addition to freeing engineers from the frustrating task of troubleshooting, Qubole relieves them of the burden of maintaining, patching, and upgrading a dedicated in-house infrastructure, and automates other administrative tasks. "Qubole takes care of that background heavy lifting, says senior data engineer Jami Amore. "so, we can focus more on providing value to the business."

Looking Ahead

Gaia is presently considering new ways to use Qubole. Analysts like Lawlor and Koblitz are particularly intrigued by the prospect of using Qubole notebooks and dashboards to enable company stakeholders to generate their own ad hoc queries to surface business insights. This practice would reduce the workload on company analysts while also allowing for more granular reporting. On the data science side, the team hopes to automatically divert even more types of data into the data lake—for example, data from email campaigns (which is currently harvested manually)—and to generate more real-time insights.

> " *It becomes more powerful as we add different types of data into our data lake, because we can combine new data with our existing data to help drive more insightful business decisions.* "
>
> **Patty Vonick**
> Senior Data Engineer, Gaia

Business Value

- Improved recommendation engine drives 50 percent lift in subscriber engagement, content conversion, and consumption.

- Greater confidence in daily business insights enables true data-driven decision-making.

- Workload-aware autoscaling of compute power ensures workloads complete successfully, providing users access with critical information and resources with significant cost-savings.

- No outages, minimal troubleshooting, and near-zero administration frees engineers to focus on generating value for the business.

Qubole is the open data lake company that provides a simple and secure data lake platform for machine learning, streaming, and ad-hoc analytics. No other platform provides the openness and data workload flexibility of Qubole while radically accelerating data lake adoption, reducing time to value, and lowering cloud data lake costs by 50 percent. Qubole's Platform provides end-to-end data lake services such as cloud infrastructure management, data management, continuous data engineering, analytics, and machine learning with near-zero administration. Qubole is trusted by leading brands such as Expedia, Disney, Oracle, Gannett and Adobe to spur innovation and to transform their businesses for the era of big data. For more information, visit us online.

For more information:

Contact:
sales@qubole.com

Try QDS for Free:
https://www.qubole.com/products/pricing/

469 El Camino Real, Suite 205
Santa Clara, CA 95050
(855) 423-6674 | info@qubole.com

WWW.QUBOLE.COM