



Federated access to data and business critical intelligence to its workforce using QDS

About Wikia

Wikia is the ultimate hub for passionate pop culture fans around the world. As a digital media company with over 190 million global monthly visitors and over 360,000 fan communities, Wikia enables fans to come together and celebrate their passion for TV, movies, games and every pop culture topic imaginable. The company was founded in 2006 and is observing tremendous growth in not only the number of users and engagement, but also the amount of unstructured user data that can be used to improve the core product and user experiences.

Challenges

The top challenge Wikia faced in its big data initiative was rapid scalability. Wikia has an extremely active user base that is growing exponentially and producing massive amounts of data. By analyzing it, the business has the ability to target users in specific communities in ways that would garner the maximum response and impact.

Previously, Wikia was using an on-premises Hadoop deployment based on Cloudera. However, as it scaled, Wikia realized it was going to have to grow its hardware footprint significantly and was unwilling to continue to make investments in a short term solution, and most of the company did not have easy access to the data. Data engineers are extremely hard to find and having them spend time on systems administration tasks was distracting them from their primary responsibilities, and without federated access to data to get both technical and non-technical users and analysts direct access to the platform, Wikia couldn't deliver the insights needed to grow its business.

Rather than spending time and even more money further building out its rigid infrastructure, Wikia instead chose to invest in growing its team of data analysts and engineers while finding a cloud alternative to run its Hadoop clusters. Wikia was already using S3 to backup data for disaster recovery, and it wanted to add an analytics platform to take advantage of the data already hosted in S3.

“ *As we made the transition to the cloud, Qubole's ability to automate the infrastructure and automatically scale to meet the demands of our users saved us time, resources and budget.* ”



WADE WARREN
SVP – Engineering & Operations, Wikia

Results

Wikia selected Qubole as its analytics platform and migrated its on-premises clusters to Amazon Web Service (AWS). Wikia was able to fully migrate its big data infrastructure and workloads in a few months and has completely eliminated the overhead needed to manage its data platform.

Due to advanced automation and optimizations possible with the Qubole Data Service (QDS), Wikia was able to provide federated access to data and business critical intelligence across a large percentage of its workforce. Thanks to on-demand scaling leveraging AWS cloud resources and Qubole's capability to auto-scale up or down dynamically in order to maintain the right cluster size to handle the workload, Wikia is able to spin up and grow capacity much easier than the on-premises tools it had used previously.

Moving to a software-as-a-service offering eliminated worries about versioning and ongoing maintenance. Wikia is able to receive updates about its analytics instantly and get detailed, in-depth support when needed. Analytics efficiency also improved dramatically because Wikia is able to co-locate its Hadoop processing and reporting with data storage on AWS using the many cloud optimizations available with QDS.

Thanks to Qubole's ability to work across a shared Hive metastore, Wikia has migrated its workflows to the cloud while maintaining the uptime they need. Wikia is also able to easily and quickly test new data engines as they constantly seek to maintain its competitive advantage as a data-driven organization.

About Qubole

Qubole is passionate about making data-driven insights easily accessible to anyone. Qubole customers currently process nearly an exabyte of data every month, making us the leading cloud-agnostic big-data-as-a-service provider. Customers have chosen Qubole because we created the industry's first autonomous data platform. This cloud-based data platform self-manages, self-optimizes and learns to improve automatically and as a result delivers unbeatable agility, flexibility, and TCO. Qubole customers focus on their data, not their data platform.