



Industry: Media

Rebuilding and Optimizing Real-Time Mobile Advertising Bidding with Qubole

Business Problem Overview

In 2017, market forces called for a shift in Spotad's business model, from a managed service to a platform as a service (PaaS) that could operate with near-real time big data and support Spotad's advanced AI and ML models. This required significant changes to the company's business and technology infrastructure and processes.

"Our prior technology infrastructure—which relied heavily on Amazon tools, like Elastic MapReduce (EMR) and Athena—was insufficient in several respects," recalls Spotad CTO Tal Maizels. Specifically, it was not adequately robust to handle the volume of data Spotad would need to process. And, the total cost of ownership was too high—not to mention impossible to control.

Spotad didn't just need a more robust and less costly technology infrastructure. To orchestrate its big data operations, the company also needed a reliable and experienced technology partner. "We needed someone we could actually talk to, brainstorm with, and get good ideas from—good people who actually understand big data," says Maizels. Thanks to its reputation in the ad-tech space, "Qubole stood out." Since then, the partnership has prospered. "I think it was a good choice," says Maizels.

Improving Data Processing

Spotad's RTB platform powered by Qubole employs a simple yet sophisticated workflow. First, the platform—which collects streaming data in various forms from a variety of sources—receives a bid request for a mobile ad impression. The platform then issues a series of queries to evaluate whether the mobile ad impression on auction matches a particular client's advertising impression objective, timing, cost, and ROI parameters. If so, the platform uses an ML algorithm to calculate how much it should bid for the impression and places the bid automatically.

Spotad handles billions of these bids each day, which translates to processing between 3 and 4 terabytes of streaming data daily. Spotad also continuously collects discovery data to inform the ML algorithm used to calculate bids, and to identify clusters of customers who are likely to respond to a particular ad. Finally, it records events associated with any ads it successfully places—impressions, clicks, downloads, and so on. This and other operational information is delivered via reports and dashboards for internal users and clients.

About Spotad

Spotad, Inc. is a demand-side mobile advertising platform that uses real-time data science, including artificial intelligence (AI) and machine learning (ML), to automatically optimize its clients' digital advertising efforts on the world's biggest mobile ad exchanges. A mobile ad exchange is essentially a marketplace for ad impressions. Companies that distribute ads buy these impressions, usually by way of a bidding process. Spotad employs proprietary algorithms to identify which impressions will yield the most value for its clients and bids on those impressions automatically. This process is called real-time bidding (RTB).

Processing this volume of data naturally requires considerable resources—and the ability to apply those resources flexibly. Qubole allows this, swiftly autoscaling up to meet increased demand and downscaling just as quickly when the peak period ends or infrastructure becomes idle. “If you need more compute power, Qubole just throws more nodes at the cluster,” says Maizels. “This is a dream.”

Lower Cloud Costs

Before partnering with Qubole, Spotad’s RTB platform employed a PostgreSQL data warehouse, Amazon EMR, and Amazon Athena—an expensive setup. “The total cost of ownership was very high,” says Maizels, citing the cost of EMR as well as Athena’s opaque pricing model. “We felt abandoned...we lacked support and didn’t have a way to control costs.” Migrating to Qubole saved Spotad more than 50 percent in its operating costs almost instantly.

All-in-One Technology

Although Spotad’s RTB platform’s workflow is simple in concept, it involves several different technologies. These include Apache Kafka, a data lake on Amazon S3, Apache Spark, Hive, Presto, Looker, and more.

Qubole provides a single platform for Spotad to leverage all these technologies. It has simplified Spotad’s efforts to retool its business model from a managed service to a PaaS. It has also made it possible to easily administer the platform with very few people—which in turn frees up time for Spotad to devise and iterate new product features.

Speaking of administration, Qubole facilitates partition management, enabling Spotad to create partitions of streaming data every minute, and provides a very robust set of schemas to prevent mistakes. This allows for quick and easy recovery in the event of a glitch. “When mistakes happen—and with big data, it’s not if, but when—we can just roll back to the last partition we saved,” says Maizels. “It takes a matter of minutes.”

“*The amazing thing is that we do a lot more with Qubole, and for less money. We’re very happy with that!*”

Tal Maizels

CTO, Spotad, Inc.

But the benefits of a single platform go way beyond administration. It’s about serving all types of users in the most efficient way. Spotad currently organizes its use cases for Qubole along three areas. The first of these is data infrastructure and data operations—tasks such as scheduling data loads, data transformations and de-duping, data checks, verification and validation of data, among others.

The second area is ad hoc analytics and business intelligence. With an always-on Presto cluster on Qubole, Spotad ensures that reporting is taken care of for all internal users and clients. Reports and dashboards are distributed through the integration of Looker with Qubole, which leverages the data model and formulas across the board. This allows Spotad to serve its customers much more quickly—particularly new ones. For example, Spotad’s China operation alone can process between 600 and 800 queries per second. According to Maizels, “it’s a very good choice for keeping the storage (data lake) and compute separate, plus it’s not much work to build reporting cubes.”

Finally, the third area is data science. Using Spark and R, Spotad engineers process and validate raw auction data in Qubole to feed its AI and ML models. The models act on events to predict best outcomes and calculate bid prices with sub-second response times. The flexibility offered through Qubole Notebooks enables Spotad’s teams to process data and algorithms and collaborate much more efficiently through a unified interface. “It’s a dream in terms of DevOps,” says Maizels.

“Our partnership with Qubole is the kind of partnership that we want and that we value.”

Tal Maizels
CTO, Spotad, Inc.

Looking Ahead

Although revamping its business model and technology has been a significant job, Spotad’s partnership with Qubole both eased and expedited the process through ease of use, administration, and automation. This freed resources and people to build other high-value offerings, such as an easy-to-use, interoperable API to help internal users and clients build their own prediction models and perform other tasks to surface critical insights.

The company also seeks to shift its focus from predictive marketing to prescriptive marketing—in other words, to convert data into intelligence to recommend actions and strategies rather than simply interpreting data to predict marketing outcomes. This will enable clients to “focus on being creative and not being a slave to grids and grids of data,” says Maizels. “As prescriptive marketing requires even more processing power than predictive marketing, support from Qubole will no doubt be critical in this endeavor.”

Business Value

- Easy retooling of business model and technology infrastructure from a managed service to a platform as a service (PaaS)
- The ability to handle billions of bids daily—resulting in processing between 3 and 4 terabytes of streaming data per day
- A single platform for all use cases: data infrastructure and data operations, analytics and business intelligence, and data science
- 50 percent cost savings over the previous technology structure, which used a PostgreSQL data warehouse, Amazon EMR, and Amazon Athena
- Out-of-the-box integration with other key technologies, like Kafka, S3, Spark, Hive, Presto, Looker, and more
- Easy administration and automation to free up time for developing new product offerings

About Qubole

Qubole is revolutionizing the way companies activate their data — the process of putting data into active use across their organizations. With Qubole’s cloud-native big data platform, companies activate petabytes of data exponentially faster, for everyone and any use case, while continuously lowering costs. Qubole overcomes the challenges of expanding users, use cases, and variety and volume of data while constrained by limited budgets and a global shortage of big data skills. Qubole offers the only platform that delivers freedom of choice, eliminating legacy lock in — use any engine, any tool, and any cloud to match your company’s needs. Qubole investors include CRV, Harmony Partners, IVP, Lightspeed Venture Partners, Norwest Venture Partners, and Singtel Innov8. For more information visit www.qubole.com.

FOR MORE INFORMATION

Contact:
sales@qubole.com

Try Qubole 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