Komli Media Improves Utilization with Premium Big Data Platform Qubole

Challenges

Komli Media's success in the digital advertising business depends on recognizing and reaching targeted audiences efficiently and at scale. The company uses Big Data capabilities built on Hadoop to understand ad campaign performance. Komli has amassed) 100+ TB of data in Hadoop, arming its business users with information such as ads served, clicks, reach, impressions, data events, and consumer behaviors. Users rely on insights derived from this data to optimize campaigns, improve real-time bidding algorithms and to identify new opportunities. Komli encountered a number of issues with its initial implementation of Hadoop centering around the ad-hoc nature of its Big Data processing requirements explains Shailesh Garg, Engineering Manager at Komli Media. "Suddenly, there's huge demand from our users. I would just have to say no because I knew that we didn't have access to the computational resources we would need to process the requests. We were too vulnerable to the nature of loads on our platform and didn't have a lot of flexibility." Another challenge was the company's sheer volume of data. Komli collected around 700 GB of raw data each day or about 21 terabytes each month. Fixed clusters made it impossible for Komli to handle requests to process a month's worth of data with an acceptable turnaround time. Even smaller queries could take as long as 15 hours to process. Because of the long wait, users couldn't rely on getting the data they needed to get their jobs done. Ad-hoc processing of Big Data was also costing Komli too much money. Monthly cluster processing costs had escalated to $15,000 per month. With no ability to add or remove compute resources based on actual usage, the company had to over provision hardware to support its variable workloads.

Summary

- Digital Advertising Platform running on Hadoop

- 30% decrease in job processing times (waiting 1-2 hours for data instead of 10-15).

- 50% decrease in monthly processing costs.

- Ability to scale to meet all user requests.

- **Application:** ad-hoc queries, cluster management and auto-scaling for a campaign analytics

- **Users:** 15 business analysts and analytics engineers

> " *Using Qubole Data Service, we have reduced our monthly cluster processing costs by 50%. And, we're getting more for our money. Advanced features such as auto-scaling and S3 I/O optimization provide more flexibility and faster turnaround time to support the needs of our business users.*

SHAILESH GARG
Head of Analytics
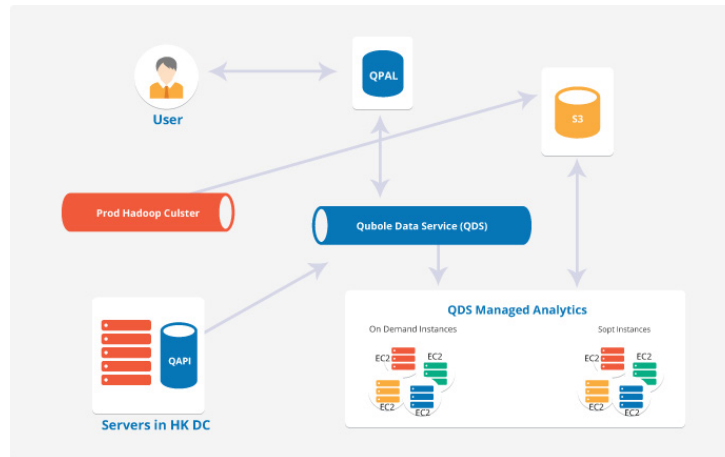
Komli realized that it needed to switch to a managed cloud service to support its ad-hoc Big Data processing. It looked at options such as Amazon's EMR, but decided that Qubole Data Service (QDS) offered some unique features that would best meet its needs. "The overwhelming reason we selected Qubole was because it is the only vendor that offers what I consider to be true auto-scaling," comments Shailesh Garg. "By true auto-scaling, I mean that auto-scaling is self-service – if the load on our cluster is high – the cluster automatically expands. Conversely, nodes are automatically removed when the load is low. This is different from manual auto-scaling where you need to pre-define auto-scaling capacity." Performance was also another area where QDS offered a higher value proposition than other options. Running Hadoop clusters on AWS would have compromised transfer speeds between S3 and Komli's Hadoop cluster on AWS using EMR. Therefore, AWS jobs would have run more slowly than running them with QDS which has extensive I/O optimization for S3. This was important since Komli needed faster query execution.

Results

QDS delivered the improvements in Big Data processing and total cost of ownership that Komli was looking for, faster performance and unlimited scale at a lower cost:

1. 30 percent reduction in job processing times
2. Scale to process all business requests using QDS auto-scaling
3. 50 percent savings on monthly cluster processing costs

"Our business has a lot more flexibility now thanks to QDS," says Shailesh Garg. "Nowadays users get their data processed in one or two hours max instead of the 10 to 15 hours it took before QDS. I'm no longer afraid when my business users ask for more data. In fact, I'm happy that they are able to depend on the availability of Big Data to make better decisions and I encourage them to experiment with even more business questions." Komli also reports that Qubole offers outstanding technical support. It is pleased with the support team's responsiveness and Big Data expertise throughout its migration to QDS.

The Future

Following its initial success with QDS, Komli is in the process of building a Big Data wish list including all the Big Data needs that it has not been able to previously support. It's also looking at potentially adding Apache Hive data warehouse software to make it easier for our business analyst to query our data

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.