Insightera Accelerates Hadoop's Time to Value Using Qubole Premium Service

About Insightera

Insightera is the first learning B2B targeting and personalization platform. Insightera helps marketers capitalize on their existing assets by personalizing onsite, social and ad network experience based on real-time discovery of prospects' industry, organization, location and digital journey. By using machine-learning algorithms, Insightera continuously improves ROI, auto-tunes campaigns and adjusts content accordingly for both known and anonymous prospects. Insightera's software works with any CMS, using any content and requires zero IT. Insightera was founded in 2009 and is headquartered out of San Mateo, California.

Customer Profile

- **Industry:** B2B in-bound marketing
- **Application:** cloud integration, ad-hoc queries, cluster management and auto-scaling for a B2B targeting and personalization platform built on Hadoop
- **Data Sources:** behavioral data stored in MongoDB and MySQL and reference data from social media and IP address lookups
- **Data:** Over 3 terabytes
- **Qubole Users:** 2 developers and 1 data scientist

> "
> *We were growing very fast as a startup and needed a way to accelerate our time to value for Hadoop," explains Mickey Alon, Insightera's CEO and Co-founder. "We wanted to focus more on data processing and turning insights into actionable results, and less on the operational side of Hadoop and Amazon S3 for tackling our Big Data integration challenges.*

MICKEY ALON
Insightera's CEO and Co-founder

Challenges

Insightera's cutting edge, inbound B2B marketing platform leverages Big Data, machine learning and predictive analytics to display the most relevant content to targeted high-yield prospects in real-time. Insightera selected Hadoop as its marketing platform's Big Data foundation, but found it to be complex and time consuming from an operational perspective.

"We were growing very fast as a startup and needed a way to accelerate our time to value for Hadoop," explains Mickey Alon, Insightera's CEO and Co-founder. "We wanted to focus more on data processing and turning insights into actionable results, and less on the operational side of Hadoop and Amazon S3 for tackling our Big Data integration challenges."

Insightera's forte is turning insights into action using its proprietary real-time CEP engine powered by predictive analytics. Its data scientists wanted to find a way to simplify Big Data operations so they could test out and release key predictive algorithms that later on turned into one of Insightera's core capabilities. Rather than hiring additional Hadoop experts who can be expensive and hard to find, Insightera decided that it would be more advantageous to use a Big Data as a Service solution.
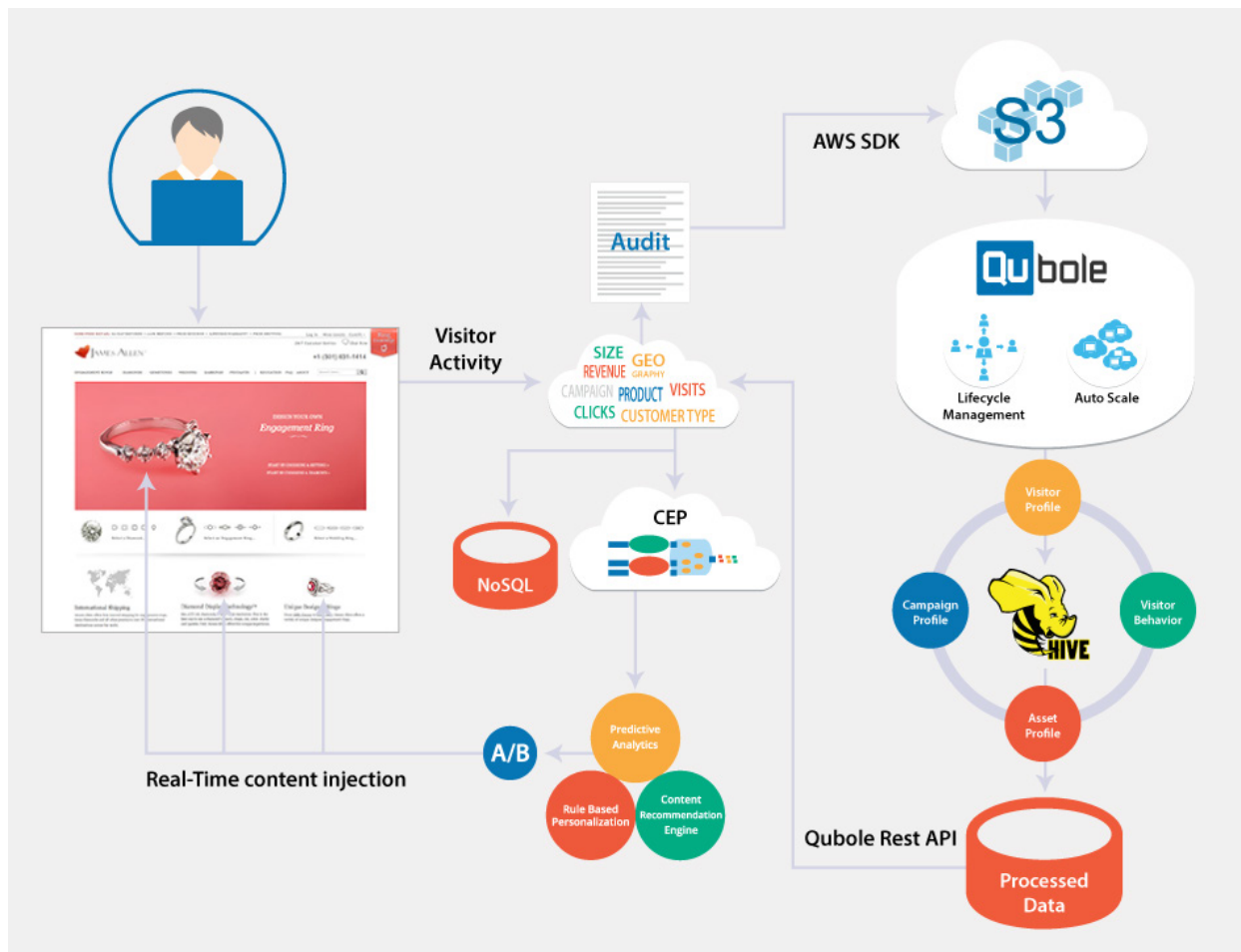
Why Qubole Data Service?

When evaluating Big Data as a Service solutions, Insightera did its homework. According to Alon, "We looked at similar offerings, but quickly discovered that Qubole offered the most mature solution. The final decision boiled down to choosing between Qubole Data Service (QDS) and Amazon EMR. We selected QDS because it provided a better user experience, auto-scaling, more flexible cloud integration and a wider selection of training resources."

With regards to the user experience, Insightera found that QDS provided the most intuitive tools and the highest degree of automation. Its query editor and visual query builder offered Insightera's developers and data scientists an easy way to access Hadoop data with no specialized MapReduce and Pig coding skills. Because QDS runs on an elastic Hadoop-based cluster, Insightera could automate cluster configuration and management so that its small Big Data team wouldn't have to spend their time on these activities.

Insightera valued QDS resource utilization features, including Amazon S3 I/O optimization, faster queries and self-service auto-scaling to scale capacity up and down as needed without having to manually reconfigure resources. Insightera viewed auto-scaling as essential not only to meet unanticipated demand from big brands in its customer portfolio, but also saw auto-scaling as a way to save on cloud compute costs.

In addition, QDS received high marks from Insightera for its cloud integration. QDS provided connectors to its MongoDB and MySQL sources for Insightera's behavioral data, offered a built-in Hive User Interface, and fast I/O performance for Amazon S3. Plus, QDS made it easy to marry behavioral data with reference data such as social media activity and IP address lookups.

Results

Insightera was able to get QDS up and running in just a few weeks. As Insightera's business grew 200 percent in just one year, its data volumes grew at the same rate, reaching over 3 terabytes. With the help of QDS, the company avoided time-consuming, complex and expensive administration typically associated with this massive amount of Big Data, realizing several benefits:

1. Reduction of time to provision Hadoop cluster instances from days or weeks to just a few minutes or hours with automated cluster provisioning and management
2. Precluding the need to hire additional operational staff. Insightera estimates that it would have had to increase its development team by 40% if it had not implemented QDS.
3. Saving 20% in cloud compute costs by eliminating unnecessary uptime using QDS auto-scaling
4. More efficient resource utilization due to higher performance, improved resource allocation and auto-scaling of capacity

Most importantly, QDS empowered Insightera's team with a Big Data platform that requires zero-IT and that resulted in faster time-to-market. QDS made configuring and managing Hadoop clusters, adding data sources, and running queries very simple. Alon comments, "My team learned how to perform data loading and run queries with QDS in just a few minutes. They became masters in a couple of days."

The Future

Insightera is currently focusing on increasing capacity to support its ever growing volumes of Big Data. To accomplish this, Insightera will be taking greater advantage of the fully automated clustering capabilities offered by QDS. This includes using its scheduler and workflows for data loading, storage and updates as well as leveraging its cluster API.

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.