

## REPORT REPRINT

# Treasure Data tightens its managed analytics message, sees customer and revenue growth

JAMES CURTIS

11 FEB 2016

Although Treasure Data started out offering Hadoop as a service, it has shifted its positioning to focus more specifically on a managed analytics infrastructure that runs in the cloud - what the company calls its 'plug and play' offering.

---

THIS REPORT, LICENSED EXCLUSIVELY TO TREASURE DATA, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR REPOSTED, IN WHOLE OR IN PART, BY THE RECIPIENT, WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2016 451 Research, LLC | [WWW.451RESEARCH.COM](http://WWW.451RESEARCH.COM)

Although Treasure Data started out offering Hadoop as a service, it has been diligently working not so much on getting customers up and running on Hadoop, but on helping them collect, manage and access data for analytics. Since our previous coverage, Treasure Data reports healthy growth in new customers and a sharp bump in revenue. A previous funding grab certainly helps as well.

---

## THE 451 TAKE

It has been more than a year since we last spoke with Treasure Data. The company may not have as high a profile as similar firms in the market, but we believe it is making improvements in this area. Treasure Data possesses strong customer loyalty, having retained its earliest customers and continued to provide value-added service. As Treasure Data continues to raise its profile and grow its customer base, it could shape up to be a strong player in the hyper-competitive big-data services space.

---

## CONTEXT

Treasure Data was founded in 2011 by Hiro Yoshikawa (CEO), Kazuki Ota (CTO) and Sada Furuhashi (software architect). The company is headquartered in Mountain View, California, but it also has a strong presence in Asia, with offices in Japan and Korea.

Since our previous coverage of Treasure Data, the company has moved away from positioning itself as a Hadoop-as-a-service provider to focus more specifically on a managed analytics infrastructure that runs in the cloud – what the company calls its ‘plug and play’ offering. The strategy appears to be working for the startup. Treasure Data reports 150 paying customers, up from 110. The company also says revenue rose by 144%, driven primarily by an increase in the average contract value. Treasure Data has grabbed \$29m in total funding. Its most recent funding round occurred in January 2015, a series B round led by Scale Venture Partners that came in at \$15m. Treasure Data previously had 55 employees but now cites more than 90 on its payroll.

## PRODUCTS

Treasure Data provides an end-to-end offering for collecting and storing data, and then making that data available – from the cloud – to data analysts/scientists from a Web console or through BI and visualization products. Early on, Treasure Data concentrated on capturing log data files but later learned that companies not only wanted the log files, but they also wanted to combine that data with other data – application data, for instance, or even product usage data. While enterprises can build environments to collect, store and integrate the data using a combination of relational database management systems, NoSQL databases, Hadoop, CRM systems and enterprise data warehouses, Treasure Data says that this process takes a good deal of time and dedicated resources to pull off.

As such, the company’s value proposition is to remove the complexity of collecting, storing and making the data available for analysis. Treasure Data says it can collect from more than 300 sources. It has put considerable effort into its tools for ingesting data. It has developed two tools – fluentd for data streaming and embulk for batch ingestion – which the company has open sourced and made available as Apache 2.0 licenses. For instance, the company uses a server agent based on fluentd to stream data to Treasure Data. Additionally, Treasure Data uses SDKs for mobile devices and Web applications, for instance, to collect data from these environments.

Treasure Data stores the data and makes it available as part of its multi-tenant cloud service. Although it stops short of calling its data repository a data warehouse, it operates in a similar vein. Underneath, Treasure Data runs its database that stores the data in a columnar-based, schema-flexible format. To access and work with the data, Treasure Data provides a proprietary Web console that enables users to carry out analysis of the data by submitting SQL queries directly into the tool’s UI, which can also be saved as workflows. However, queries can also be executed using a command line interface.

Users have a number of options for exporting the results. Data can be exported to popular cloud platforms such as Amazon Web Services S3 and Google Cloud Storage; other databases such as MySQL, Postgres, Amazon Redshift and NoSQL stores such as MongoDB; cloud applications such as Salesforce and MailChimp; and BI tools such as Tableau Server.

## USE CASES

Treasure Data claims some flexibility in how the offering gets deployed. It has had some success with gaming companies because these were some of the company's first customers and have remained loyal. One gaming customer had pertinent data sitting in separate data silos, which made it difficult for the data science team to access all of the data. Treasure Data's offering enabled the customer to aggregate its data sources plus collect data from the gaming server and mobile devices in which the games were played. The customer then used Tableau for BI analytics and visualization. Collecting, storing and making data available is essentially Treasure Data's value play. The company says ad tech is another use case that also integrated with Amazon Redshift. It cites e-commerce and IoT, such as for telematics, wearables and energy, as other use cases.

## COMPETITION

Treasure Data offers a managed service, an analytics infrastructure offering in the cloud that the company refers to as plug and play, suggesting an easier way to collect, store and analyze data. Although we are unaware of direct competitors, there are a number of vendors that offer data analysis as a service. Altiscale, Qubole and Cazena come to mind. Altiscale offers managed services for both Hadoop as a service and Spark as a service. Qubole offers 'data as a service' through its Qubole Data Service offering that is designed to give organizations a choice of computing engines based on data stored in the cloud. Cazena, on the other hand, offers a cloud-based computing service for Hadoop (Cloudera or Hortonworks), massively parallel processing (Amazon Redshift) and Spark for in-memory processing.

The Hadoop cloud providers might also be considered competitors, but not directly. Amazon Web Services has Amazon EMR and Microsoft offers HDInsights on its Azure cloud platform, while Google introduced Google Cloud DataProc in September 2015. IBM offers BigInsights on Cloud on its SoftLayer cloud platform and Bluemix PaaS. While these vendors have cloud offerings, the difference is that Treasure Data is a managed service, and although Hadoop in the cloud negates some on-premises administration, it still requires some technical expertise to manage. Moreover, there are the known Hadoop vendors, notably Cloudera, Hortonworks, MapR, Pivotal and IBM. These vendors offer their wares in the cloud as well as on-premises, including offering services associated with their respective deployments. Treasure Data is cloud-based, but the company also considers Hadoop data complementary in that it's another data source to feed its cloud data store.

There are also a few Hadoop managed services firms such as Teradata's Think Big and MetaScale. Although these firms do provide a managed service, the focus tends to be on Hadoop with customized development work also offered.

## SWOT ANALYSIS

### STRENGTHS

Treasure Data has zeroed in on providing a specific data capture, storage and analysis use case, buoyed by the company's strong data-collection and querying tools.

### WEAKNESSES

Treasure Data's profile is still on the light side, although it is growing with renewed marketing efforts and expanding its use cases.

### OPPORTUNITIES

The company has opportunities to tap into firms that do not want to mess with managing an on-premises environment yet have data-analysis needs that can be addressed with Treasure Data's data capture and collection capabilities.

### THREATS

Treasure Data tends to get lumped in with the Hadoop vendors, even though the company's offering targets a broader data-analysis play.