

REPORT REPRINT

Primary Data amps up efficiency and accessibility with enhanced data virtualization

HENRY BALTAZAR, LIAM ROGERS

18 AUG 2017

The data management vendor has raised additional funding on the heels of the version 2 release of its DataSphere data virtualization and management tool. Will Primary Data capitalize on the growing need for data and workload optimization tools?

THIS REPORT, LICENSED TO PRIMARY DATA, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT 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 RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2017 451 Research, LLC | WWW.451RESEARCH.COM

Data management vendor Primary Data recently secured a new round of funding while also releasing an early-access version of its DataSphere 2.0 data virtualization platform. DataSphere aims to use automation to appropriately position data on available storage to facilitate more efficient migration and break down silos between primary and secondary storage.

THE 451 TAKE

The increasing relevancy of data management is in parallel with the ongoing growth of the sheer volume of data that enterprises must deal with. Data management is also being fueled by advances such as flash and object storage that can enhance performance and cost-effectiveness. While offering a panacea for the plight of sprawling data is no easy task, effective data management provides efficiency and mobility at scale, and this is the goal for DataSphere. Data virtualization and management are tools to optimize existing storage to keep pace with the woes of data growth, and Primary Data is one of several startups addressing this sector. However, the company is leveraging machine learning in its latest DataSphere release to further draw on the utility of metadata, and the additional funding provides Primary Data with the means to boost its sales efforts. Although earlier attempts at information lifecycle management failed to gain traction over a decade ago, evolving market needs could make automated and intelligent data movement and placement a key requirement for clients in the future.

CONTEXT

Primary Data was founded in 2013 and came out of stealth in 2014. The company is led by several previous executives from flash-based storage proponent Fusion-io, which was acquired by SanDisk in 2014. Current CMO Rick White and CTO David Flynn were both founders of Fusion-io, and it even includes Apple cofounder Steve Wozniak among its ranks as chief scientist. Primary Data is based in Los Altos, California, and has a headcount of about 80.

The company just closed its most recent round of funding valued at \$20m led by Pelion Venture Partners, with participation from Accel Partners and Battery Ventures, bringing its current VC haul to \$80m. Primary Data also secured an additional \$20m in a line of credit. Its customer count is currently less than a dozen, with average deal size in the six-figure range. The vendor claims the largest DataSphere installation to have over 1.67 billion files under management, and is confident in the scalability of its architecture in the enterprise and service-provider markets. With its latest release, Primary Data is also looking to address smaller enterprise environments with more modest budgets.

STRATEGY

The crux of Primary Data's DataSphere strategy is focused on intelligently delivering the right data to the right place at the right time, which is easier said than done given the complications of data growth and siloed storage environments. The company is somewhat broadening this vision by offering DataSphere in two versions to pursue smaller clients and expand its potential customer base. Some of Primary Data's target verticals include media and entertainment, oil and gas, and service providers.

The company claims that one function of the additional funding will be to strengthen sales efforts and facilitate the formation of additional partnerships. As we previously reported, Primary Data has maintained a few partnerships with the likes of VMware, NetApp and Trace3. In conjunction with the DataSphere 2.0 release, the vendor also announced partnerships with Amazon Web Services, Cloudian, Google and Scality. It also offers DataSphere via resellers.

PRODUCTS

Primary Data's products include DataSphere 2.0 and DataSphere Extended Services (DSX). DataSphere 2.0 is the company's flagship metadata engine, an out-of-band storage-agnostic tool for data management and migration that automates data flow. It is now available in two versions, enterprise and line of business (LOB). The newer LOB version is for smaller deployments that may not need to operate on the same scale as many enterprises and inte-

grates up to four storage nodes. DSX is a data orchestration tool that fills four roles: cloud connector, data mover, data portal and data store. It moves data based on criteria set in DataSphere, also allowing data to be moved to S3-based cloud and object storage.

DataSphere 2.0 introduces several new features, with an emphasis on making the data management experience more intelligent. Snapshots have been enhanced so that they can be de-duplicated, compressed and moved to and stored in the cloud. DataSphere 2.0 also adds supports Active Directory, Windows ACL support, SMB 2.1 and 3.1, as well as ID mapping between Windows and Linux user domains. Network connectivity enhancements include VLAN, IPv6 and IPv4 support. Additionally, this version includes functionality based on objective expressions that can be used to provide finer control over data placement and can streamline tier creation through the deployment of predefined objectives.

Pricing is \$35,000 for a single node for the LOB version for one year, while Primary Data's enterprise software typically has an average deal size in the six-figure range. The vendor claims that for some customers the price of its software is dwarfed by data migration costs, which can sometimes exceed \$100,000 per operation.

COMPETITION

Primary Data competes with incumbent storage providers that offer data management and optimization products. These include Dell-EMC (Enterprise Copy Data Management), IBM (Spectrum), NetApp (OnCommand Insight) and HDS (Data Management). These vendors' offerings target a range of uses, from data retention and efficiency to providing deeper management and mobility capabilities.

Other storage firms playing in the data management space include Actifio, Catalogic, Cohesity, Datagres, Delphix, Komprise, Rubrik, Talon Storage and Virtual Instruments. Actifio has been moving toward a data-as-a-service model, taking aim at making secondary storage and disaster recovery more efficient. Cohesity and Rubrik have both landed additional funding in the past year and are leveraging hyperconverged infrastructure alongside data management platforms. Komprise similarly focuses on unstructured data analysis and management with an emphasis on turning data visibility into useful insights. While Primary Data encounters Komprise in some deals, the company claims that Komprise's offering is targeted more at the lower end of the market and notes that it focuses on NAS and file servers.

SWOT ANALYSIS

STRENGTHS

Primary Data is targeting a pain point for many enterprises as it looks to break down storage silos to make data easier to manage and in turn more useful. Integrating machine-learning functionality could prove to be a differentiator.

WEAKNESSES

The company's customer count is still relatively low, and it will have to ensure that it is marketing its product accordingly to both storage infrastructure professionals and generalists struggling with data management.

OPPORTUNITIES

Enterprises are aiming to get the most value out of their data, and to do so efficiently, and this presents market opportunities for vendors willing to cater to customers' ongoing data management challenges as they migrate to hybrid cloud environments.

THREATS

The number of data management providers in the market is increasing as more secondary storage players look to offer greater management capabilities.