

REPORT REPRINT

VAST Data aims to be fastest-growing storage startup of all time

MARCH 28 2019

By Henry Baltazar, Liam Rogers

Claims don't get bolder than what VAST has said out of the gate in its debut. The vendor believes its unique architecture and data-reduction technologies can make flash a viable and efficient replacement for multiple disk storage tiers.

THIS REPORT, LICENSED TO VAST 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.



Summary

VAST Data is strutting into the market with \$80m in funding, and it has already closed a handful of \$1m+ deals. The vendor's Universal Storage platform leverages NVMe over Fabrics, inexpensive quad-Level cell (QLC) flash and high-performance storage-class memory (3D XPoint), and the company claims it can replace many disk storage use cases with a single-tier-storage approach. VAST's global compression is a key element that allows its products to be price-competitive with disk-based object and scale-out NAS offerings that handle many unstructured data storage workloads today. VAST has found initial traction for its systems in the biosciences and financial services markets, and the vendor is targeting emerging use cases such as data storage for artificial intelligence and machine learning.

451 TAKE

Any storage startup entering the market screaming 'disk is dead' is effectively painting a giant bullseye on its back, but this is something the VAST Data team appears to relish. In a market where venture capital for storage has dwindled in recent years, the 'go big or go home' mentality is important not only to build awareness of its platform but to excite prospects and investors alike. The key to success for any storage startup is its ability to carve out a foothold in the highly contested storage space, and VAST's early wins are even more important than the impressive early-stage funding it has received. Naysayers will likely question VAST's architecture, which leverages QLC flash. QLC flash is inexpensive but has low endurance ratings that could create concerns about long-term reliability. Competition in the all-flash storage space also continues to ramp up, where products such as Pure Storage's FlashBlade are also targeted for object storage and scale-out NAS use cases.

Context

VAST Data was founded in 2016 by a team of three: CEO Renen Hallak, former XtremIO VP of R&D; Shachar Fienblit, VP of R&D, who previously served as Kaminario's CTO; and Jeff Denworth, VP products and marketing, who most recently led marketing for CTERA. Former Dell EMC and XtremIO veteran Michael Wing joined the company last fall as president, along with VP of operations Avery Pham, who most recently had the same position at Pure Storage. Prior to launching the company, the team interviewed hundreds of customers ranging from enterprises to cloud hyperscalers to get an understanding of how flash was transforming storage in their datacenters.

The company has raised \$80m in series A and B funding (two \$40m rounds) from Norwest Venture Partners, Goldman Sachs, Dell Technologies Capital, TPG Growth and 83North. It currently has a headcount of about 70 and has its headquarters in New York, with offices in San Francisco and Tel Aviv. In its first 90 days of sales, VAST has closed a handful of \$1m+ deals and claims it is on pace to becoming the fastest-growing infrastructure startup in its first year of operation.

Strategy

VAST is initially targeting four key vertical markets: life sciences, financial services, content creation and distribution, and high-performance computing. In the life sciences market, Ginkgo Bioworks is using VAST to facilitate gene sequencing operations. Zebra Medical Vision is using VAST's high-speed file storage in conjunction with NVIDIA GPUs and Zebra's AI algorithms to process CT and MRI scans to help radiologists identify high-risk patients earlier to reduce cost of care and facilitate disease-prevention programs. In the financial services space, VAST has a hedge fund customer that is using the platform to enable adaptive trading using machine learning to react to changing market conditions.

REPORT REPRINT

Government IT integrator General Dynamics Information Technology has deployed a multi-petabyte cluster of VAST's Universal Storage to handle a customer's large analytics workloads. VAST goes to market exclusively through its resellers and has announced a handful of partners including Enterprise Vision Technologies, FusionStorm, Link Solutions Group and Trace3.

VAST is targeting data-intensive applications in big data (Hadoop, Spark), log analytics (Splunk, Elastic) and AI/ML. The vendor also sees an opportunity in the enterprise backup space, where it claims its Global Compression can further reduce deduplicated backup payloads from vendors such as Veritas and Commvault.

Technology

Although VAST's Universal Storage leverages new technologies such as 3D XPoint and QLC flash, which are not common in the enterprise storage space today, VAST's main innovation is in the software stack, which takes advantage of the benefits of these solid-state technologies while working around media limitations.

QLC's write endurance is a major issue since this medium can only handle a few hundred write cycles, in contrast to triple-level cell (TLC) flash, which can endure 1,000 writes and multi-level cell (MLC) flash, which can handle 10,000 writes before becoming unable to accept new write requests. On the positive side, QLC is attractive because it costs half as much as TLC on a dollar-per-gigabyte basis and will be 1/10th the cost of MLC flash. We note that flash media endurance is not impacted by reads, which means once data is written, it can be read many times over without burning out the media. This quality, along with the high density of QLC flash, makes the medium attractive for unstructured data use cases where data is not modified often. In the Universal Storage architecture, 3D XPoint provides buffering to ensure full-stripe writes to eliminate flash wear from read-modify-write operations, which is essential to prolonging the endurance of inexpensive QLC media. To alleviate flash endurance concerns, VAST is providing a 10-year endurance warranty for its QLC flash, which is also a key point of differentiation from disk-based storage systems that are often replaced every three to five years.

VAST's disaggregated shared everything (DASE) architecture has two building block components: storage enclosures that provide persistence and loosely coupled server nodes that are stateless and are responsible for storage processing and other capabilities such as metadata management. In DASE, an NVMe fabric provides connectivity and allows any server processor to be able to see data on any storage node in the Universal Storage cluster. Universal Storage has a global namespace that allows the servers to access data directly from the storage nodes over the NVMe fabric without needing to worry about cluster cross-talk, while also providing scalability to support up to 1,000 storage enclosures and 10,000 nodes running VAST software.

Two additional software components that are unique are VAST's Global Compression and Global Data Protection. The Global Compression software is application-aware and looks for similarities in data by creating data fingerprints after it has been written to the storage-class-memory media. After comparing fingerprints, the software clusters similar chunks together and then extracts and stores byte-level deltas, which allow the system to present data on demand using reference blocks and the deltas. VAST's Global Data Protection is an erasure code algorithm that the vendor claims can reduce data protection overhead to as low as 2% in large clusters, in contrast to alternative algorithms that have 33-66% overhead.

VAST has server and storage appliances available at launch, although Universal Storage can be deployed as software only for large-scale customers with more than 100PB of storage. A third option is a mixed configuration where VAST storage enclosures work together with VAST containers running on commodity servers. For the appliances, the compute nodes are composed of quad server chassis, which have eight 50Gb Ethernet network ports, while the storage nodes contain 675TB of raw flash

REPORT REPRINT

and four 100Gb Ethernet/InfiniBand network interfaces. Universal Storage provides storage access via NFS or through an Amazon S3-compatible object storage interface, and a block storage interface could be added in the future.

Competition

VAST is attacking the storage status quo with its Universal Storage architecture. In the first iteration, VAST storage clusters will serve as an alternative to disk-based scale-out NAS and object storage systems. Dell EMC, with its Isilon scale-out NAS platform, continues to be a market leader in the space and has been selling all-flash versions for a few years now. There are several other contenders, including large players such as NetApp, Hitachi Vantara and Red Hat, and some smaller vendors such as DataDirect Networks, Qumulo, Nexenta, Panasas and WekaIO. Pure Storage's FlashBlade all-flash NAS and object storage platform is a likely competitor as well.

VAST currently does not have block storage capabilities, but the vendor believes the platform's ability to deliver NFS with sub-millisecond latency will allow it to handle performance use cases that typically use block storage. That being said, in this iteration, we do not see VAST's Universal Storage as a direct competitor to mainstream AFA products such as Dell EMC's PowerMax, Pure Storage's FlashArray, HPE's all-flash Nimble and 3PAR systems, and NetApp All-Flash FAS. At some point, VAST will add support for various virtualization and orchestration platforms to expand its target use cases.

SWOT Analysis

STRENGTHS

VAST Data has an experienced management team that's had success disrupting the storage market before. The company has a substantial amount of venture capital and has hit the ground running with some large deals out of the gate.

WEAKNESSES

The vendor's technology has not been on the market for long and still needs to establish a track record for reliability and scalability. The platform currently lacks features such as asynchronous replication.

OPPORTUNITIES

Organizations are looking to simplify their storage infrastructures while reducing their opex and capex, which VAST storage clusters could help alleviate with their scalable architecture. New use cases in AI, analytics and containers could also drive adoption for highly scalable storage platforms, which can deliver high performance.

THREATS

There are many strong and established players in the storage market that will not be easy to displace.