Datrium proves that storage architecture isn't its only unusual quality

TIM STAMMERS

25 APR 2017

The company's midmarket primary storage system ushered in a very unusual architecture in 2016, and it has been updated with new features that are also uncommon. Datrium is backing its claims of strong sales growth with a long list of named customers.

451 Research®

Startup Datrium entered the general-purpose primary storage market in 2016 with a product called the DVX, which features a distinctive scale-out architecture and is aimed at service providers and midsize and larger enterprises. Datrium says it has won more than 80 customers, which it claims is faster growth than even high-flying Nutanix or Data Domain saw in their first year of shipping.

The company has also renamed the DVX as the DVX Rackscale, although it has not yet felt the need to raise the maximum 32-node, 180TB size of the product. It has unveiled free updates to the software that underpins the system, adding data replication, snapshots and encryption, among other features. Building on the DVX's existing per-VM, policy-based management, the snapshots are managed with the help of a backup catalog that Datrium says scales to hundreds of thousands of objects. The encryption is 'blanket' encryption, which goes further than the at-rest encryption offered by other storage systems by also protecting data on the host and in flight. Within the next 12 months, Datrium plans to extend the replication and blanket encryption to link on-premises DVX systems to Amazon's public cloud.

THE 451 TAKE

Datrium's sales numbers for its first year of shipping are certainly encouraging, and its first round of major updates are promising, showing that it is not just the architecture of the company's storage system that is unusual. That is because the updates don't just tick off standard RFP boxes for enterprise storage features – they have also been implemented in an impressive way. The new and unusually scalable snapshot catalog continues Datrium's focus on making its product simpler to manage in virtualized environments than legacy storage systems. We note that the same focused approach has already been a major factor in the success of hyperconverged systems (but don't confuse Datrium's product with hyperconverged storage – it is not the same). Datrium's blanket encryption is equally unusual and impressive, and should appeal to multiple industry sectors.

BACKGROUND

Founded in 2012 and based in Silicon Valley, Datrium has raised $110m in three funding rounds. The most recent series C round was announced in December 2016, and raised $55m. It was led by New Enterprise Associates, with contributions from existing investor Lightspeed Venture Partners and an unnamed investor. Datrium's CEO is Brian Biles, who was founder and VP of product management at Data Domain – a startup that pioneered de-duping backup appliances and was bought in 2009 for $2.4bn by EMC, only eight years after it was founded.

Datrium began shipping its product just over a year ago. List prices currently begin at $118,000, and Datrium says it had sold over 100 DVX systems. According to Datrium, that is 50% and 80% more sales than Data Domain and Nutanix had achieved at the same point in their respective histories. Datrium says Nutanix's IPO filings indicated that it won 38 customers in its first year on the market, while Datrium itself won 70 customers in its first year. The company now has 80 customers (roughly 60 of which have been named) that span the finance, media and entertainment, hosting, healthcare, state and local government, education, technology, and manufacturing sectors.

The unusual DVX architecture is similar to – and at the same time very different from – both hyperconverged and conventional stand-alone storage. Datrium itself calls its product an 'open convergence' system. It consists of VMware ESXi host servers fitted with flash drives that store read-only cache copies of active or hot data, linked via 10Gb Ethernet to a back-end disk array that stores the master copy of all data and handles all write operations. One back-end disk array can be linked to up to 32 host servers, and provides 60-180TB of effective capacity, according to data-reduction ratios.

All storage services – such as snapshots, RAID, in-line de-duplication and background compression – are completed on the host servers by Datrium software. Until April, the only hardware sold by Datrium was the back-end disk array, which comes with unlimited free licenses for the host software.

Customers previously sourced their own host servers, as well as the flash drives that were fitted to them. They can continue to do that, but there is now an option to buy servers from Datrium that are 16- or 28-core x86 devices pre-installed with up to eight 480GB SATA flash drives, alongside VMware software and Datrium's storage software. Customers can mix their own servers with servers bought from Datrium.

BLANKET ENCRYPTION

Datrium's encryption is completed in software using an Intel instruction set, and has almost no impact on performance, according to the company. That is relatively unusual, since it distinguishes it from the hardware-based encryption used in self-encrypting drives (SEDs), which are currently the most common way to implement encryption.

The Datrium feature is much more unusual in two other ways – it protects data while it is on the host, in flight, and at rest, and it does so without disabling cost-slashing data de-duplication and compression functions. 451 Research is not aware of any other stand-alone, converged or hyperconverged storage systems that provide native blanket encryption, let alone with support for data reduction.

This level of security will be appealing in sectors such as federal and local government, financial services, and healthcare. Not only do these sectors typically operate under heavy regulations, they are also frequently the subject of cyber attacks. Given that de-dupe and compression typically reduce the effective cost of storage media four- or fivefold – including the server-side flash used by Datrium – the ability to continue using those two functions with blanket encryption prevents the latter from carrying a cost penalty.

Stand-alone or SAN-attached storage systems cannot encrypt data on the host or in flight because they have no footprint on the host by which to do this. Third-party host software such as VMware ESXi can deliver blanket encryption – on the host, in flight and at rest – but when the data reaches the stand-alone storage, it cannot be de-duped or compressed. Encryption disables those two functions by removing the duplication and bit patterns that they rely on.

Blanket encryption is possible with hyperconverged systems, but 451 Research is not aware of any major hyperconverged systems that do so. Instead, the major hyperconverged storage products encrypt data only at rest – in some cases encrypting data in software, and in others doing so in self-encrypting drives. In some cases, de-dupe and compression are not disabled, but in others they are disabled.

SNAPSHOT CATALOG

Until now, the DVX did not offer snaps or system-to-system data replication for disaster recovery. These are table-stakes functions, but it is very common for storage startups to begin shipping systems without them, and add them later. That is what Datrium has done, delivering those two new functions and others via a DVX software update. However, Datrium claims it is now delivering more than mere snapshots and replication.

Alongside the ability to schedule snapshot creation and retention on an individual or group basis – using predefined or custom policies for levels of data protection – Datrium has added a searchable snapshot or backup catalog. While such catalogs already exist in third-party backup software, Datrium says DVX is the only storage system that it is aware of that includes such a catalog, which can scale to hundreds of thousands of snapshots. This is needed because a DVX system may be a platform for thousands of VMs. Each of those VMs can comprise multiple virtual disks, and each of those virtual disks will generate an even larger number of snapshots, and will be accompanied by other objects, such as templates and ISO and OVA files.

Those VMs, snapshots and other objects will be stored in multiple LUNs or data volumes, and therefore need storing in a namespace with database-style wildcard search functions, according to Datrium. Without such a catalog, customers looking for a snapshot from which to restore data would first need to determine which LUN or volume the data originated in, before mounting and searching the snapshots taken from that LUN.

REPLICATION AND CLOUD SUPPORT

The replication added to the DVX is asynchronous and bidirectional, allowing 1:n and n:1 replication of snapshots, for RPOs down to 30 minutes. As for the other DVX data services, the replication is managed by policy for groups of objects.

By March 2018, Datrium plans to deliver the ability to replicate snapshots to Amazon's public cloud, with policy-based management of where snapshots will be kept, and with the existing snapshot catalog embracing objects stored both on-premises and in the public cloud. The data stored in the public cloud will be subject to the same blanket encryption as data stored on-premises in Datrium systems.

COMPETITION

The market for midrange primary storage is dominated by hybrid disk-and-flash stand-alone SAN-attached storage sold by incumbent suppliers such as Dell-EMC, NetApp, HPE and IBM. All of these suppliers are selling hybrid SAN-connected storage that was originally designed to be powered entirely by disk, and has been retrofitted with flash drives and policy-based tiering software. Per-VM management in these devices is also being retrofitted, via VMware's VVOLS framework, with varying results.

Datrium is not the first startup to see the opportunity to compete in this market with newer designs. Other startups, such as Nimble Storage (now owned by HPE), Tintri and Tegile, have developed stand-alone hybrid storage that was purpose-designed to marry disk with flash, sometimes with VM-level management also built in from scratch.

Although it is significantly different in some important aspects, Datrium's architecture bears a superficial resemblance to hyperconverged infrastructure (HCI). Vendors of HCI are chasing the same customers and storage budgets as Datrium, and include VMware, Nutanix, SimpliVity (now owned by HPE), Cisco, Pivot3 and HyperGrid (formerly GridStore). However, HCI sales still only account for less than 10% of the total enterprise storage market. Among the advantages that Datrium claims over HCI are better scaling, flexibility and capacity utilization. Another startup whose product shares Datrium's quality of being tough to categorize is Cloudistics, which describes its unusual architecture as 'super converged.'

SWOT ANALYSIS

STRENGTHS
Datrium has already presented strong arguments about the virtues of its unusual architecture, and is now building on those foundations with equally unusual and impressive functionality.

WEAKNESSES
At present, the DVX only supports VMware's ESXi hypervisor, although Datrium has said that it plans to support others.

OPPORTUNITIES
The rapid growth in the sales of hyperconverged storage have shown that there is a demand for an alternative to the conventional OEM systems that continue to dominate the market for mainstream primary storage. It remains to be seen how far upmarket Datrium's DVX Rackscale system can play.

THREATS
Datrium's architecture is similar to (but different from) that of hyperconverged storage, and there is a risk that potential buyers will not appreciate or understand the differences.