Solution Showcase

Modernizing Data Protection for Converged Systems with Datrium

Date: April 2018 **Authors:** Edwin Yuen, Senior Analyst; and Mike Leone, Senior Analyst

Abstract: Many of the most compelling hyperconverged and converged infrastructure solutions boast built-in data protection capabilities. For the purposes of this paper, the two architectures will be collectively considered and referred to as "HCI" unless specifically noted. Built-in data protection capabilities within some HCI offerings are often seen as differentiators and additional benefits of consolidating IT solution components versus using self-built and protected infrastructure. Datrium offers a solution that combines a tier 1 HCI system with scale-out backup and cloud DR, bringing large-scale mission-critical application performance to petabyte-scale storage with built-in backup and instant restart.

Whenever You Modernize Production, You Must Modernize Protection

For the past several years, ESG's annual IT spending intentions research has shown that *increased use of server virtualization* and *improved data protection* have consistently been two of the most commonly cited IT priorities. According to respondents to the most recent survey, virtualization and data protection are the two likeliest data center modernization initiatives to drive spending in 2018.[1]

Historically, modernizing production platforms has necessitated modernizing protection solutions because the legacy protection mechanisms typically prove themselves inadequate to protect the modern workloads or platforms. This has been true for the last several IT platform shifts (*mainframes > midrange > NetWare > Windows Server*) and even today is a struggle for IT teams as they continue to modernize through virtualization (see Figure 1).[2]

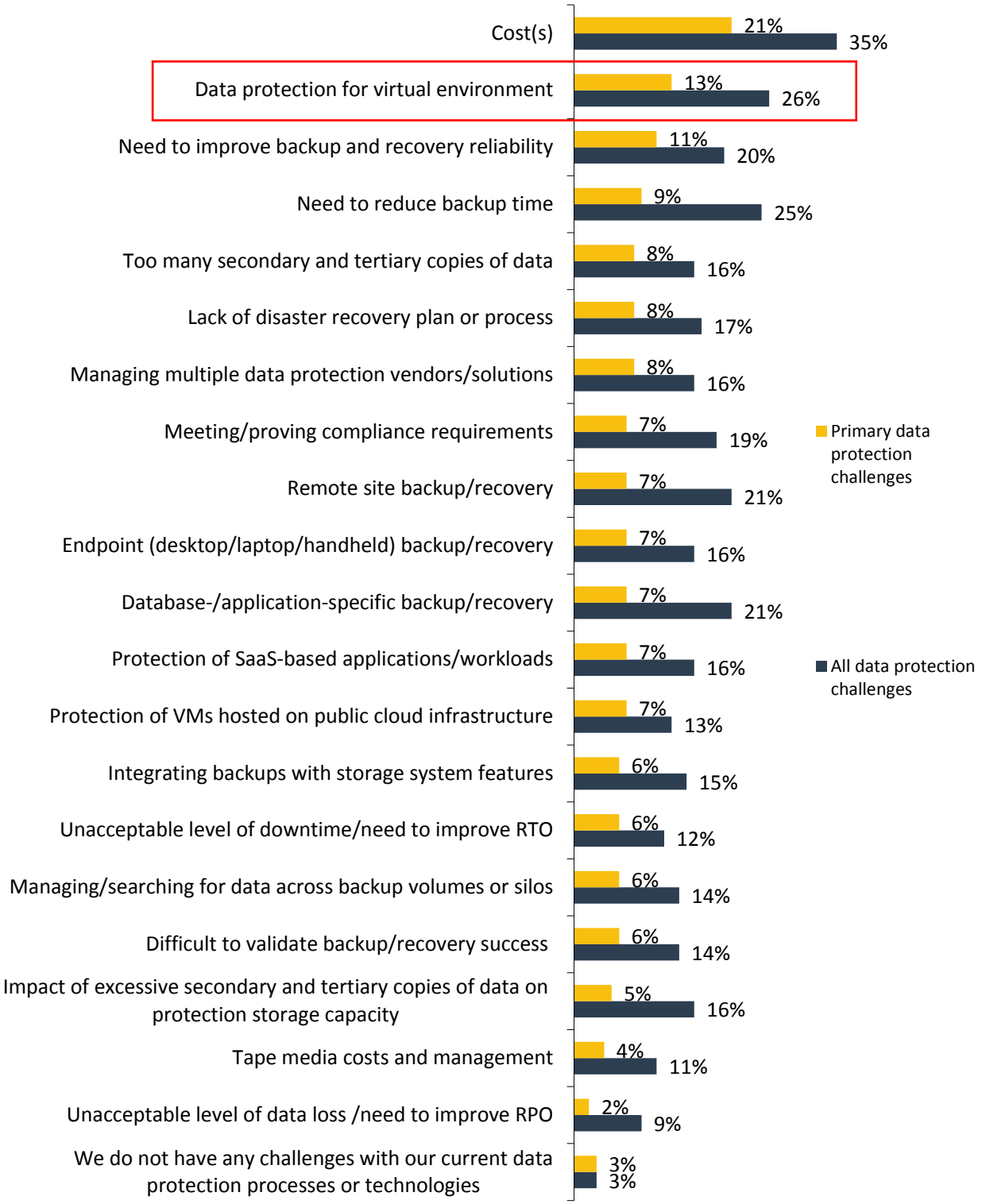
Admittedly, while "virtualization"—including server virtualization most commonly, but also storage or networking virtualization—is compelling for a variety of reasons, the same abstraction that benefits and accelerates production purposes often makes protection tasks more challenging. Considering that the next leap forward in virtualization (i.e., HCI) creates even more abstraction and agility for production purposes, it can also create even more challenges or complexity for protection tasks. That means IT professionals should be even more mindful of data protection when investigating HCI.

[1] Source: ESG Master Survey Results, *2018 IT Spending Intentions Survey*, December 2017.
[2] Source: ESG Research Survey, *Data Protection Modernization Trends*, December 2016.

This ESG Solution Showcase was commissioned by Datrium, based on the previously published ESG Brief, *Why and How to Protect VMs within Hyperconverged and Converged Infrastructures*, January 2018, and is distributed under license from ESG.

Figure 1. Data Protection Challenges

Which of the following would you characterize as challenges with your organization's current data protection processes and technologies? (Percent of respondents, N=387)



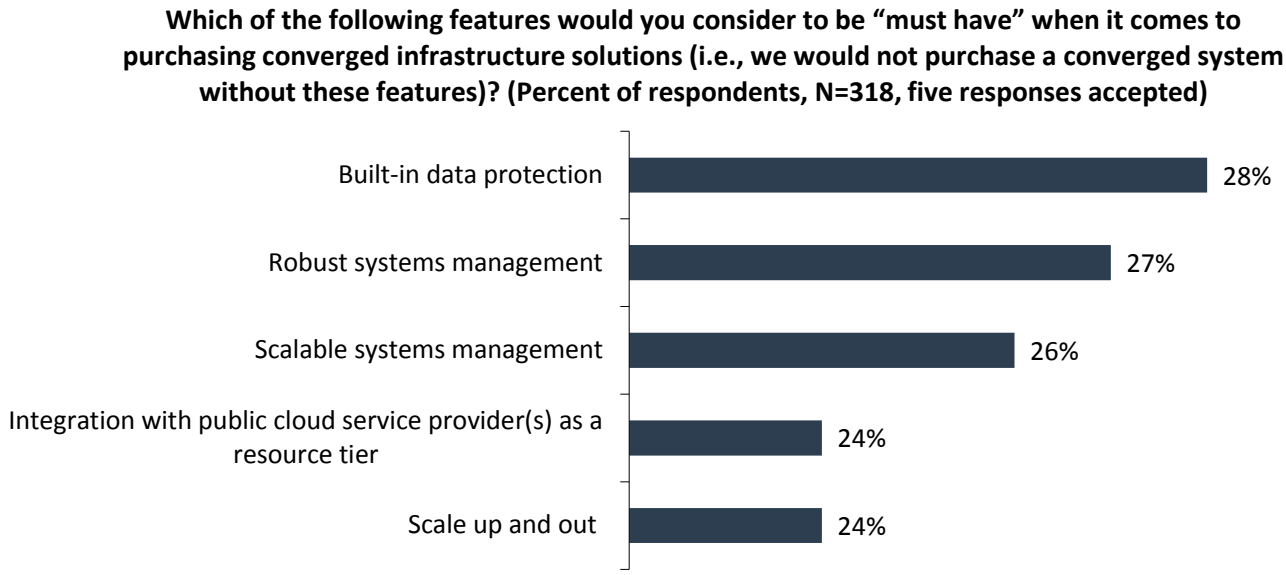
Source: Enterprise Strategy Group, 2018

Don't Forget About Data Protection when Considering HCI

Consolidating IT infrastructure into HCI "building blocks" isn't just about convenience or ease of acquisition for "*Day One*" activities such as deployment. For most IT professionals, the real power of HCI is in the comprehensive and integrated approaches to ongoing management for "*Day Two*" (and every day afterward). Many who are embracing the consolidated manageability aspects are looking to consolidate as many management functions and as much instrumentation as possible, including data protection—and IT operations professionals see it as part of a holistic systems management strategy, instead of a standalone IT process.[3] As such, it should not be a surprise that when IT professionals who are responsible for CI were asked about their "must have" features for new solutions, *built-in data protection* was the most often cited in the list (see Figure 2).[4]

Figure 2 reveals built-in data protection as the most cited purchase criterion among converged infrastructure systems.[5]

Figure 2. Top Five "Must Have" Converged Infrastructure Technology Features



Source: Enterprise Strategy Group

Figure 2 specifically measured interest within converged infrastructure (i.e., larger systems aimed at enterprise data centers). Similar ESG research reveals a similar sentiment among those seeking hyperconverged appliances that they are also looking for built-in data protection.

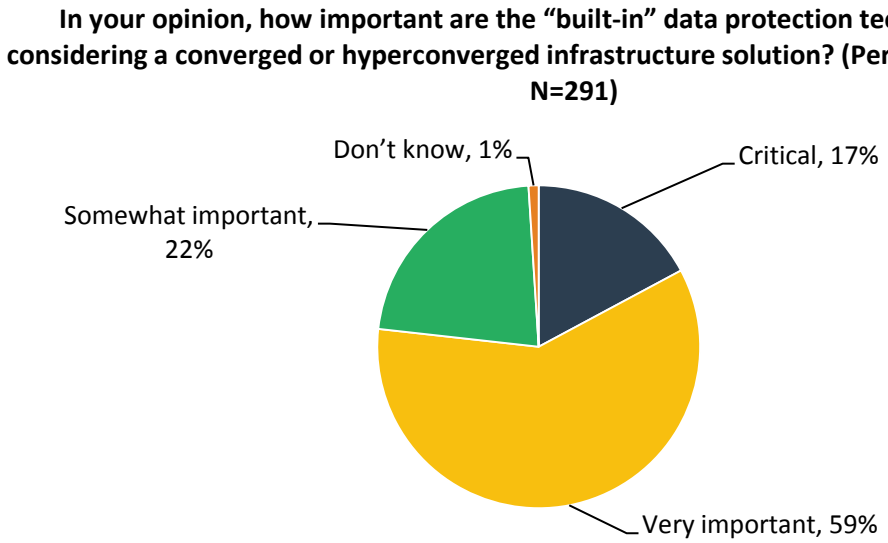
How Should "Built-in Data Protection" Affect Your CI/HCI Strategy

It is worth noting that between the challenges in backing up virtual machines and the ever-heightening requirements for faster and more agile recovery, only one in nine (11%) environments rely solely on VM backups. The other 89% supplement VM backups with other data protection mechanisms, including snapshots and/or replication, delivered within the production hypervisor or storage components. Said another way, if a majority of IT teams are looking to supplement their VM backups with snapshots/replication anyway, why not look for those capabilities to be "built in" to the HCI solution itself?

[3] Source: ESG Brief, *Data Protection Should Be Part of Your Systems Management Strategy*, January 2018.
[4] Source: ESG Master Survey Results, *Converged and Hyperconverged Infrastructure Trends*, October 2017.
[5] ibid.

This is a topic that IT professionals who are responsible for converged infrastructure technology (see Figure 2) and those responsible for data protection agree on (see Figure 3).[6]

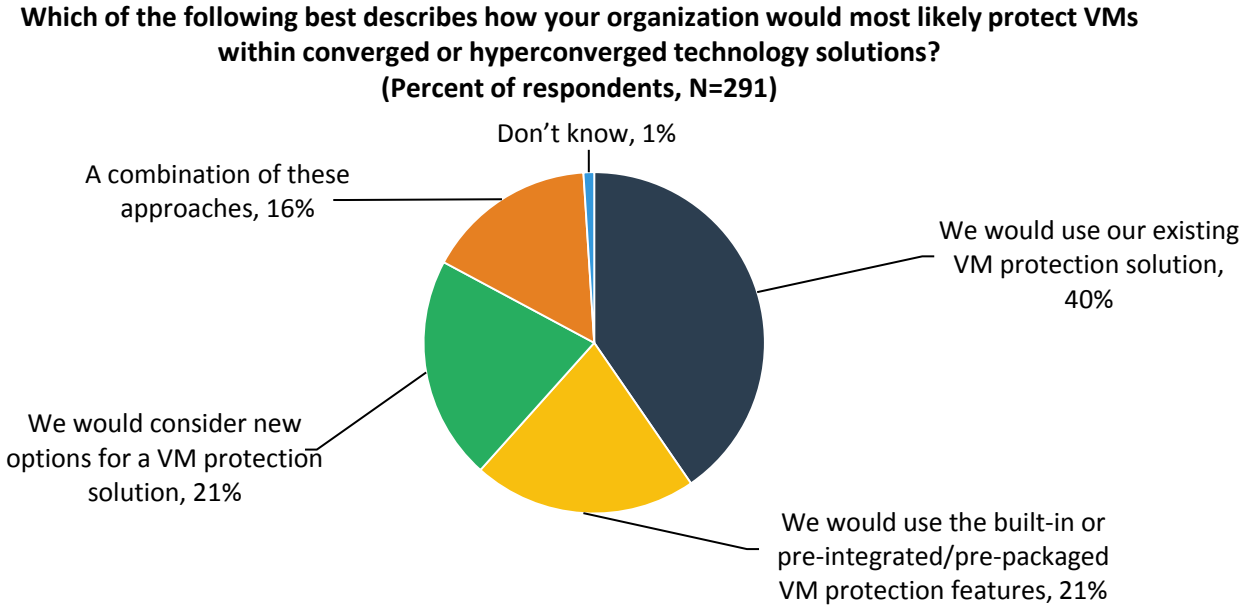
Figure 3. The Importance of Built-in Data Protection Technologies within CI or HCI Solutions



Source: Enterprise Strategy Group

With all of that said, it would be a huge error to assume that built-in data protection features negate the need for any other VM-backup solution. In fact, only 21% of IT professionals responsible for data protection would rely solely on the built-in capabilities and 16% would utilize both the built-in and existing VM-backup solutions (see Figure 4).[7]

Figure 4. Protection of VMs within Converged or Hyperconverged Solutions



Source: Enterprise Strategy Group

[6] Source: ESG Brief, *Reliable Virtualization Protection Continues to Elude Many Organizations*, October 2017.
[7] ibid.

Some of the reasons behind these positions when it comes to protecting converged infrastructure-based VMs include:

- **Confidence in one's existing VM-backup solution**. IT organizations struggling with VM-backup capabilities (see Figure 1) may be more likely to embrace (switch to) HCI's built-in data protection features, while those that already have reliable and agile VM recoverability would likely supplement with the built-in capabilities or stay with their existing solution.

- **VM fluidity.** For many organizations, HCI will supplement their self-built (DIY) infrastructure. In some of those environments, VMs will be initially provisioned in DIY and later migrate to HCI or vice versa. As such, those organizations may want to use a third-party VM-backup solution that can protect VMs regardless of which DIY or HCI stack they currently reside on; though that still doesn't preclude embracing the built-in capabilities within HCI solutions nor the snapshot/replication capabilities of the DIY storage array(s) or hypervisor(s). That said, others may presume that HCI is a more stable and durable architecture, thus presuming that VMs might move to HCI (but not back to DIY again).

- **Operational and economic constraints.** Built-in data protection within HCI is not "free" from a storage consumption perspective. Backing up even a handful of VMs, whether using built-in or third-party tools, still requires additional storage consumption. For those with an established backup mechanism, the first backups using built-in tools will consume an appreciable secondary storage footprint somewhere, which may not be part of the initial deployment investments. Organizations planning on "switching" (from third-party to built-in) should be mindful of this additional storage while operational-recovery versions are doubled in both tools until the switch period of weeks to months is complete.

- **Missing DP features or capabilities.** Not all HCI vendors' built-in DP capabilities are equitable, and not all are comparable to the myriad third-party backup solutions that may have VM backup capabilities; thus, switching is not always viable. That said, most HCI built-in capabilities are powered by storage technologies (i.e., snapshots, replication, or both), which again is what the vast majority of organizations are supplementing their VM backups with. In addition, HCI systems lack a cost-optimized place to store backups for any period of time. Backups are stored on HCI nodes optimized for primary workloads, not low-cost backup, and are configured with SSD capacity, compute, and software licenses.

- **Recovery capability.** Much of the discussions around data protection focus on backup but the more critical part of any data protection solution is the ability to recover the backup. The recovery capabilities of any solution are dependent on the ability of the DP software to leverage the capabilities of the HCI-connected storage technologies. Organizations need to evaluate the benefit of using existing third-party solutions with any additional integration or features from the built-in systems, especially as it impacts their RPO and RTO goals.

Datrium DVX

Datrium DVX offers a solution that addresses both the need for high performance and data protection by matching the speed and scale of a tier 1 HCI-like system for mission-critical applications, with built-in secondary storage, Backup/Instant Restart, and cloud DR. While a single DVX system supports over 1 million vDisk-granular backups, they are compressed, deduped, and redirect-on-write (RoW) always, so retention and efficiency has no impact on speed. Instant recovery is handled at the application level, not the volume level. DVX is thus the first system in the industry that can eliminate the need for separate backup storage and backup applications. More important for recovery time, snapshots are retained within the primary system as backups in a scalable, searchable catalog database. RTO is near instantaneous versus using a separate backup application to copy data from a dedicated backup repository back to the primary system. Cloning or restore are single click operations, guided by built-in snapshot search. Snapshots can be scheduled as often as every 10

minutes, and recovery of a VM takes just seconds. Restart gets critical business applications back online an order of magnitude faster than copy-back restore.

The built-in Backup/Instant Restart feature provides a single console for VM consolidation, acceleration, and protection, and more importantly eliminates the tedious workflow and re-work with a dedicated backup team. The integrity of every backup is validated multiple times per day so recovery can happen when needed. Over a million VM and/or container snapshot backups can be made for instant recovery points in a single DVX. All protection is handled automatically by DVX as a simple plugin to vCenter, through flexible backup policies which can be set up by class of application. VMs can then be automatically included into existing policies based on pattern matching rules to enable data protection at scale. For example, all VMs with "SQL" in their names will inherit a given protection policy set for all database VMs.

Finally, Backup/Instant Restart eliminates dedicated backup devices and software, saving hundreds of thousands of dollars in capital costs and weeks of associated management time. Datrium DVX is the first system in the world to successfully converge primary storage with secondary storage and deliver both low latency primary performance and cost-optimized secondary protection in the same system. Highly efficient, secure and flexible asynchronous (snap-based) replication is also included by default for DR planning. In addition, using Cloud DVX for offsite backups can eliminate an entire second data center deployment. Virtually all arrays and many HCI deployments rely on third-party backup, DR, and WAN optimization because long term snap retention is simply too expensive and may limit performance on their respective systems.

Datrium DVX goes beyond traditional HCI/CI solutions, combining a system for high performance, tier 1 workloads with a cost-optimized backup system, all in one converged solution. Datrium DVX can:

- Provide an integrated data protection solution with the full capabilities of standalone data protection software, including granular backups, snapshots, file-level recovery, always-on global deduplication, compression, erasure coding, and asynchronous replication for DR.
- Offer VM fluidity, with choices in compute nodes, storage media, and hypervisors.
- Provide compute and capacity independence, as well as on-premises and public cloud choice, as key elements of built-in data protection.
- Enable faster application recovery with RTO in seconds.
- Provide high-end performance and security features without compromise, including always-on erasure coding, end-to-end data encryption, performance isolation, and independent compute and storage cluster node scaling well beyond other HCI/CI solutions.

The Bigger Truth

In considering this topic, there are a few truths:

- Converged and hyperconverged infrastructure is and should be part of most modern on-premises data centers.
- Reliably protecting and ensuring recoverability of virtualized infrastructure is still daunting to many, in part due to antiquated approaches, as well as the protection complexities caused by the same abstracted layers that make virtualization so important for modern production.

- Systems management strategies are usually better when one “pane of glass” provides comprehensive insights and actionability across as wide a range of IT functionality as possible; with data protection being an increasingly compelling facet of a holistic approach to systems management.

Put together, it is not surprising that many HCI solutions are including built-in data protection capabilities as part of their solutions—and a lot of organizations ought to be thrilled about that, as long as those IT teams don’t lose track of their broader data protection responsibilities and recovery/availability requirements. Datrium offers a new breed of converged solution that combines high performance compute and resilient storage, with built-in data protection and cloud DR capabilities, and that has unique advantages over standard HCI/CI solutions. Organizations that need a scale-out HCI/CI solution and value a data protection solution that can scale with it should look towards Datrium DVX for their converged system needs.