White Paper

Using HPE SimpliVity Hyperconverged Infrastructure to Improve Data Protection and Recovery Effectiveness

Sponsored by: Hewlett Packard Enterprise

Phil Goodwin
April 2019

IDC OPINION

Data protection operations are undergoing fundamental change in the era of virtual computing. Most organizations designed their data protection infrastructure based on concepts from the client/server era, anchored by traditional backup/recovery (B/R) software that runs once per day and delivers a 24-hour recovery point objective (RPO) and a 24-hour recovery time objective (RTO). Because these service-level agreements (SLAs) are insufficient in most cases, IT organizations have implemented complex systems of snapshots, mirrors, remote replication, purpose-built backup appliances, and cloud backup to address the data availability requirements of the business. Further complications arise from redundant tools on multiple platforms, system incompatibilities, and the operational complexity of making the whole thing work.

Unfortunately, these client/server era technologies are proving inadequate to meet the dynamic computing nature of virtual machines (VMs) that can be deployed across multiple datacenters, public clouds, and SaaS implementations. Data loss threats are also becoming more complex, challenging IT managers to come up with "anywhere, anytime" data recovery schemes. Moreover, there is a new "race to zero" for service levels where business managers are expecting zero RPO and zero RTO.

In light of the changing compute environment plus the complexity and inadequacy of older data protection methods, IT organizations need to take a fresh look at their approach to data protection and availability and unbind themselves from the limitations of client/server era technologies. IDC believes that organizations should seek to evolve their current data protection strategies with the following considerations:

- Tightly integrate the compute infrastructure and data protection.
- Reduce operational complexity.
- Meet the evolving data loss threats, such as ransomware.
- Deliver near-zero service levels.

IN THIS WHITE PAPER

This white paper examines how IT organizations can leverage HPE SimpliVity to improve and simplify their data protection and resiliency capabilities. Proof points have been provided from results of an independent survey of HPE SimpliVity customers conducted by IDC. HPE offers the HPE SimpliVity hyperconverged system, a turnkey hyperconverged infrastructure platform that consists of HPE SimpliVity built on the HPE ProLiant DL380 or HPE Apollo compute platforms. All benefits of the Data Virtualization Platform referenced in this white paper apply to both platforms.

April 2019, IDC #US44972419

SITUATION OVERVIEW

Data protection remains one of the most problematic, labor-intensive, and least loved activities in the datacenter. Administrators must manage scores or hundreds of daily jobs, troubleshoot problems, recover jobs, respond to user requests for restores, and load/unload/manage tapes. B/R is just one of the elements to data protection because datacenter managers also utilize snapshots, mirrors, and replication as supplemental means of data protection. IDC's research has found that a typical organization has thousands of individual database and file system images, each often used for a different purpose — all having their own schedules for snapshots, mirrors, backup, and the like.

The B/R technologies commonly in use today were designed in an era when 24-hour RPOs and RTOs were sufficient to meet most organizational needs. The advent of internet commerce dramatically increased the number of transactions processed by businesses and related financial institutions. A 24-hour RPO could result in a loss of thousands of customer orders for businesses and a loss of millions of transactions per hour for financial institutions. This loss could dramatically increase the cost of downtime for these organizations in not only personnel productivity but also lost business that could never be regained.

Because of the better data protection and availability of SLAs required by business leaders, storage vendors responded with array-based capabilities such as snapshots and mirrors (clones). Snapshots provided a significantly better RPO, such as an hour or less, but could not survive an array or site failure. Mirrors could be made to alternate arrays and sites, but due to long processing times and overhead, they were generally conducted only every 12 hours or so. To meet the continuum of both data loss scenarios and SLA requirements, IT organizations have implemented complex schemes of snapshots that may involve hundreds per volume plus local and remote mirroring. Moreover, neither local mirroring nor remote mirroring has replaced traditional B/R because neither of them can provide the full backup capabilities needed for regulatory or other requirements. Managing this complex scheme requires full-time administrators in most organizations, meaning thousands of man-hours per year.

All this effort is fundamentally unproductive to an organization's overall mission. By this, we mean that data protection activities do not generally add to the competitive advantage of the organization, except to the extent that the organization may be able to bring an application back online faster. However, consumers do not flock to buy widgets simply because the widget maker has great backup. Consequently, the organization seeks to optimize its backup while minimizing the manpower needed to manage it. To the extent that labor related to data protection can be reduced, that freed-up labor can be directed to other projects that may be more strategic to the organization.

Because snapshots offer the best RPO compared with other technologies and often better RTO as well, IT organizations have come to rely on them for a large proportion of their routine data recoveries. This use has exposed the following weaknesses with typical snapshot technology:

- The maximum number of possible snapshots may limit the time frame for recovery. For example, in a system where 256 snapshots are the maximum and 1 snapshot is taken every 15 minutes, the total time covered by the snapshots would be limited to 2.67 days. Any requirement to recover data prior to that time must depend on other methods, such as B/R, with a much higher RPO and therefore lost data.
- Traditional snapshots do not protect against data corruption. Any changed blocks that are corrupt will be captured by the snapshot.
- Snapshots are not "free." Total snapshot overhead is determined by the rate of data change. However, just a 1% rate of change would result in a 2.5 times increase in consumed capacity

for a sequence of 256 snapshots. As a result, many organizations limit the sequence to much fewer than 256 snapshots. This reduces the recovery window even more than that noted in the first bullet point.

- To recover, snapshots need the entire chain up to the point of recovery, much like an incremental backup recovery. If a snapshot chain is broken for any reason (i.e., hardware failure and user error), the chain is unrecoverable and an alternate recovery method will be required, again likely resulting in a higher rate of lost data.

Because traditional backup/recovery is both labor intensive and inadequate to meet business requirements, some IT organizations are transitioning their data protection scheme away from the defined third-party backup model to a model where data protection is built into the compute/storage stack. In many cases, this includes geographic distribution of federated systems that propagate data across the various systems to automatically protect against system and site failures while providing very low RPO. Cloud may be one of the stops in these schemes, which may be for either data protection or long-term data archive. In addition, disaster recovery (DR) sites (i.e., disaster recovery as a service) may be included at a relatively low incremental cost.

The client/server era was characterized by discrete technology stacks, such as server, storage, networks, and applications; backup/recovery was one of these application stacks managed largely as its own entity. The emergence of converged infrastructure did not fundamentally change this dynamic, especially regarding data protection. The existing schemes and product implementations remained largely the same, with no change in service-level delivery.

More recently, the implementation of hyperconverged systems, which include not only a preconfigured hardware system but also an integrated operating environment, offers an opportunity to simultaneously simplify data protection and improve service-level delivery and certainty. With hyperconverged systems, data protection does not need to be an add-on application, requiring separate testing, qualification, and management. Instead, data protection can be built into the operating environment. Thus data protection is below the hypervisor, aligning it with the system operations and applications. Policies and operations are aligned with the virtual machine, providing a VM-centric backup and recovery capability. Offsite data storage and DR become a by-product of the overall scheme where policies can be managed globally.

FUTURE OUTLOOK

Hyperconverged infrastructure is much more than a bundling of components. By tightly integrating the hardware infrastructure with the software operating environment, hyperconverged systems simplify deployment and can significantly reduce the labor needed to manage them. From a data protection perspective, data protection schemes, methods, and implementations are consistent across the environment and globally managed. Nodes may be federated within the datacenter and across datacenters as well as geographically dispersed, and they may include cloud implementations. The HPE SimpliVity architecture has been designed and optimized for efficiently managing data across a federation. A federation consists of systems that are both local (within the same datacenter) and remote (dispersed datacenters). A key capability of this federated architecture is efficient data mobility within and across datacenters, which results in improved data protection and availability.

HPE SimpliVity is a hyperconverged system designed to provide enterprise system capabilities with cloudlike economics. In summary, this platform provides resilient infrastructure and data services for virtualized workloads running on it. It includes integrated data protection that eliminates the need for

separately licensed backup/recovery or remote replication software or purpose-built backup appliances.

HPE SimpliVity data protection and disaster recovery technologies differ significantly from conventional backup, recovery, or snapshot methods. In many respects, it offers the best of both worlds: the complete protection of backup and disaster recovery software with the speed and RPO/RTO of snapshots. Highlights include:

- There is no practical limit to the number of backups that can be created and retained. HPE SimpliVity engineers have tested thousands of backups, and the upper tested maximum currently exceeds 200,000.
- Each copy is a full backup. There is no chain of changes that can be broken, corrupted, or needed for replay. This means that you can delete the original VM and the backup will still be functional.
- Backups include a clone of the VM parent object. The full VM tree is logically copied in its entirety, and all VM trees are the same length:
  - If the original VM is deleted, it can still be recovered from the backup set.
- Storage capacity improves over time (i.e., reduced) because all data, metadata, and pointers are deduplicated:
  - Rehydration is not necessary as long as the data is on HPE SimpliVity.
- Backup images can be transferred to an offsite system for DR:
  - Two copies are automatically kept on different nodes.
  - Copied blocks cannot be deleted.
- In addition, HPE offers HPE SimpliVity Rapid DR — an optional, standalone disaster recovery orchestration tool that was developed specifically to work with HPE SimpliVity. It was designed to automate the failover of a preconfigured set of VMs from the production site to a secondary site.
- Data and metadata corruption can be detected with a "fingerprint" of checksum and hash values.
- Each HPE SimpliVity node is functionally an independent system:
  - Corruption on one node cannot affect another.
- Because HPE SimpliVity deduplicates, compresses, and optimizes data the first time it is written to disk and maintains it in that state for its life cycle, backup copies are already deduplicated and compressed, thereby eliminating the need for third-party software or purpose-built backup appliances.

The HPE SimpliVity approach to data protection is to provide all the benefits of a full backup and data replication at the speed of a snapshot. The methodology is also an excellent defense against ransomware because all prior images are saved as full images, permitting recovery at a point prior to the ransomware (or any other) infection.

IDC was engaged to conduct an independent survey of HPE SimpliVity users. The purpose was to determine whether, and to what degree, organizations benefited from the company's approach to data protection. We asked several questions regarding the customers' before-and-after experiences with HPE SimpliVity. A total of 83 respondents participated in the survey. The findings of the survey are described in Figures 1–4.

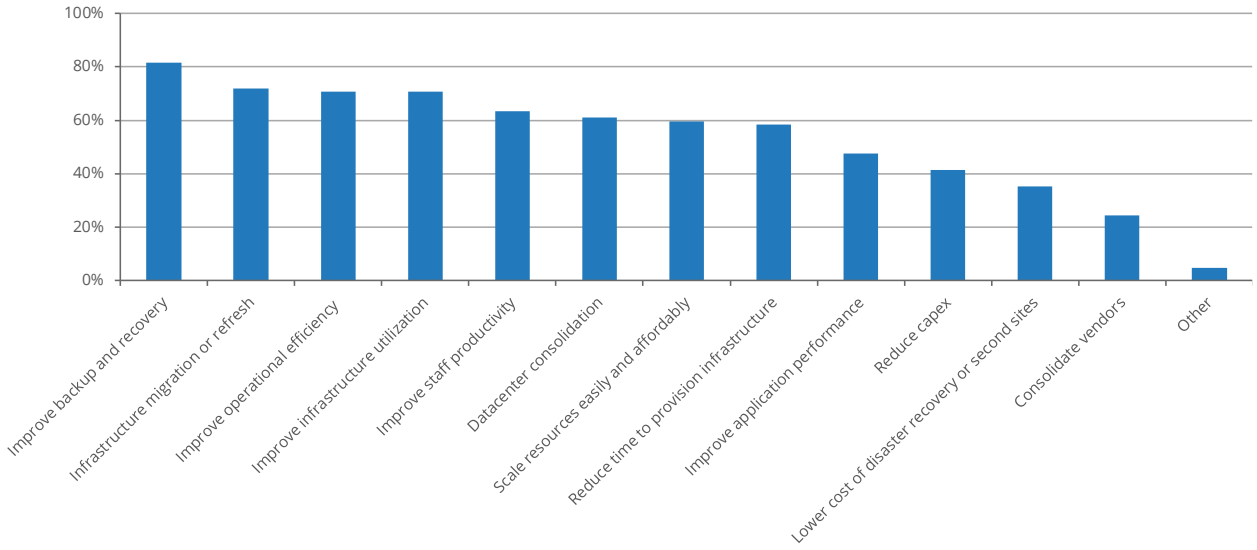
We asked respondents to tell us what system deployment benefits they had expected to attain by implementing HPE SimpliVity. The results are shown in Figure 1.

Figure 1 illustrates that improved backup/recovery was the number 1 objective cited by respondents, with 82% of respondents including it on their list of objectives (multiple answers were permitted). Interestingly, of the top 4 items, two items apply directly or indirectly to data protection operations (improve backup/recovery and improve operational efficiency).

FIGURE 1

Challenges Addressed with HPE SimpliVity Deployment

Q. What are all challenges that your organization sought to address with the use of an HPE SimpliVity hyperconverged infrastructure?



n = 83

Source: IDC, 2018

Of course, expectations and results are two different things. So we asked respondents about their experiences regarding data protection with HPE SimpliVity.

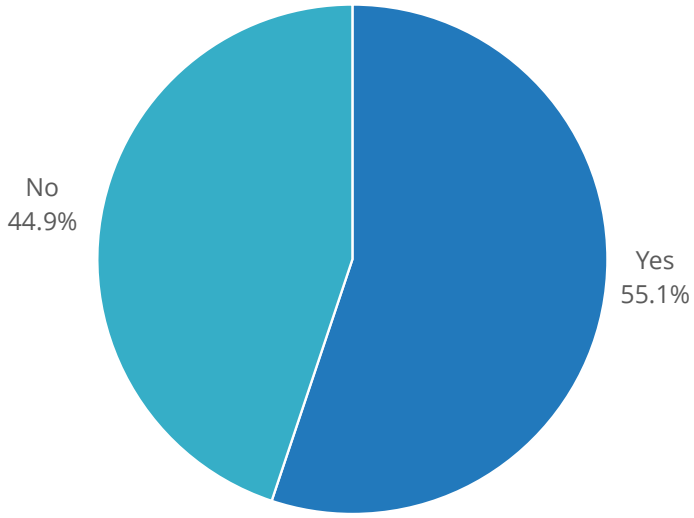
One measure of an organization's improvement to backup operations is the ability to consolidate backup tools and retire redundant products. In this regard, we wanted to know how many organizations had been able to retire their current backup and/or replication solutions as part of simplifying operations. The results are shown in Figure 2.

Figure 2 shows that 51% of respondents were able to entirely retire their current backup software and replace it with the HPE SimpliVity solution. The elimination of a separate backup application can reasonably imply a reduction in licensing and maintenance fees and simplification of the operating environment.

FIGURE 2

Backup/Recovery Tool Retirement

Q. *Have you retired the use of third-party backup and/or replication solutions for workloads running on HPE SimpliVity hyperconverged infrastructure in lieu of HPE SimpliVity built-in data protection?*



n = 83

Note: For more information, see *HPE Datacenters Leverage HPE SimpliVity to Drive Operational Simplicity, Improved Performance, and Other Critical Datacenter Benefits* (IDC #US44216718, August 2018).

Source: IDC white paper, sponsored by HPE, 2018

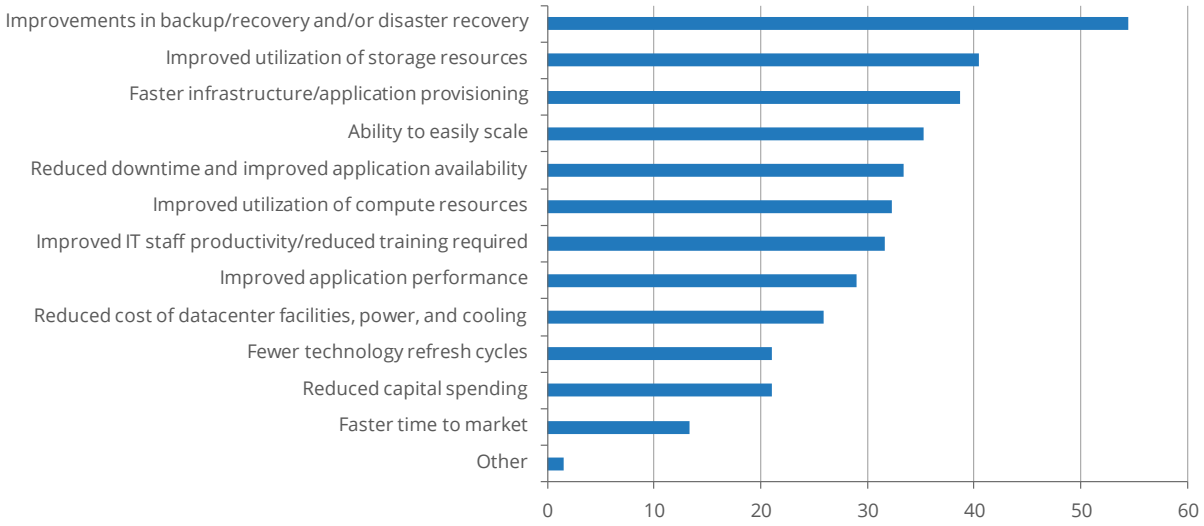
To drill deeper, we wanted to know not just whether an improvement in data protection had been made but also how much of an improvement. So we asked respondents to tell us what *percentage* improvement they had experienced in various areas. The results are shown in Figure 3.

As illustrated in Figure 3, HPE SimpliVity customers indicated a 54.5% improvement in their B/R and DR operations. This is a significant improvement. In addition, customers reported a 40.4% reduction in storage resource utilization, a 33.4% reduction in downtime, and improved application availability.

FIGURE 3

Percentage Improvement by Area

Q. What percentage improvement has your organization experienced from the use of HPE SimpliVity hyperconverged infrastructure in any of these areas?



n = 83

Note: For more information, see *HPE Datacenters Leverage HPE SimpliVity to Drive Operational Simplicity, Improved Performance, and Other Critical Datacenter Benefits* (IDC #US44216718, August 2018).

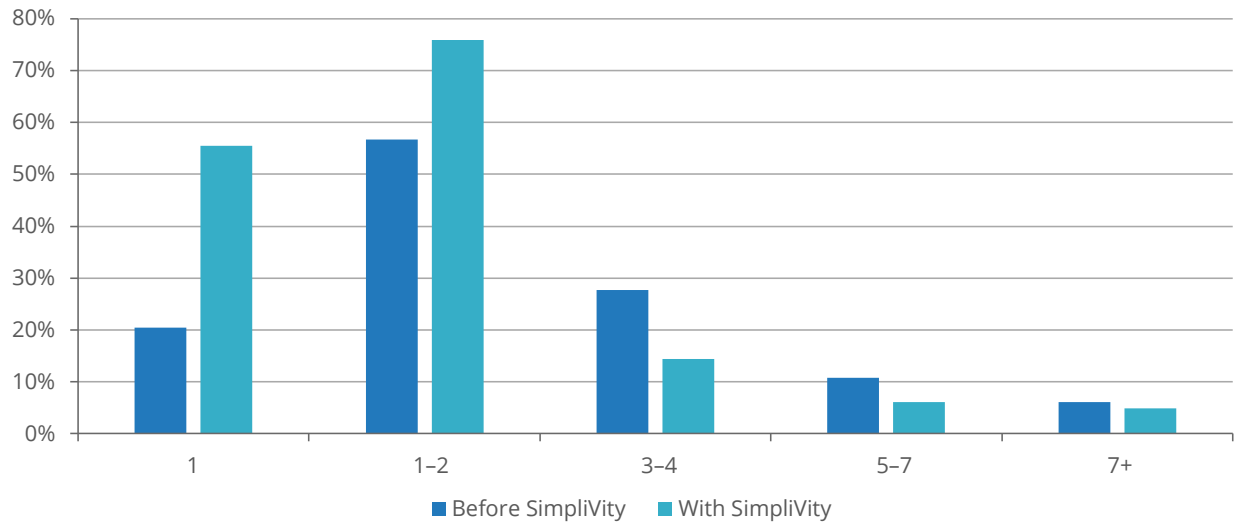
Source: IDC white paper, sponsored by HPE, 2018

Staff productivity is clearly a major area of concern for all IT operations, so we asked respondents to tell us about their staffing requirements for B/R activities before and after implementing an HPE SimpliVity solution. Figure 4 shows the shift from higher levels of staffing before HPE SimpliVity to lower levels after HPE SimpliVity.

Prior to implementing the HPE SimpliVity platform, customers reported an average of 2.11 full-time equivalents (FTEs) required to manage their infrastructure, which dropped to just 1.48 FTEs after implementing the platform. This statistic covers all infrastructure management, not just B/R or DR, but given the improvements seen in backup operations, it is certain that these areas were significant contributors to the results. In fact, by drilling further, we were able to determine that the amount of time spent on backup operations dropped from 18.7% of staff time to 10.4%, a 44% reduction. This means that IT staff members are now available to take on other, more strategic tasks.

FIGURE 4

IT FTE Infrastructure Management Comparison



n = 83

Note: For more information, see *HPE Datacenters Leverage HPE SimpliVity to Drive Operational Simplicity, Improved Performance, and Other Critical Datacenter Benefits* (IDC #US44216718, August 2018).

Source: IDC white paper, sponsored by HPE, 2018

CHALLENGES/OPPORTUNITIES

To benefit most from a new technology, IT organizations must be willing to rethink their infrastructure and application deployment models to overcome "the way we've always done it" syndrome. Indeed, changes to infrastructure impact both people and processes, and these changes require adaptation of staff and users. Even though the ultimate benefits may be compelling, effort will be required to achieve them. In addition, the data protection market using HCI systems is becoming more crowded; so HPE will need to develop new capabilities rapidly and in line with market requirements to remain competitive.

In addition, it is unlikely that most organizations can, or will be willing to, change their entire infrastructure to HPE SimpliVity at one time. Thus IT managers should look for an application-by-application implementation that is controllable and provable as they move to hyperconverged infrastructure over time. To address this, HPE is expanding its HPE SimpliVity partnership ecosystem, including Veeam and MicroFocus, to address key use cases beyond what HPE SimpliVity currently offers.

CONCLUSION

IT organizations must fundamentally rethink their data protection and resiliency strategies to meet increasingly stringent data and application availability requirements. At the same time, managers must find ways to make existing staff more productive and focused on strategic activities. Anytime an opportunity comes along to simplify and reduce or eliminate a task is an opportunity that must be taken.

The predominant backup/recovery methodologies, where the backup environment is separate from the compute environment, were designed for the client/server era. While it is possible to use these same techniques on newer architectures, they are unlikely to bend the cost and labor curves enough to make a big impact on overall IT operations. Hyperconverged systems include built-in resiliency and data protection schemes, which automate operations such that they become a by-product of the compute process, offering a real opportunity to improve data protection, simplify operations, and minimize costly downtime.

IDC's survey of HPE SimpliVity customers revealed that these organizations have experienced tangible, measurable results in terms of higher data availability, better operations, and reduced labor related to data backup and DR. In more than half of the cases, organizations were able to retire existing data protection solutions. In addition, the 31.6% reduction in time spent managing the B/R environment (refer back to Figure 3) is "found" time that can be utilized for more advantageous purposes. An organization with just two full-time storage administrators would find 1,315 FTE hours of freed-up time per year. We can certainly say that the HPE SimpliVity systems have lived up to the expectations of IT managers in the organizations that we surveyed.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com