White Paper

Taking Convergence to the Next Level with a Self-protecting Enterprise Cloud

Collapsing Traditional Storage Silos for Enhanced Disaster Recovery and Simplicity

By Christophe Bertrand, ESG Senior Analyst

May 2019

This ESG White Paper was commissioned by Datrium and is distributed under license from ESG.

Contents

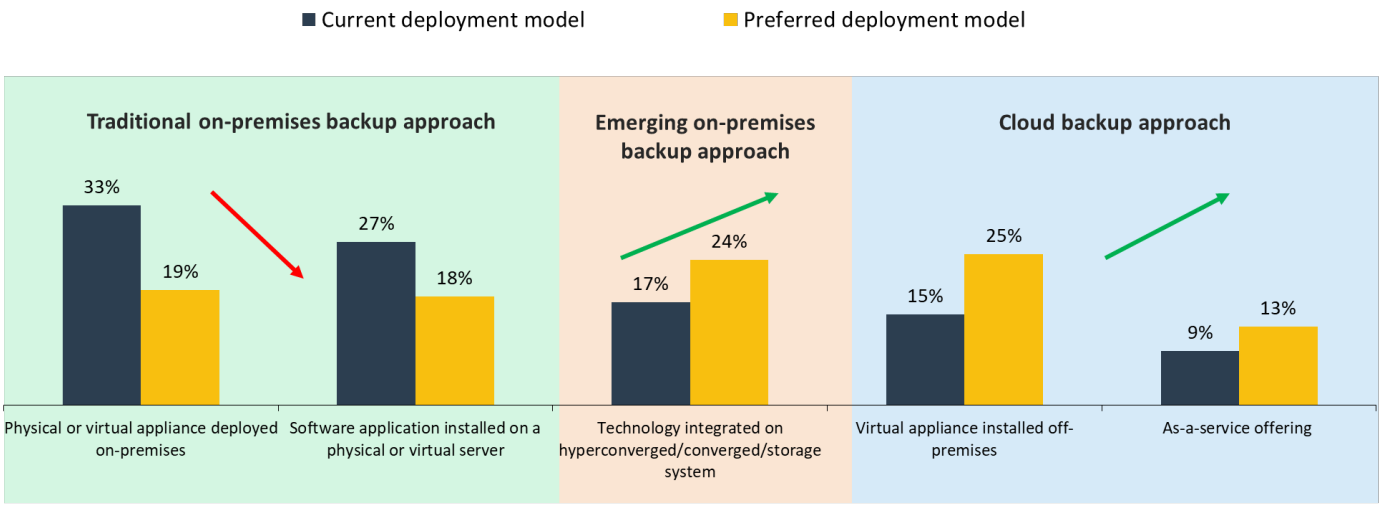
Market Landscape

In ESG's 2019 IT Spending Intentions report, 66% of respondents acknowledged that IT environments are more complex or significantly more complex than they were two years ago.[1] One of the ways organizations are alleviating these operational pressures is by adopting newer approaches in their infrastructure such as converged infrastructure (CI) and hyperconverged infrastructure (HCI) deployments. Additional ESG research reveals scalability, simplified management, improved TCO, and faster provisioning are among the top advantages of converged technology reported by respondents, making CI and HCI platforms a great fit in the fight against complexity.[2]

Not surprisingly, where critical data lands, the need for adequate backup and recovery mechanisms naturally appears. Most hardware platforms provide built-in mechanisms and our research indicates that the new converged environments are following this trend, redefining conventional data protection in the process. According to ESG's 2018 Data Protection Landscape survey, legacy backup deployment models currently dominate in the data center but are fading in popularity. Physical and virtual backup appliances deployed on-premises and software installed on physical or virtual servers are giving way to CI/HCI deployments and cloud-based data protection (see Figure 1).[3]

Figure 1. Most Common and Preferred Backup Solution Deployment Models

What is your organization's most common deployment model for solutions that back up on-premises systems? Going forward, what will likely be your organization's backup solution deployment model preference (i.e., what would you be most likely to standardize on)? (Percent of respondents, N=320)



Source: Enterprise Strategy Group

Another notable market trend picked up in ESG research is growing concern for the high volume of secondary storage that organizations are accumulating and using for data analytics and machine learning. Twenty-six percent of survey respondents reported that managing high volumes of secondary and tertiary copies is among their top data protection challenges.[4] The rapid growth of secondary storage (backups, email, shared files, archives, and the like) is also having an impact on storage vendor offerings. To further alleviate the impact of complex data-related processes, vendors of primary storage and HCI solutions are incorporating secondary storage applications, such as backup and disaster recovery orchestration, into their products. Secondary storage vendors are making moves into primary storage territory. Converged,

[1] Source: ESG Master Survey Results, *2019 Technology Spending Intentions Survey,* March 2019.
[2] Source: ESG Master Survey Results, *Converged and Hyperconverged Infrastructure Infrastructure Trends*, October 2017.
[3] Source: ESG Master Survey Results, *2018 Data Protection Landscape Survey,* November 2018.
[4] ibid.

scale-out storage solutions are breaking down traditional storage siloes to support primary and secondary storage needs, which in turn impacts how data should be protected, and with what processes. And a majority of vendors are adding cloud-friendly capabilities, including deduplication and cloud backup extensions, as the public cloud becomes a top destination for secondary storage. These implementations will vary by vendor.

Security is a key requirement, challenge, and concern for enterprises and their infrastructure. Against a backdrop of constant and elaborate cyber attacks, organizations need to invest in protecting their environment and primary data. According to ESG research, nearly two-thirds of surveyed organizations across North America and Western Europe experienced a ransomware attack at some point in the previous year, with 22% reporting weekly attacks.[5] The attacks have helped make cybersecurity a target of IT investment, and spending is accelerating. ESG research also shows that strengthening cybersecurity tools and processes is the top focus area for IT in 2019. In fact, improved cybersecurity continues to be the top business driver for technology spending and also remains the key IT investment justification.[6]

Being an IT leader is not getting easier, with 66% saying that IT is more complex compared to two years ago.[7] Any solution that can help simplify the infrastructure as organizations digitally transform will help redirect resources and efforts toward supporting business outcomes.

Finally, as is often the case in IT, money and budgets matter! Paralleling the growth in IT environment complexity is an increasing focus on costs. Indeed, more than one-third of IT professionals reported cost as one of the challenges their organization faces in its data protection environment, and specifically, storage hardware is the most cited cost concern, reported by 66% of organizations. Reducing cost is also a mandate from IT leadership, reported by 35% of respondents, second only to security (42%), a topic that is top-of-mind across a wide swathe of IT disciplines.[8]

Convergence Revisited

The ESG 2019 Technology Spending Intentions Survey noted that digital transformation continues to be in full swing. The anticipated business value from digital transformation encompasses operational efficiencies, better and more differentiated customer experiences, the creation of new data-centric products and services, and the development of new business models. The potential for such sweeping, organization-wide benefits has meant line-of-business managers are getting more involved in technology buying decisions.

However, as is evident from the survey, these same managers have an increasingly dim view of IT as a competitive differentiator. Deployment complexity and time-to-market for new applications are the primary grievances business users have with in-house IT. Not surprisingly, IaaS and SaaS cloud-based offerings are also gaining in popularity, allowing users to realize benefits such as reduced administration, faster deployment, increased flexibility, and cost efficiencies that make IaaS and SaaS infrastructure approaches more responsive to the needs of digital transformation.

HCI is a direct response to the complexity and deployment challenges that have plagued IT. HCI solutions integrate compute, storage, and network components on commodity hardware, with software-defined networking and storage management. This comprehensive, integrated approach eliminates physically distinct infrastructure silos in the data center. When asked about their most problematic IT skills shortages, organizations routinely list disciplines associated with managing legacy infrastructure silos, most notably in database, network, and storage administration.[9] Converging components and abstracting administration complexity through holistic management software alleviates IT's dependence

[5] Source: ESG Research Report, *2018 IT Spending Intentions Survey*, February 2018.
[6] Source: ESG Master Survey Results, *2019 Technology Spending Intentions Survey,* March 2019.
[7] ibid
[8] Source: ESG Master Survey Results, *2018 Data Protection Landscape Survey,* November 2018.
[9] Source: ESG Master Survey Results, *2019 Technology Spending Intentions Survey,* March 2019.

on silo-specific expertise and enables greater use of automation, significantly easing the burden of provisioning and managing resources.

Converged solutions bring the scale, cost, and flexibility benefits of IaaS and SaaS into the data center. This gives in-house IT organizations access to the capabilities that business users have been demanding for digital transformation initiatives. However, HCI only addresses the first level of inefficiencies from legacy data center silos. Storage systems have long been siloed into primary and secondary storage, based on the data's performance, access, and protection characteristics. Current marketplace trends show primary, backup, and DR orchestration integrated into a single converged storage solution, extending the benefits of convergence even further.

Converging Secondary Storage

Secondary storage has conventionally demanded less of its storage architecture than its primary counterpart. That notion is changing. The scale-out architecture of converged infrastructures makes it feasible to run and protect data in the same place. With rich built-in data protection, data management, and data reduction capabilities, converged solutions can satisfy all storage needs from a single pool and make third-party backup and recovery software and hardware unnecessary.

HCI vendors use erasure coding, built-in data compression, and deduplication algorithms to make efficient use of available storage capacity, with the automatic verification protections that you would expect from leading backup solutions. Many vendors are moving to all-flash storage back-ends in order to provide performance at scale, but cost can be a factor that limits a solution's ability to converge secondary storage. A tiered approach that combines flash and disk can solve this problem, providing high performance and scale while maintaining the benefits of converged primary and secondary storage.

As is often the case in a market with multiple vendors, all of these characteristics need to be evaluated very closely because not all vendors provide the same capabilities. Areas that end-users should scrutinize to cut through the marketing hype include deduplication methods, erasure coding efficiency to handle drive failures, and the capacity associated with these processes.

Converging DR Orchestration

Tier-1 enterprise workloads are moving to converged, private cloud infrastructures, and as a result, SLAs are dictating the need for effective replication and disaster recovery (DR) capabilities. These mission-critical applications frequently need near-zero recovery time objectives (RTOs) and recovery point objectives (RPOs) measured in minutes. Converged infrastructure solutions by their very nature may be able to implement comprehensive DR orchestration capabilities since they manage both compute and storage. DR orchestration features must be built-in and cloud-native if they are to avoid tradeoffs that complicate copy data management and automated failover. One of the major benefits of having DR as part of a converged solution is that the data and the operational environment are in sync, enabling complete and rapid recovery of VMs and their associated data.

Converging primary, backup, and DR storage on a single system is the logical next step from a converged infrastructure. By eliminating siloed storage, IT organizations can make better use of available capacity, dramatically simplify storage, data protection, and disaster recovery management, and at the same time increase operational efficiency, optimize IT investments, and reduce costs to give a better ROI and TCO. The software-defined approach to data management and DR orchestration creates a more agile and flexible way to provision and protect storage and to manage data across the entire lifecycle. The long-term benefits of this approach to storage include improved SLAs, reduced downtime, reduced risk, and greater regulatory compliance.

A Platform Optimized for All Data

Conventional three-tier storage systems are often too complex, too expensive, and too difficult to manage. The silos of expertise needed to operate these systems have stretched IT to the point where the shortage of skilled staff needed to manage the variety of data hosted by most enterprises is creating a problem.

Converged infrastructure solutions are a step in the right direction but they don't go far enough to solve escalating demand for data and storage. Part of the problem is the rapid increase in the number of copies of data across the enterprise for applications such as data protection, dev/test, data analytics, archiving, and compliance. This is one of the ways the Datrium solution can help. Datrium's platform converges the data lifecycle across on-prem and hyperscaler clouds in a single, secure platform. It eliminates the need for parallel hardware and software backup and DR stacks by integrating all components and aspects of the backup and DR into a single system with unified data management.

The proliferation of data copies is a data management issue. Datrium addresses this by applying compression, deduplication, and encryption the first time data is written, which helps maintain efficiency and security throughout the data lifecycle. Universal deduplication also optimizes WAN data transfers, for remote site replication, and cloud storage costs. The company was recently assigned a US patent for its high-performance deduplication primary data storage system technology. A global catalog makes it possible to find objects easily and retrieve them, no matter where they are.

Datrium integrates primary, secondary, and DR on a software defined converged infrastructure system using a tiered media approach to deliver high-performance and cost- and space-efficient storage. The converged storage cluster uses separate layers, with high-performance host-based SSDs for primary I/O in the performance layer and a highly resilient protection layer using durable disk for dense data storage with dual components, double fault protection, and more. Using high-performance flash in the performance tier enables the converged system to rival the fastest all-flash storage arrays. And the linear scalability of the cluster will provide for high-bandwidth data ingestion that competes with specialized secondary storage appliances.

From an architectural perspective, a self-protected Datrium system and one or more accompanying systems deployed at another location or in the cloud are managed by a unified cloud DR orchestration application.

VM-based data protection and DR orchestration in the converged solutions can eliminate the need for third-party backup and replication tools. Fully integrated DR orchestration enables frequent verification and compliance checks. Unified data management reduces storage management overhead by managing data at the VM level across clouds. This eliminates the storage administration requirement to manage LUNs, SAN, and other storage artifacts.

As highlighted, security is a critical mandate for IT. In this area, Datrium provides unique built-in encryption that persists from inception, through the network, at rest, over the WAN, and in the cloud.

The Bigger Truth

Digital transformation is shaking up the IT world and driving in-house IT organizations to compete with the simplicity, speed, and cost savings of the on-demand economy. Traditional IT infrastructures are slow to change and cumbersome to manage. Converged infrastructures, on the other hand, mimic the massive infrastructure implementations of cloud service providers, such as Amazon and Google. They give in-house IT teams a private cloud in the data center and the flexibility to extend to public clouds when they need to respond quickly to business user needs.

The rapid growth of corporate data is straining existing data protection and DR infrastructures. Converging primary, backup, and DR enables IT to protect data at enterprise scale. Historically, IT would have to choose either a high-cost performance architecture or a lower-cost protection architecture. Newer converged storage technology like Datrium's offers both.

Converging primary, backup, and DR is a hard thing for most IT vendors to handle. But emerging vendors like Datrium have solutions that make it both easy and scalable. These solutions will help solve the growing problem of fragmentation of systems and overall IT complexity.

Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides actionable insight and intelligence to the global IT community.