

White Paper

Microsoft Azure StorSimple: The Next Step for Hybrid Storage

By Mark Peters, Senior Analyst

July 2014

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



Contents

The Growing Challenges of Data Growth	3
The Arrival of Hybrid Storage	3
Microsoft Azure StorSimple: Contemporary Hybrid Storage	4
Market Relevance and Analysis	6
The Bigger Truth	7

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.

The Growing Challenges of Data Growth

The world of enterprise storage—an extensive range of storage arrays, Fibre Channel SANs, iSCSI SANs, and NAS boxes, all with networking components, applications, operating systems, and, increasingly, cloud backup targets—is a complex one. And as each of these systems gets upgraded with more functionality, such as replication and deduplication, it grows even more complex and harder to manage.

Meanwhile, the amount of data—especially contemporary, critical data—continues to grow at an astounding pace. Much of it is unstructured data, the sort produced by office and document applications, e-mail, content management systems, graphics, and streaming data. According to ESG research, the rapid growth, and subsequent management, of such unstructured data remains a primary storage environment challenge for many enterprises.¹ Enterprise storage managers are running short of viable options to keep up with the continually escalating data growth, application demands, and user expectations, while simultaneously having to deal with the less-glamorous issues of data security requirements and budget constraints. Even with feature-rich and highly scalable storage solutions, storage management tasks are becoming increasingly cumbersome and resource-intensive. For many of these overworked storage managers, their job must at times seem like a workplace version of the Whack-a-Mole arcade game, where, to maintain control, they have to keep endlessly pounding the varmints back into their holes, day in and day out!

IT teams find themselves constantly acquiring storage, and are then often shocked to find that their recently installed high-ticket storage systems almost immediately need capacity upgrades to keep up with all the data. This isn't just a capacity and budgeting problem; it also creates risks for application performance, availability, and backup. Furthermore, strict regulations and compliance requirements make it imperative for corporations to store historical copies of data for much longer periods of time than before, often maintaining its availability to application servers as well. While simple file sharing cloud storage can represent a cost and capacity answer, it invariably also leads to a further set of challenges when it comes to control, performance, and security. In recent years, a better cloud storage option—a cloud-*integrated* approach—has been available to address the growing number of storage-related challenges in a manner that still delivers and maintains enterprise quality control.

The Arrival of Hybrid Storage

Including cloud services as part of an overall storage solution spreads the benefits associated with clouds throughout the business. The scalability and pay-as-you-grow elements of a cloud model mean IT users do not have to pay for a sizeable storage upgrade if only a small amount of additional capacity is needed. And, if a lot more capacity *is* needed, it's still easier to add it than to purchase more storage arrays. This is a leading explanation for why 23% of ESG research respondents cited cloud computing as one of their most important 2014 IT priorities and 72% of respondents expected to increase spending on it in the upcoming year.²

Cloud storage services also typically have built-in redundancy. Since a single upload of data is automatically replicated as part of the service, storage managers don't need to upload it numerous times for backup or archive purposes. Not only that, they can keep data on cloud storage for as long as they want. They don't have to worry about copying it from system to system over the years as they wear out and are replaced (the systems, not the managers!).

However, transferring data to cloud storage is not without some deficiencies. Data uploads and downloads tend to have much higher latencies with less throughput than traditional enterprise storage. That means it's good for the light to moderate workloads of (for instance) file servers, but not so good for the heavy workloads and high I/O processing requirements of (as an example) transactional database servers.

When people discuss data growth, they tend to focus on all the *new* data being created (and certainly it is often "short-term critical"), but in reality, there's also a significant share of the overall problem that is the older, "colder" data that it all becomes and that keeps accumulating. Incoming data and the data that applications are actively

¹ Source: ESG Research Report, [2012 Storage Market Survey](#), November 2012.

² Source: ESG Research Report, [2014 IT Spending Intentions](#), February 2014.

using comprise the working set and this is constantly morphing as new data is added to the system and old data becomes inactive. If we view the epidemic data growth problem in terms of these two types of data, we can envision a balanced solution that integrates SAN storage with cloud storage, putting the working set on the SAN side and the inactive, colder data on the cloud side.

The blending of data center storage and public cloud storage was initially referred to as “cloud-integrated storage” (CiS), and one of the leading developers of that technology was StorSimple, which Microsoft acquired in November 2012. CiS combines the two types of storage into one effective, managed pool with the cloud storage elements effectively treated as an online remote tier. Following the acquisition of StorSimple, Microsoft naturally combined its new CiS feature set—including primary storage data deduplication, integrated data tiering, data compression, encryption, backup and disaster recovery, automated storage management, dynamic capacity expansion, and automated offsite protection—with the existing Microsoft Azure cloud to deliver what it called hybrid cloud storage (HCS). While the cloud-based storage within HCS can’t yet replace high-performance local storage systems, it can act as a semi-active tier where applications and users access or restore data from the cloud without using backup catalogs or archival restore procedures.

In a nutshell, HCS allows IT to extend its storage infrastructure in order to exploit the cloud’s scalability for application data, without sacrificing performance or central management.

Microsoft Azure StorSimple: Contemporary Hybrid Storage

Keeping with the mission to simplify data management, the next logical step in the development of HCS is a hybrid storage offering that Microsoft has named Microsoft Azure StorSimple. The concept of Microsoft Azure StorSimple is to move the storage control mechanisms *from* on-premises *to* the Azure cloud while simultaneously delivering more capacity and performance on-premises. Microsoft Azure StorSimple arrays enable its enterprise customers to integrate Azure Storage with more data center servers and applications to address such problems as data growth, capacity management, data protection, archiving, and disaster recovery.

There are three key points to be made with respect to the differentiation and value of this new manifestation of Hybrid Storage:

Storage Management Is Handled within Azure

With traditional on-premises storage, it’s nearly impossible to know the status of storage systems across a distributed or complex enterprise. To manage storage globally with traditional approaches, an enterprise needs multiple people at multiple sites, most often with a heterogeneous range of skillsets as well. Microsoft Azure StorSimple has been designed to streamline and unify storage management across the enterprise by providing top-down management through a cloud-based management portal. The Microsoft Azure StorSimple Manager service *runs within Microsoft Azure* and communicates with all the storage devices in the enterprise, rolling up alerts, device status, and other information in order to provide a dashboard of the entire Microsoft Azure StorSimple environment. This facilitates global management of all of the storage resources in the enterprise: the arrays, all the data, and the policies that govern data protection and retention. An administrator can manage multiple sites from a centralized place (which of course is running in Microsoft Azure). If there’s a problem anywhere in the storage environment, the centralized management point will allow the administrator to see it.

In terms of managing incessant data growth, Microsoft Azure StorSimple helps enterprises to reduce their storage capacity consumption using inline data reduction (inline data deduplication *and* compression). And when the amount of deduped, compressed primary data exceeds its limit, the least recently used data is automatically and transparently migrated to the more cost-effective Azure Storage. Furthermore, data protection copies will invariably consume cloud storage capacity instead of data center capacity. *Cloud snapshots* are available to replace redundant, costly and complicated traditional data protection models and to enable backup, snapshot, archive, and offsite data protection to all be consolidated into a single, policy-driven repository.

Hybrid Storage Is Enhanced and Performance Is Increased

The local storage component within the Microsoft Azure StorSimple solution is an on-premises iSCSI hybrid storage array with a back-end that accesses objects in Azure Storage services. Two versions of StorSimple 8000 series arrays were introduced as part of Microsoft Azure StorSimple solutions:

- StorSimple 8100, which has 15 TB of usable, pre-dedupe local capacity. This extends to 200 TB with the integrated cloud capacity.
- StorSimple 8600, which has 40 TB of usable, pre-dedupe local capacity, which reaches 500 TB with its integrated cloud capacity.

The actual operating local capacity of these arrays will of course depend on the deduplication and compression ratios that are attained. Both StorSimple 8000 series arrays are fully featured iSCSI SAN arrays with the same components and capabilities you'd expect from any enterprise hybrid storage system, including redundant controllers, power, fans and network ports, SSDs, SAS drives, automated tiering, thin provisioning, application-consistent snapshots, and support for Windows, Linux, and both VMware and Hyper-V environments.

Microsoft has given the StorSimple 8000 series arrays a boost to make them larger and faster than its previous generation StorSimple products. Although Microsoft is not yet promoting detailed performance specifications for these arrays, it expects them to compete favorably in terms of IOPS for tier-2 and -3 workloads (which, as already mentioned, include a great deal of enterprise data, but exclude high-performance transactional database applications).³

A New Virtual Appliance That Runs as an Azure Service

Within its Microsoft Azure StorSimple offering, the company is also providing the Microsoft Azure StorSimple Virtual Appliance, a software-only version of the StorSimple array that runs in Azure as a service. This software-only version will allow the rehydration of, and access to, data in the cloud. Rehydration refers to the ability to recreate the data in its original, un-deduped form. Put very simply, if users are going to move their data to the cloud, then they ought to be able to use it there! The Virtual Appliance is intended to be used for DR in the cloud, development/test purposes, data analysis and discovery, and to help customers migrate on-premises applications to Azure.

Notes on Archiving and Disaster Recovery (DR)

Archiving and DR are two common IT responsibilities for which Microsoft Azure StorSimple is perfect. While most everyone would acknowledge these as necessities for their organizations, often they are neglected because they require a larger investment of time and resources than the IT team can make available. Both archiving and DR have similar foundational demands:

- Reliable, inexpensive, long-term retention of high volumes of historical data; and
- Easy management of the process, and rapid access to data when it is needed.

Cloud storage per se is obviously inherently good for reliable, inexpensive data retention but it is the capabilities of Microsoft Azure StorSimple that add the functional richness to deliver easy management and rapid access as well. Storage administrators set up StorSimple arrays to automatically run daily (for example) cloud snapshots that can serve for both backup and archive purposes. When data in the cloud is needed it is either accessed directly from the current volume view (for online, inactive data), or restored by mounting a cloud snapshot view (for historical archive and DR data). Where archiving and DR are concerned, the operational aspects are essentially the same; access to data in the cloud uses a synthetic volume view of the data as it was at a specific point in time and access to it is location-independent, so it can be done from either any site with a StorSimple array or in Azure using the

³ This paper is being written while Beta testing is still in progress: however, ESG Lab will be conducting testing (currently scheduled for the latter part of 2014) to evaluate the performance and operational value of these systems in detail.

Virtual Appliance. IT organizations deploying Microsoft Azure StorSimple can avoid spending money on a custom recovery facility, because Azure can be a recovery facility. The approach of mounting a synthetic volume view bypasses one of the common concerns about using “plain” cloud storage for DR for a full recovery (i.e. full data download); the concern is that such an operation could overwhelm the throughput capabilities of both the system and the network. The *instant recovery* capability of the Microsoft Azure StorSimple solution addresses this concern as it only restores the working datasets that applications need, while leaving the data that is not immediately needed in the cloud: this saves time and money, yet simultaneously allows much more aggressive RTOs to be achieved.

Market Relevance and Analysis

Many current approaches to cloud storage amount to little more than trusting your data to a cloud vendor and hoping it will be there later. Microsoft Azure StorSimple takes a very different approach. Microsoft has integrated management services that are top-down from the cloud to all the storage devices in an enterprise. What this means is that you’re not just “parking” data in the cloud, and then retrieving it from the “valet” when you need it back (hoping, of course, they are not busy with someone else at that moment....and that they didn’t damage anything... and didn’t lose the keys!).

With Microsoft Azure StorSimple, when you move data into the cloud, you can actually manage it there, and use it there.

Hybrid storage systems that integrate the cloud can essentially be viewed as bridging technologies that tie the present to the future. They represent a way to unite the best of traditional (i.e., data centers and storage arrays) and emerging (i.e., cloud, service-oriented approaches) worlds, providing the economic attraction and flexibility of the latter while retaining the control, familiarity, and performance of the former. The technical foundation of Microsoft Azure StorSimple is the erstwhile StorSimple CiS element that spans the gap.

- For storage users who want to stay with their legacy arrays or keep existing processes in place, using the cloud as a storage tier behind what appears to be a conventional array gives them the extensibility and price points of cloud without having to radically alter processes or retrain staff.
- For users who need a more comprehensive DR strategy but can’t afford a remote site—or the system and staff to manage it—Microsoft Azure StorSimple provides a more affordable, far simpler, low risk, and yet (at least) equally capable alternative. They have merely to make a few clicks within their local storage array management GUI to automatically copy their data to the cloud in encrypted form.

Many new approaches and technologies initially attract their share of naysayers. This is understandable—at their time of introduction, some technologies are not fully baked. Think of iSCSI, SATA, servers, perhaps Microsoft Windows (remember the “blue screen of death?”)...and even the cloud itself! But when they come from the right corporation, one with real determination and sufficient resources to succeed, products that are initially a little raw on the inside get “done” over time and become standards. There have been many cases of a technology that seems so good that, even though it comes from a tiny start-up, users decide to gamble on it, then cross their fingers that the start-up doesn’t go bust or get purchased by the wrong bigger company. Obviously that’s not the case here. That this Microsoft solution includes Azure as its public cloud platform will make most prospective users feel secure in their decision because Microsoft is (viewed as) a stable and credible provider with the sort of critical mass and commitment necessary to offer the levels of quality, investment, and longevity that are crucial to selecting a cloud provider. Extending the idea of hybrid storage to the cloud might be seen as a big step by some, but it is also one that shows the likely direction of enterprise storage infrastructure.

The Bigger Truth

For decades, enterprise storage has been managed via what can best be described as a “least worst” series of options. But you can only whack so many moles before surrendering to the futility of it all. Hopefully the various prevailing knee-jerk norms for managing storage will be viewed as aberrations in the not-too-distant future. After all, today’s science fiction is tomorrow’s reality. How many ideas that were once only dreamt about are now everyday realities in our work world? Video conferencing, IP telephony, smartphones, hot desking, and even working from home have become so ingrained in the way we get work done that we have difficulty remembering the time when they didn’t exist.

Of course, as more full hybrid storage offerings appear that genuinely embrace and integrate the cloud, enterprise storage managers will continue to look for capabilities they already need and use, such as thin provisioning, snapshots, remote replication, data reduction (which also minimizes network traffic), and high availability. Consequently, Microsoft is to be applauded for aiming at a comprehensive functional suite for its Microsoft Azure StorSimple offering.

In today’s IT world, growing data capacities, rampant virtualization, and increasing globalization are all making storage management dramatically more challenging. Users of this new Microsoft offering will gain comfort from knowing that the foundational cloud that is being used and integrated into the overall solution is Microsoft Azure, with its plentiful skilled professionals and capabilities. Equally, however, adopters of Microsoft Azure StorSimple will also be relieved to know that they do not have to relinquish the ultimate control of their data, just the daily grind of reacting to unplanned storage problems. This is all a very different proposition than abdicating responsibility to a service “out there somewhere in the cloud.”

With Microsoft Azure StorSimple, the cloud moves from playing a passive role in enterprise storage to a highly active role. Now, users can extend the capabilities of their storage and systems into the cloud, which becomes a true extension of their infrastructure.

The bottom line is that, with Microsoft Azure StorSimple, Microsoft is taking genuine hybrid storage to a new level. The solution can be viewed as the next (and a most welcome) step on the road to storage as an application rather than storage as a problem.



Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel: 508.482.0188 Fax: 508.482.0218 | www.esg-global.com