White Paper

Evergreen Storage: A New Approach

By Scott Sinclair, Storage Analyst

March 2016

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

Contents

Introduction

For decades, the limitations of spinning disks have held back data center efficiency and constricted architecture decision-making. These performance bottlenecks have impacted not only storage infrastructure decisions, but also decisions involving application deployment, server, and networking configurations, and power and cooling. To meet the performance requirements of business-critical workloads, IT administrators often employed what could be considered today as unnatural acts to improve storage performance. For example, disk striping and the various RAID configurations, such as RAID 10 and 50, that leverage disk striping were designed to attempt to aggregate multiple spinning drives to increase the collective performance. It was also common for organizations to deploy far more capacity than was necessary or to significantly reduce the accessible capacity of a drive via short stroking all to increase performance. These techniques were not only common, but also best practices recommended by application vendors. Thankfully, those days are fading into the sunset as the dawn of a new era of storage rises with the emergence of solid-state storage solutions, such as the all-flash array.

Even as more organizations adopt solid-state storage and integrate it into their environments, however, the remnants of old storage architectural design paradigms exist. One example, the scale-out storage architecture, enables a storage pool to scale performance and capacity simultaneously by adding incremental storage nodes that offer a combination of processing, memory, and disk capacity. This model provided simplicity to spinning disk environments, where incremental capacity was often required to increase performance. But in the era of solid-state storage, this model often introduces inefficiencies. Additionally, though not as antiquated as short stroking, many of the benefits of scale-out storage designs offered to spinning disk media can be achieved by leveraging alternative means.

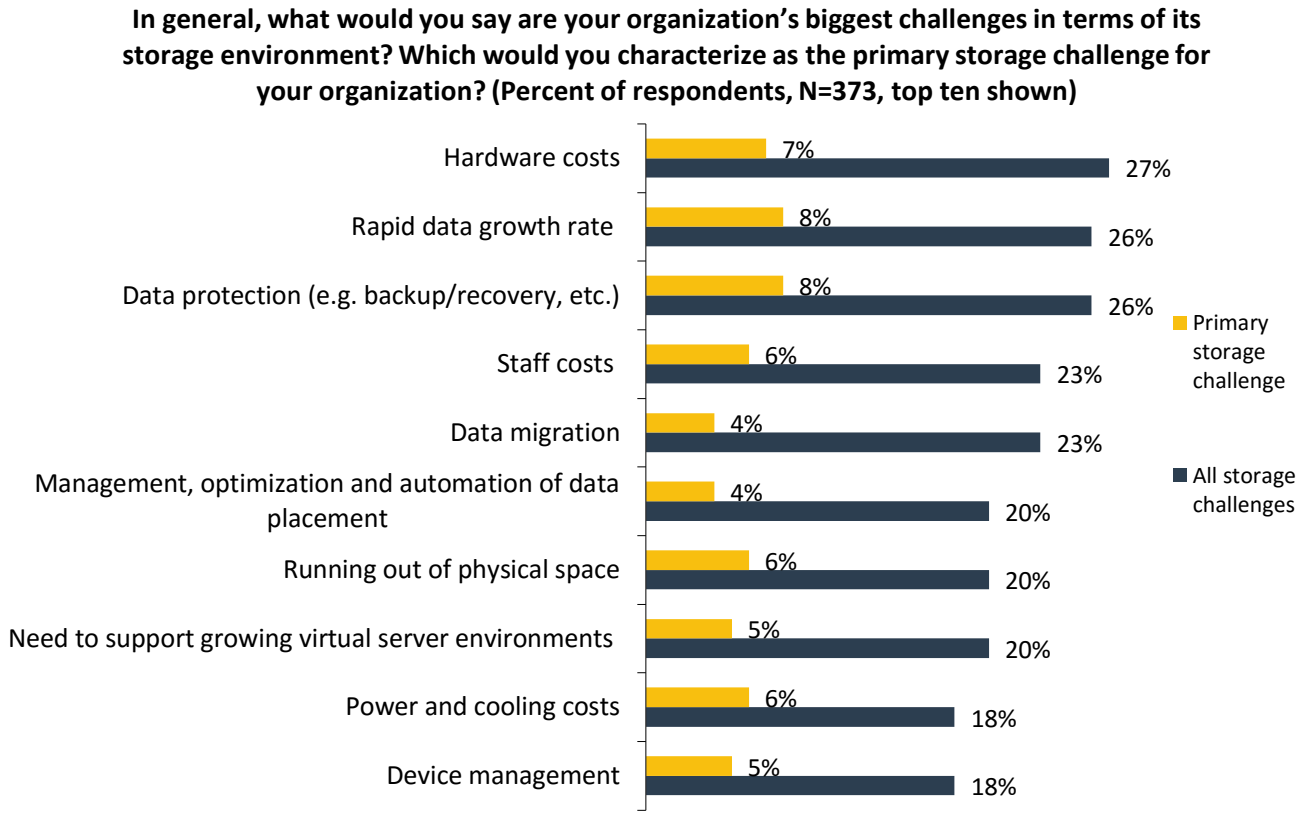
One all-flash array provider, Pure Storage, recently introduced its FlashArray//m array with a modular design. This announcement is coupled with a new program, Evergreen Storage, which is designed to extend the life of the storage deployment to more than ten years. Pure Storage seeks to accomplish this feat by enabling IT organizations to augment existing FlashArray//m deployments with upgraded processing, memory, and solid-state storage as new innovations emerge, all while the data is online and available. The net result is intended to allow existing configurations to scale performance and capacity over time as more advanced and less expensive technologies emerge. For many organizations, the combination of Pure Storage's architecture and Evergreen Storage provides the benefits that scale-out provided to spinning hard-drive environments in an efficient manner, without locking capacity and performance scaling together. The ultimate result introduces essentially a third storage option, Evergreen storage, beyond just scale–up and scale-out.

Data Growth and the Challenges of Enterprise Storage Environments

ESG recently conducted a research study investigating general storage industry trends. As part of that study, ESG surveyed 373 IT decision makers responsible for their organization's data storage environments. During part of the study, respondents were asked to identify their organization's biggest storage challenges. The results, shown in Figure 1, found that organizations continue to struggle with the challenges of rapid data growth.[1] Additionally, when considering the remainder of the responses that comprise the top ten, it appears that the impacts of data growth are being felt across the data center: Challenges with hardware costs, data protection, staffing costs, and data migration, along with the others identified, can all be considered symptoms of data growth.

[1] Source: ESG Research Study, *2015 Data Storage Market Trends*, October 2015. All ESG research references and charts in this white paper have been taken from this research report.

Figure 1. Organizations' Top Ten Biggest Challenges in Terms of Storage Environment



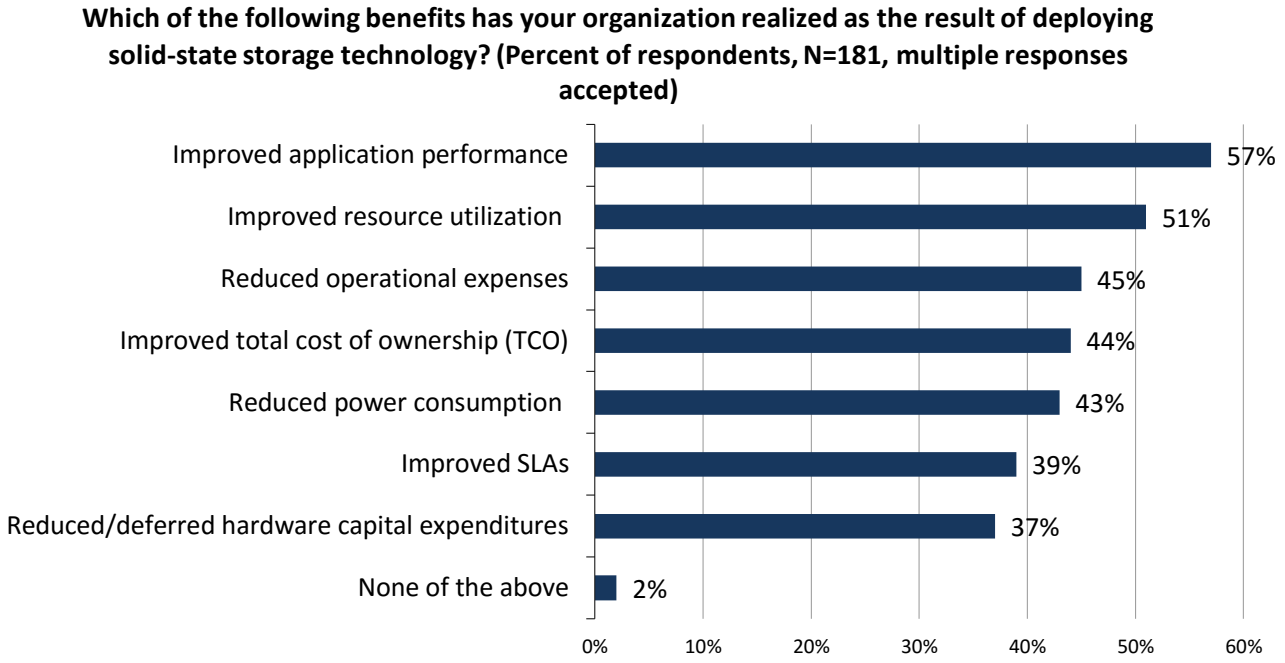
Source: Enterprise Strategy Group, 2016.

This list of challenges—to a large extent—illustrates IT organizations' basic need for more efficient data storage infrastructure. Data growth will continue; therefore, it is the storage architectures that must evolve to ease the effects of that growth. This efficiency represents a large portion of the initial promise of solid-state technology: the ability to serve transactional workloads with storage infrastructure tailored to the needs of the organization. As a result of this new low-latency brand of storage, organizations are experiencing a number of benefits that expand beyond simple performance gains.

Solid-state and Evergreen Storage Alters the Scale-out Storage Paradigm

Over the past few years, organizations have adopted the various forms of solid-state storage at a rapid pace. As the price continues to decline, it is easy to see how the rate of deployment will increase. After only these few initial years of wider adoption, organizations are already reaping benefits, which are not limited to performance. As mentioned previously, ESG conducted a storage industry research study in 2015 in which ESG investigated multiple storage technologies, including solid-state storage. As part of that study, current solid-state users were asked to identify the benefits their organizations have realized from the technology. While the most often identified response to this inquiry, improved application performance (57%), will not likely surprise anyone, the interesting responses were the ones that followed: The second, third, and fourth most-cited responses were improved resource utilization (51%), reduced operational expenses (45%), and improved total cost of ownership (TCO) (44%) (see Figure 2).

Figure 2. Benefits Realized as a Result of Deploying Solid-state Storage



Source: Enterprise Strategy Group, 2016

Eliminating the bottleneck created by spinning disk media allows solid-state storage to set in motion a cascading series of benefits for the data center, which translates not only to faster performance, but also an ability for other aspects of the infrastructure to simply do more. Each server can serve more applications. Application licenses can serve more data. Each data center investment can better achieve its potential. Though often procured to solve a particular performance need, the resulting impact on the data center can be far greater. Many of these benefits are in part the result of the efficient manner in which solid-state storage provides performance. As an electrical component—rather than a mechanical one—performance is an innate benefit and not the result of massive scaling. Even a small capacity of solid-state storage can improve performance, minimizing the data center real estate, and power and cooling footprint. This ability to deliver efficient performance improvement translates into a new way of deciding how to deploy storage capacity and performance.

Evergreen Storage: The Third Option beyond Scale–up and Scale-out

For spinning disk media, scale-out storage architectures emerged to solve a number of specific data storage challenges. The direct correlation between the number of spindles and disk performance offered an opportunity to simplify storage scaling. Organizations could expand a storage pool by adding processing, memory, and disk capacity combined as a single node. When performance demands increased, the existing storage pool could expand to meet those demands. Since it was common to deploy excess capacity to meet performance needs, deploying more capacity than necessary was often seen as standard operating procedure. Additionally, when the storage pool needed to expand for capacity reasons, scale-out storage nodes could be architected to only offer the right level of processing and memory for the performance the spinning disks could provide. With solid-state storage, however, increasing capacity to meet the performance demands of most applications is not necessary. And when applications demand increased capacity, the solution often does not need to be burdened by the expense of adding more performance. It is these distinctions that place limits on the benefits that can be derived from scale-out solid-state storage architectures for today's typical IT data center environments.

For spinning disk-based storage environments, simple scalability is not the only benefit associated with scale-out storage systems. In large capacity storage environments, scale-out storage offerings, such as object storage or scale-out file systems, can also allow organizations to mix and match hardware generations in a single storage pool to ease the integration of new hardware technologies (e.g., faster processing, faster memory, or lower cost storage).

Also, it is common for these object and scale-out file systems solutions to achieve greater resiliency by spreading the data across multiple storage nodes, which allows for data to remain accessible in the event of a node failure. While these capabilities are common in massive capacity scale-out storage solutions, it is important to note that these features are not typically present in scale-out solid-state storage solutions. As such, it is not safe to assume that all the benefits associated with scale-out storage are actually available in scale-out solid-state solutions.

Pure Storage recognizes this technology trend, and has designed its modular FlashArray//m to deliver the benefits typically associated with the promise of scale-out storage, even though the current reality of most scale-out storage solutions may not live up to those expectations. Pure Storage has combined its FlashArray//m architecture with its Evergreen Storage program to offer benefits such as scalability and the simple integration of new or lower cost hardware technologies in an efficient package. With Evergreen and the FlashArray//m, upgrading to a higher-performing controller option allows IT organizations to expand performance by scaling across hardware generations. Additionally according to Pure Storage, upgrades are performed while maintaining 100% performance and data availability, reducing the hurdles often incurred with upgrading controllers. Also, with Evergreen Storage, Pure Storage has announced that IT organizations receive a free controller upgrade upon the three-year maintenance renewal after the third year

The net result is that Pure Storage's Evergreen Storage falls into an alternate (third) storage category, beyond scale-out or scale-up. Scale-out storage locks performance and compute together, offering scale only through incremental nodes, which often creates inefficiencies. Scale-up storage solutions, even those with all-flash technology, can scale capacity separately, but the performance can still remain limited behind the capabilities of two controllers. Pure Storage's Evergreen Storage resides in a separate category, designed to augment only the capability required, only when it is required, without disruption.

Pure Storage FlashArray//m and Evergreen Storage

When deploying an all-flash array—or any storage array—IT storage decision makers consider a number of aspects. Can the solution provide the performance and capacity to meet my application needs today? What about in three years? What if my needs change? How does the solution impact the bottom line? Despite the critical importance of these questions, too often technology evaluations over-emphasize specific features, rather than focusing on the net benefits those features provide. Ultimately, whether a storage system is scale-up, scale-out, or something else entirely doesn't matter. The net benefits that the storage system provides to the organizations are what truly matter. The following section will review several—but not all—of these specific technology benefits and discuss how Pure Storage can deliver these capabilities to many organizations in a manner equivalent to, or even superior to, that which is possible with scale-out solid-state storage.

- **Low-latency performance**: While low-latency and high input/output request per second (IOPS) performance often go hand in hand, they offer different advantages. The latency provided by a storage system indicates how quickly a particular input/output request (I/O) is serviced, while IOPS refers to how many of those I/Os can be serviced in a second. Obviously, if an I/O can be serviced in a shorter amount of time, then more I/Os can be serviced in a second. For a large percentage of workload environments, especially those in the mid-range, latency is the more critical of the two because it impacts the application regardless of how many I/Os the application actually demands. It is, however, important to ensure that the solution provides enough IOPS to support any application demand spikes and future demand growth expectations. All-flash storage systems, including those from Pure Storage, provide low-latency performance that enables substantial levels of IOPS. While some scale-out solid-state storage solutions can achieve a higher aggregate IOPS number, if the organization does not need that level of IOPS, the overall capital investment for these scale-out solutions may not translate into the best results for the business.
- **Efficient scaling:** As storage demands scale over time, systems that can most efficiently adjust and scale to changing workload demands will likely offer superior business results. As mentioned previously, the emergence of solid-state storage, such as all-flash arrays, has allowed for the separation of application performance from capacity—to a large extent. No longer requiring organizations to add more capacity than necessary to meet the performance demands of transactional workloads, all-flash arrays can offer more

efficient storage infrastructure usage. Performance or capacity can be added or expanded when necessary, rather than being tied to each other in lock step. This efficient capacity delivery also applies to initial deployments, when IT organizations often do not want to pay higher infrastructure costs than they need at that time.

Scale-out storage solutions, on the other hand, often add in the cost of additional controllers with capacity expansion regardless of whether the performance of those controllers is actually required. Scale–up solutions can add capacity, but typically performance cannot be expanded non-disruptively. The Pure Storage FlashArray//m modular design allows for new storage controllers, and more performance, to be integrated into the system non-disruptively, so organizations can take advantage of future advances in processing and memory without having to deploy new capacity. With the modular design of FlashArray//m, the underlying data can remain in place and active.

Additionally, although storage capacity is often measured by the amount of raw storage residing on the system, what the application sees and has access to is more important. To this point, deduplication and other data reduction techniques can become paramount. The benefits of Pure Storage's data reduction technology are well documented and are a key component in the drive to make all-flash storage cost-effective and ultimately more feasible for a larger percentage of workloads.

- **Resiliency:** Enterprise storage systems are defined, in part, by their ability to offer high-availability, often delivering data resiliency in excess of five 9s. As mentioned previously, the improved resiliency achieved by scale-out storage, object storage, and scale-out file system environments is often not available in scale-out solid-state storage deployments, specifically in the form of the ability to support a node failure. In other words, if both controllers in a node fail, the data may not still be available or the storage pool may not still be active. In some scale-out solid-state storage environments, not only will the loss of both controllers in a single node failure make the data on that node unavailable, but it can also bring down the entire storage pool. In these cases, the scale-out nature of the solution increases the failure domain and can actually make the system as a whole less resilient. While dual controller failures typically have a low likelihood of occurrence, each additional node increases the chance of a node failure, which in turn increases the chance of the entire pool going down. In these cases, a single dual controller storage deployment, such as Pure Storage's FlashArray, can be more resilient. In the extremely rare event of a dual controller failure, only the data tied to that node becomes unavailable, reducing the failure domain and minimizing the impact to the data center.
- **Future-proof and performance scaling:** Innovation in storage—especially flash storage—technology, continues to evolve and mature. At the same time, application and data accessibility requirements persist and often grow. Storage architectures that can more easily integrate new technologies allow for organizations to gain access to the benefits of advances in processing, memory, and solid-state storage more quickly and easily. In turn, the reduced disruption can translate to a significant reduction in management and operating costs. A significant percentage of storage solutions, both scale-up and scale-out, require forklift upgrades. This requirement often results in the purchase of new hardware and software to replace the existing solution, while requiring that all of the data also be migrated to that new solution.
- Solving this challenge is part of the idea behind Pure Storage's modular FlashArray//m architecture and Evergreen Storage program. With Evergreen Storage, storage architecture can scale through hardware generations without a data migration. Pure Storage intends for IT organizations to deploy a Pure Storage FlashArray//m once and then upgrade the features and the hardware capabilities, such as performance or capacity, independently. Pure Storage claims that this modular capability allows the expected life span of each deployment to exceed ten years. If a solution requires more performance, more memory, or more capacity, the solution can simply be upgraded, adding only what is needed. With Evergreen Storage, Pure Storage is able to provide a single pool of storage that spans hardware generations, where organizations receive a free controller upgrade upon the three-year maintenance renewal. The net result can significantly reduce the total cost of ownership over the life of the solution.

Despite the inefficiencies ingrained in scale-out solid-state storage architectures, some environments may benefit from the design. While solid-state does not require scale-out to provide the IOPS necessary for transactional workloads, the massive IOPS performance levels achievable by scale-out solid-state storage can present opportunities for very large environments. That being said, existing scale-out systems can introduce limitations. For example, scale-out systems often increase the failure domain by placing more data at risk in the event of a storage node failure. This added risk is in addition to the scaling inefficiencies mentioned previously that arise by tying capacity and performance scaling together. Finally, it is important to consider how any storage solution will handle a hardware generation upgrade. If not supported, scale-out solutions simply increase the pool of data that must be migrated when the next generation of hardware is released.

The Bigger Truth

As the IT industry continues its transition to solid-state storage, organizations are starting to better understand the full extent of the technology's benefits. Though often initially procured to provide a needed boost to performance-hungry applications, many adopters of solid-state storage are realizing benefits that cascade throughout the data center. As experience with and understanding of solid-state storage grows, like any new technology, many of the preconceived notions derived from older technologies must be cast aside. To a large extent, this is the case with scale-out architectures. For many environments, tying capacity and performance expansion together can be considered a relic of the "rotating spindles" era. While some environments can benefit from the massive levels of IOPS scaling these solutions provide, an architecture that limits deployment flexibility and increases the failure domain can and likely will generate inefficiencies as workload demands grow and evolve over time. And those inefficiencies lead to increased costs. When evaluating storage solutions, IT organizations are often best served by first assessing their needs, and then evaluating multiple products to determine which product and architecture best fits those needs. Injecting architectural bias towards one model (scale-out, for example) into the storage decision process can result in a missed opportunity.

Pure Storage, with its Evergreen Storage program and its modular all-flash array design, understands that not only will the needs of the data center change, but so too will the flash technology that these storage solutions are built upon. As organizations seek to become more efficient and leverage new innovations in processing, memory, and solid-state storage, Pure Storage allows these innovations to be deployed while data remains in place and available. The net result can be increased flexibility and simplicity as the solution can augment only the capability that is needed, when it is needed. In the end, what matters is the value that a storage architecture—whether it's a spinning disk solution or an all-flash array—offers to the IT organization and to the business as a whole. When looking to an all-flash array, organizations would be best served to ensure that it provides the right level of performance, flexibility, and cost-effectiveness for their particular environment, for today and for the future.

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