White Paper

Architectural Design Decisions Directly Impact IT Agility for Pure Storage FlashArray and FlashBlade Users

Sponsored by: Pure Storage

Dave Pearson
October 2023

IDC OPINION

All-flash arrays (AFAs) have become an increasingly popular choice across the enterprise storage ecosystem for many high-performance and business-critical application and workload storage needs. AFAs have certain key benefits versus traditional spinning media (hard disk drive [HDD]-based arrays) – higher performance, reduced operational costs, and greater environmental sustainability are some of the chief reasons buyers look to flash rather than spinning disk media.

Early thinking about AFAs – "higher performance at a higher price" – is an oversimplification that buyers cannot afford to make. Core AFA architecture design choices have significant impact on the usability and long-term effectiveness. Not all AFAs are created equally – from the types of flash media within the AFA to the form factor to the hardware and software managing data on the AFA. The selection of a storage vendor can have direct and long-term ramifications on how a company's resources are utilized, from the initial AFA deployment to the operational life span of the AFA – a period that can exceed 3, 5, or even 10 years. The usability and manageability of an AFA is also critical, as well as the architecture's impact on upgradability and sustainability.

Taking a disciplined and prescriptive approach, so there is a clear understanding of the underlying architecture, helps with the selection of the correct storage infrastructure, across the enterprise and on a workload-by-workload basis. This approach must not only solve for the problems of today but also provide information technology (IT) capabilities and agility that organizations need for the longer term.

The challenges IT executives face are complex:

- How can their IT teams meet application performance and availability SLOs today as well as be prepared for future data needs and growth?
- Will the underlying architecture meet future storage requirements and provide an easy, nondisruptive upgrade path with minimal impact to operations, financial resources, and IT teams?
- What can they do to achieve energy consumption and real estate–related efficiency goals?
- Can the architecture reduce the burden of maintenance and management tasks so that IT teams can focus on higher-value, strategic IT projects that provide tangible benefits to the organization?

These challenges are only magnified in difficult macroeconomic environments as literally every project, resource, and kilowatt consumed is under financial scrutiny while trying to satisfy governance, compliance, and ESG standards. Legacy approaches to storage deployments, while familiar to IT teams, can greatly limit IT agility not only through technical constraints but also through disruptive procurement, installation, operation, and upgrade processes. Legacy storage architectures can require "forklift" (remove and replace) upgrades that are complete overhauls of the IT infrastructure: End-of-life or outdated storage systems are ripped out and replaced with newly purchased technology, requiring data migration between old and new systems. The process is complex, disruptive to workload and application access, and extremely IT intensive. Forklift upgrades are driven by legacy architectural incompatibilities between storage generations or limitations in scale-out architectures that can't add new nodes to existing arrays or clusters.

This white paper takes a closer look at Pure's alternative to the forklift upgrade, made possible by the vendor's Evergreen architecture. We will primarily focus on Pure's scale-up block and file product FlashArray; however, we will also point out where the shared design philosophy extends to its scale-out, unstructured data FlashBlade product family. We will delve into four very specific areas: software-driven architecture, modularity of components, simplicity of design, (which are common to both products) and stateless controllers, which is a key enabler of FlashArray's nondisruptive upgrade capabilities. We will also highlight some of the unique upgrade capabilities in FlashBlade//S; however, we have gone into much more detail on its architecture in *Pure Storage's Next-Generation FlashBlade//S Delivers a Huge Leap Forward for Unstructured Data Storage* (IDC #US49102422, May 2022), and we explore its nondisruptive upgrade capabilities in *Pure Storage's Evergreen//Forever Subscription Ushers in New FlashBlade//S as a Premier Platform* (IDC #US49103522, June 2022). Real customer benefits — whether in terms of better performance, higher availability, easier deployment and management, improved efficiencies, nondisruptive multigenerational technology refresh, and/or lower costs — result from these architectural choices. These benefits are also a key factor in Pure Storage's industry-leading Net Promoter Score, an objective third-party metric that measures the quality of the overall customer experience (CX).

The typical legacy enterprise storage life cycle for a storage array is 4-5 years, after which a need to move to newer technologies commonly drives an often-painful forklift upgrade. Pure Storage aims to double the typical enterprise storage life cycle for its products to 8–10 years and beyond, simplifying the enterprise storage life-cycle experience and lowering the total cost of ownership. The flexibility to accommodate newer technologies nondisruptively is a key determinant of array longevity and IT agility, and Pure Storage's architectural design decisions reflect that goal. As enterprises consider new storage platforms, IT buyers should seek a modernized version of the enterprise storage life cycle and should expect vendors to validate their ability to deliver it.

IN THIS WHITE PAPER

We will be discussing the following points in this white paper:

- Legacy storage upgrade life cycle
- An overview of Pure's Evergreen architecture
  - Software-driven architecture
  - Stateless controllers
  - DirectFlash storage design
  - Modularity of components
  - Simplicity of design
  - The impact of FlashArray life-cycle extension

- AFA challenges and opportunities
- Architecture questions to ask and what to look for
- Conclusion

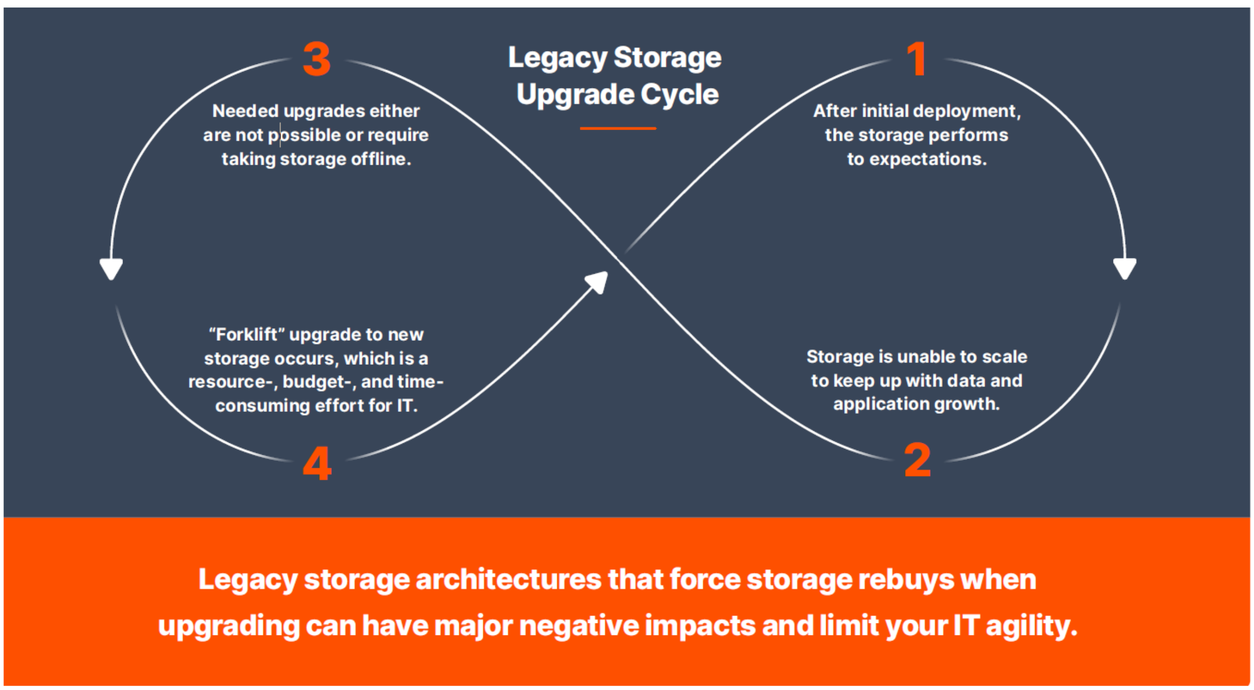
This white paper summarizes Pure's architectural decisions, discussing implementation as well as the resulting impacts on CX. In addition, the paper explains the reasons for the architectural choice and summarizes the relevant impact on performance, availability, efficiency, upgradability, and/or cost. Four families of FlashArray products are currently available (//E, //C, //X, and //XL, as well as two families of FlashBlade [//E and //S]). The capacity and performance of each differs to support a variety of structured and unstructured data workloads and use cases, but the architectural decisions around the systems as well as Pure's use of custom DirectFlash Modules (DFMs) remain consistent.

SITUATION OVERVIEW

There are typically four stages to upgrading legacy storage systems (see Figure 1).

FIGURE 1

Overview of Upgrade Cycle of Legacy Storage



Source: Pure Storage, 2023

At the point of initial deployment, legacy architectures are able to meet application and workload environments – though often accompanied by overprovisioning for future capacity needs, reducing efficiencies in both economics and utilization. Inevitably, as data and application growth increases, the storage is unable to keep up with the demands of the workloads, from both capacity and performance standpoints. This can result in a forklift upgrade, which is complex, wasteful, and time-consuming.

Complete overhauls of the IT infrastructure are required because of limited, rigid architectures that were not designed for upgradability — especially between storage product generations. With traditional storage architectures, this cycle can repeat indefinitely and undercuts the IT agility that organizations are demanding.

Architectural decisions can make a considerable difference in providing the performance, availability, efficiency, upgradability, and simplicity that effectively solve for the issues found in legacy storage systems. Pure Storage's FlashArray and FlashBlade are the result of a number of architectural decisions designed to allow these storage platforms to deliver a customer experience more in line with customer needs and expectations than the cycle described previously.

Information technology organizations that manage their own enterprise storage systems are very familiar with the legacy enterprise storage experience and life cycle. Over the years, storage managers have almost become inured to working with less flexible hardware-defined architectures, paying separately for various data services (snapshot integration, replication options, etc.) and managing systems through "nerd knobs" that require significant storage expertise. They have also endured maintenance price increases as a system ages, and then roughly every 4 or 5 years they are forced to go through a time-consuming and disruptive forklift upgrade process to move to newer technologies. Although not all of these aspects of enterprise storage ownership flow directly from system architecture decisions, many of them do. In a word, architecture matters.

Evergreen Architecture Overview

Pure's core design philosophy delivers a modular, software-driven stateless architecture that was built to be upgraded nondisruptively. Controllers, blades, flash, software, and even chassis are upgradeable to deliver performance needed to meet modern workloads and keep storage on the cutting edge, with zero impact to applications and near-zero effort for IT staff. With cross-generational compatibility built in, different flash geometries can be mixed within an array, and newer controllers can work seamlessly with older flash modules, increasing the life span and utilization of existing infrastructure. Pure has customer testimonials demonstrating these capabilities across eight hardware generations, over 10 years, and with more than 10,000 controller upgrades.

When the first enterprise-class AFAs began to ship in the 2011-2012 time frame, the use of solid state storage as a persistent storage medium set these systems apart from their legacy brethren that were still architected for (and used) hard disk drives. Some storage vendors just inserted NAND flash-based, general-purpose solid state disks (SSDs) into systems that had originally been designed around HDDs. Others made more sweeping architectural changes that made better, more efficient use of the new solid state storage medium. As one of the AFA start-ups that architected its systems from a "blank sheet of paper," Pure Storage made a number of design decisions intended to improve overall performance, availability, efficiency (of resource utilization), and simplicity that were directly linked to the decision to use only persistent flash media. One of the business objectives of Pure Storage's design intent was to at least double the useful life of an enterprise storage system, further enhancing the cost benefits associated with the more "flash optimized" design.

In considering Pure Storage's architectural decisions that have a direct impact on extending and simplifying the entire storage ownership and usage life cycle, IDC notes that the vendor created and marketed an entire ownership program, called Evergreen Subscriptions, to underline the advantages its design offers across a number of areas. The name "Evergreen" was chosen because the offering did at least double the useful life of a storage system while allowing that system to be upgraded nondisruptively to accommodate newer storage technologies as they became available — in other words, the system stayed "evergreen" throughout its entire life cycle. Many of Pure Storage's competitors have characterized Evergreen as "just a program," but a deeper look at the vendor's system architecture disputes that claim.

Software-Driven Architecture

In recent years, software-driven infrastructure has become a popular topic among vendors, particularly software-defined storage (SDS) vendors. Compared with older, more hardware-defined designs, software-defined designs put a storage system's functionality and features (beyond just persistence) into software. With software-defined functionality, system capabilities in the areas of performance, availability, capacity, and storage management can be upgraded via software. This functionality makes it not only simpler but also less expensive to add new features to an existing system in the field because it does not require the purchase of any new hardware. This does not mean that system enhancements do not come from new hardware — for example, buying additional storage devices adds capacity — but software-defined functionality provides greater flexibility to upgrade system capabilities without requiring additional hardware investment. And even in the case of additional capacity, software upgrades that improve data reduction ratios can add effective capacity to a system without requiring any hardware purchases.

Limited support for different storage device geometries can be a problem with some legacy systems, precluding them from supporting newer storage technologies that may deliver faster performance, increased density, or other attractive capabilities. When storage device support is in some way limited by the storage controllers, controller upgrades may be required to support newer device geometries. Some legacy systems do offer a "midlife kicker" that includes a "new" controller at some point in a system's life cycle that may introduce incremental improvements. But these limited upgrades are often not able to accommodate significant advancements like moving from SAS to NVMe or accommodating larger-capacity devices and/or new media types — in other words, generational improvements. In some systems, the only way to obtain a controller that can support desired new technologies is to perform a forklift system upgrade. When the software is the critical factor in storage device support (rather than the controller), there is significantly more flexibility to accommodate newer technologies.

Software-defined designs also enable more flexibility in handling metadata, or the system's data about your data. If the metadata is extensible, a system can accommodate on-disk format changes that improve the performance, endurance, and/or efficiency of operation gains without requiring any data migration. If the metadata is static and tied to the media format, then there is little opportunity to upgrade it to improve system capabilities without migrating data. Some storage administrators may remember cases not that long ago involving legacy storage products where they were forced to migrate all of the data off of an array to obtain needed new "software" features because the on-disk format was changed along with the new software release. Vendors are well aware of the pain this causes to customers, but in some cases, the need to extend the capabilities (usually the capacity scalability) of the system is deemed to outweigh the pain.

The ability to take advantage of new technologies that offer significantly improved capabilities through just a software upgrade is one aspect of extending the storage system life cycle, and Pure Storage has an excellent track record here. Since 2012, new software releases of Purity, the storage operating system used in both FlashArray (Purity//FA) and FlashBlade (Purity//FB), have improved performance, increased scalability through (among other things) support for new device geometries, accommodated new solid state media types (MLC, TLC, and QLC), increased data reduction ratios, and enabled key new features, like cloud-based predictive analytics and support for native file systems, replication, and stretch clusters. The array's metadata is extensible, and over the years, the company has continuously evolved the metadata to make its systems more performant, scalable, and efficient. All of these changes have been implemented in a manner that allowed the installed base of systems to upgrade nondisruptively, helping keep the systems "evergreen."

In the case of FlashArray, its ability to simultaneously support multiple device geometries offers what Pure Storage refers to as "Capacity Consolidation." For an older FlashArray, a customer may at some point want to move from lower-capacity to higher-capacity (i.e., more dense) storage devices. This is an evolution that often cannot be done in competitive SAN systems and, even if it can be, often requires manual data migration. On the FlashArray, a customer can insert a new disk shelf (assuming that space is available) of the newer storage devices, and the system will, over time, migrate data from the older to the newer devices as an automatic background process so as not to impact application performance. Once the data movement is complete, the administrator is notified, and the older shelf may be removed. In this way, a system that takes up the same amount of rack space can be expanded over time not only to higher-performance controllers (using the Evergreen//Forever and Evergreen//Flex subscriptions' Ever Agile and Ever Modern features) but also to increased storage density (using the Capacity Consolidation feature). If customers are consuming their Pure storage via the Evergreen//One storage as-a-service subscription, the same capabilities are employed by Pure to upgrade their storage service nondisruptively.

Stateless Controllers

Many legacy SAN arrays use a stateful controller design, an architecture that complicates controller replacements and upgrades. For example, if the memory that holds controller state information is held in the controller itself, removing a controller severely impacts array performance or even the ability to remain online. In the FlashArray, controller state information is kept in separate cache cards. When a controller is replaced, data on the NVMe-based cache cards is very rapidly reloaded onto the controller, and the array is moved back into normal operation mode. This stateless controller design is a key element in enabling FlashArray to be upgraded nondisruptively via controller swap out. The fact that state information is kept separate from the controller also provides better data security than stateful controller designs.

FlashArray also uses a modified "active/passive" controller architecture where, during normal operation, both controllers accept I/O from the initiator side. The two controllers coordinate their operations such that data is moved to and from storage devices in a system through only a single controller. If a controller fails, the remaining controller continues to service all I/O with no change in performance. Storage administrators familiar with "active/active" controller designs may observe that having both controllers active during normal operation drives higher performance, and this may be true, depending on the power of the respective controllers being compared. Readers should note that in building each class of storage controller, Pure Storage has designed each single controller to be able to handle the same amount of I/O as its competitors' dual controller designs (when those competing systems are operating normally). In other words, this design decision means that Pure Storage engineered around the trade-offs and can provide the benefits of the "active/passive" design (e.g., a failed controller causes no performance impact) without the downsides (a system runs only half as fast as during normal operation). It also makes possible the nondisruptive controller upgrade process.

FlashArray controllers can be upgraded at any time during the life of an array to move to more powerful, latest-generation controllers within the same class or to move to a higher, more powerful class of system, also of the latest generation. (The "upgrade at any time" option under the Evergreen//Forever and Evergreen//Flex subscriptions, which includes full trade-in credit for old controllers, is called Ever Agile. These subscriptions also include a feature for no-cost controller upgrades to the latest generation of the same class of controller, called Ever Modern.) For more information and IDC commentary on the Evergreen//Forever subscription, see *Evergreen//Forever Subscription Continues to Drive Industry-Leading Customer Experience as a Differentiator for Pure Storage* (IDC #US49553922, August 2022).

Pure Storage's approach to controller design in FlashArray has some inherent advantages relative to "active/active" designs.

First, upgrading a system to new controllers does not impact application services. Controllers are removed and replaced one at a time during the upgrade process, in essence creating a planned "failover" that the system handles automatically, without disruption due to the stateless, active/passive design. This is true both for upgrading to the latest generation of the currently installed controllers and for scaling up to a higher class of controller. There are scenarios in which customers require more performance, not to mention higher capacity potential, from their arrays as workloads increase and requirements change over the life of an array. If that need arises, customers have the option to nondisruptively upgrade controllers to get to a higher performance level (i.e., move from an //X20 to an //X50 or to an //X70, etc.) without a forklift upgrade, data migration, or impacting application performance. Readers should note that this upgrade capability leverages not only the stateless, active/passive controller design but also Pure's modularity of components (discussed in the Modularity of Components section), both of which are featured in FlashArray.

Second, this capability has a positive impact on the reliability of the controller pair. During normal operation, only one of the controllers is communicating with the back-end storage devices, not both. Because systems are operating normally instead of in failure mode most of the time, there is less overall stress on the controller pair (which combined run at a lower level of utilization) than with competitive "active/active" controllers in the same class (which are both handling initiator and target traffic during normal operation). Running components at lower levels of utilization has a positive effect on reliability.

DirectFlash Storage Design

Pure Storage uses a technology called DirectFlash, which combines the Purity storage operating system software with custom-built storage devices, called DirectFlash Modules. In 2017, Pure Storage moved away from commodity off-the-shelf SSDs to the new custom DFMs. All along, the Purity storage operating system software had used proprietary algorithms to manage the flash media for improved performance, higher endurance, and better capacity utilization. But the company could not always get the access it wanted to the underlying media when working with off-the-shelf SSDs. The SSD controller and flash translation layer provided by the storage component suppliers were often "in the way." Moving to the custom DirectFlash software and hardware technology provides the ability for the Purity software to manage the underlying flash media directly (hence the "Direct" name), driving performance and efficiency improvements. The custom form factor has allowed Pure Storage to introduce higher-capacity devices for increased storage density than has been generally available with off-the-shelf SSDs.

When FlashBlade was introduced in 2017, it made use of the same DirectFlash technology initially integrated directly into its blade design. Starting with FlashBlade//S in 2022, it began using the same DFMs as does FlashArray.

Today, the largest enterprise-class SSD in common use is around 30TB, while Pure Storage has released 75TB QLC DFMs for dense, high-capacity environments and higher-performance 36TB TLC DFMs. In all systems, DirectFlash Modules are hot pluggable. Pure also claims that with its direct-to-flash management, it can take advantage of these higher densities without trading off performance or reliability, which is often the case with SSDs.

The ability to manage the underlying solid state media directly in a media-agnostic manner is important for several reasons. First, it allows Pure Storage to implement I/O algorithms that are best suited for efficient use of the underlying media, given the I/O patterns in its target markets and associated use cases. Pure Storage is not forced to work through SSD-resident controllers and flash translation layers that tend to be more attuned to the high-volume consumer rather than enterprise market requirements. This has implications not only for latencies but also for how garbage collection (i.e., free space allocation) and media endurance are managed — issues that are particularly important as Pure Storage utilizes NAND flash–based media (TLC and QLC) that have lower performance and endurance than SLC. Second, by optimizing media endurance, the company can lower media overprovisioning ratios to cut costs out of the storage devices without sacrificing device-level longevity (the vendor also provides a lifetime-of-the-system guarantee on all solid state devices it sells as part of its Evergreen subscriptions). Last, the ability to manage underlying solid state media allows Pure Storage to more easily accommodate new media types as it builds them into the DFMs and qualifies them for production use faster than competitors that are dependent on storage component vendors to build and release a device before they can begin their own system-level qualifications.

In contrast, competing systems take a more commoditized view of storage, largely treating flash SSDs as if they were spinning disks. That leaves the media management, I/O translation, and maintenance behaviors up to each individual SSD — just as was the case with spinning disks. This approach creates a lot of overhead processing and introduces performance bottlenecks into the storage system.

The issue often becomes more pronounced with competitive systems when they use larger SSD sizes: The larger the drive size, the more that performance can be adversely affected. As a result, the performance, density, and reliability of competitive SSD-based systems are all significantly impacted. These trade-offs often lead competitive system configurations to use higher quantities of smaller drives — which translates to higher space and energy use — to meet a user's capacity and performance needs. Increasing capacity through DFM density rather than via the number of devices allowed Pure to launch the Evergreen//One energy efficiency SLA (service-level agreement) guarantee in early 2023, which provides service credits as well as density and consolidation services at no additional cost.

Modularity of Components

FlashArray and FlashBlade//S employ a very modular design in which major components that affect system performance, availability, and scalability can be hot-plug replaced quickly and easily, without tools. Controllers (in the case of FlashArray) or blades (in the case of FlashBlade), cache cards, storage devices, disk shelves, power supplies, and fans are redundant where meaningful (the disk shelves aren't necessarily redundant) and provide the hardware foundation for the system's "six nines" availability track record.

The FlashArray backplane can be field replaced as well, providing nondisruptive upgrades even between model lines. For example, Pure customers have upgraded nondisruptively from the FlashArray//X to the //XL (which uses a different chassis than the //X). Pure has also built modularity into its backplane before. Starting in 2016, FlashArray could support both SAS and NVMe storage devices. After Pure Storage shipped its first NVMe AFAs (NAFAs) in 2017, existing customers began upgrading their installed base systems from SAS to NVMe, all nondisruptively. Although its competitors began shipping their first NAFAs in 2018 and through 2019, none of them ever allowed their installed base of SAS systems to be nondisruptively upgraded in place to NVMe — they all required a forklift upgrade to get to NVMe. FlashArray's nondisruptive SAS-to-NVMe upgrade path was unique in the industry, and no other vendor has offered this level of "evergreen" upgradability for in-place scale-up systems.

Systems that are more modular in design offer more opportunities to upgrade different storage components to newer technologies over the life of an array, without requiring a total replacement. Pure's cache cards are an example of this design planning. Like many enterprise storage vendors, Pure Storage uses solid state storage for its cache, sending all writes there first to take advantage of the extremely fast write acknowledgements that the medium can provide. Competitive caches in an array are often on the storage controllers, but Pure Storage's use of separate cards allowed the company to upgrade cache cards without upgrading controllers and work with higher-capacity caches that could handle all writes and also a good number of reads. For example, in early systems, when the rest of the array was still plumbed with SAS, the cache cards used NVMe to provide lower latencies.

Simplicity of Design

FlashArray leverages a simpler design than the offerings of most of Pure Storage's competitors in the large enterprise systems market. First, the company has stayed with a dual controller design, while many of its competitors moved to a more complex design that allowed them to scale controllers beyond just two. Through the use of either a crossbar switch or a highly specialized ASIC, competitive systems could scale up to four and sometimes as many as eight controller pairs to extend the performance of their systems in place. In Pure Storage's case, customers needing more performance could just quickly and nondisruptively upgrade to a higher class of storage controllers. The crossbar switch and/or ASIC designs added significant complexity and even reliability issues to the competitive systems, both of which must be managed. Because of cache coherency overhead and other east-west traffic, these multi-controller systems did not scale performance linearly. The FlashArray uses existing system resources more efficiently with the much simpler dual controller design. Pure's design can indeed support multiple petabytes at very high performance and reliability levels in a single system, which is especially true for the high-end FlashArray//XL. In reality, many customers actually prefer to split such large workloads between the fault domains of multiple systems to ensure high availability. However they choose to implement their storage solution, Pure provides this high performance and capacity in a system that has been designed to be intuitive to install, manage, and upgrade.

Traditionally, large enterprise storage systems have included administrative interfaces that offer very deep and granular configuration flexibility. These "nerd knobs" (as they are often referred to by storage administrators) were originally needed because there was such a disparity between the performance of compute and the HDD-based storage resources that customers needed to continuously fiddle with settings to optimize performance as workloads evolved over time. When all-flash systems were introduced, one of the first things that storage administrators noticed was that with the predictably consistent sub-millisecond latencies AFAs delivered under load, almost all storage performance problems just went away. Vendors that introduced AFAs by upgrading their existing HDD-based architectures kept these granular controls because they were already a part of their storage operating systems, but many AFA start-ups took the opportunity to think differently about what they demanded from storage administrators.

Pure Storage took this latter approach with both of its all-flash products. Since it was taking a "blank sheet of paper" approach to all-flash storage, instead of making everything configurable, the vendor built its Purity software with a set of defaults that allowed a system to be unboxed and deployed very rapidly. Most of the enterprise-class data services, which in the past had to be individually configured, were now just "on" all the time. With HDD-based systems, there was a noticeable latency impact with many data services choices, but the performance of flash (and especially software that was purpose built to manage flash storage) made that a nonissue at the application level for almost all workloads in AFAs. For example, in the case of FlashArray, instead of having to decide which storage device-based data protection method to use from a list (which included RAID 1 [mirroring], RAID 5 [single-parity RAID], RAID 6 [dual-parity RAID], and RAID 10 [mirrored stripes]), and actually implement that method, Pure Storage assumes and implements a modified RAID 6 implementation for all data — automatically and without any operator intervention required. Features that were performance sapping in HDD-based and legacy-retrofitted AFA environments, like thin provisioning and inline data reduction, were instead written into the Purity software to be "always on," without impacting application performance in FlashArray.

While this is a great enabler for less experienced storage administrators (which by the way are becoming much more popular as the responsibilities for storage management in many IT shops are migrating more toward IT generalists), what about seasoned storage administrators that are comfortable with "nerd knobs?" Generally, administrators that may express a preference for them up front do so because, based on experiences with HDD-based arrays or legacy architecture AFAs, they do not trust that default configurations will give them all the performance they need for particular workloads and thus may keep them from meeting certain service-level agreements.

Pure Storage's all-flash design addresses this issue. Customer prospects concerned about this issue can talk to their colleagues that are using Pure Storage arrays (there are tens of thousands of systems in the field). They will find that Pure systems can deliver sub-millisecond latencies at scale (i.e., while pushing millions of IOPS). Ultimately, even sophisticated storage administrators come to appreciate the simplicity of FlashArray's default configurations once they see that they do not, in most cases, impact their ability to meet storage SLAs.

In the case of FlashArray, having features like inline data reduction on at all times, which can be guaranteed by Pure Storage and drives an average 5:1 data reduction ratio in mixed workload environments, produces very desirable economics. While FlashBlade utilizes inline data reduction, the nature of unstructured data does not yield results quite as dramatic. But it still offers many features and capabilities by default that greatly simplify the task of managing storage while delivering high performance. These are all critical parts of the Pure Storage value proposition — that customers will not have to give up performance to get simplicity of management, efficient use of storage resources, or a compelling total cost of ownership (due in part to the intuitive nature of the storage).

On a side note, when evaluating the total cost of ownership of any system, from Pure Storage or any other vendor, IT managers should focus on the effective capacity (the capacity available for use after formatting, data protection, and inline data reduction have been applied), not the raw capacity. Raw capacity advantages can quickly be overcome by higher data reduction ratios, particularly in larger configurations.

There are other simplicity aspects of the FlashArray CX that should be pointed out. Each independent system is one physical box with only six cables, which can be installed in 30 minutes. All array software, even advanced features like ActiveCluster (Pure Storage's stretched cluster option), are included with the base price of the array. Features like thin provisioning, inline data reduction, quality of service, and FIPS 140-2-compliant end-to-end encryption are always on. ActiveCluster uses a cloud-based mediator to resolve split-brain syndrome, removing the need to buy and maintain a separate quorum device for these configurations. Snapshots are "portable" in the sense that they are incremental, space efficient, and self-describing and can be shared between other FlashArrays, FlashBlades, third-party NFS storage, or the public cloud. Pure1, an AI/ML-driven cloud-based predictive analytics platform providing full-stack analytics, is included with the base price of each array. And Pure arrays include APIs that enable integration with other automation/orchestration tools, such as VMware and OpenShift.

The Impact of Storage Life-Cycle Extension

IDC has commented at length about the features and benefits of Pure Storage's Evergreen//Forever in *Pure Storage's Evergreen//Forever Subscription Continues to Drive Industry-Leading Customer Experience* (IDC #US49553922, Aug 2022). The original Evergreen program has been expanded into a subscription portfolio that includes Evergreen//One, Evergreen//Flex, and Evergreen//Forever to address the demands of customers in terms of balancing ownership, control, and SLAs for their storage environments. Aspects of the Evergreen//Forever subscription, like all-inclusive software and guarantees for data reduction ratios and fixed maintenance pricing for the life of the array, simplify the overall experience of owning and managing an enterprise storage array. Other features extend the life cycle of a FlashArray while enabling customers to take advantage of new technological advancements. The Evergreen subscription provides an included controller or blade upgrade every 3 years under the Ever Modern feature, and the Ever Agile feature allows customers to upgrade controllers or blades at any time with full trade-in credit for old components. The Capacity Consolidation option lets customers move to denser and more modern storage devices to expand system capacity without increasing footprint. Pure Storage provides trade-in credits for up to the full amount paid for the components being retired toward the purchase of the new components, which means customers are not rebuying storage equipment and/or capacity at regular intervals — investment in these components is truly preserved.

It is the combination of these features with the Evergreen architecture that doubles the storage life cycle of a Pure Storage array into the 8- to 10-year range and beyond. In competitive systems, forklift upgrades are required when existing technologies reach capacity or performance limits that cannot be upgraded in place — a cycle that happens roughly every 4–5 years. Limitations around backplanes (SAS only), controllers, and supported storage device geometries are all factors that drive this legacy forklift upgrade, but on a Pure array, these factors do not impede technology refresh. A SAS-based FlashArray system purchased in 2012 from Pure Storage that used off-the-shelf SSDs could be nondisruptively upgraded to an end-to-end NVMe-based system that uses DFMs and supports newer device geometries (for increased storage density) and different media types (MLC, TLC, and QLC) — all without having to rebuy performance and capacity that are already paid for. Performance and capacity upgrades do generate some incremental costs, but those are always defrayed by the guaranteed trade-in credit amounts that customers know up front when they buy their FlashArray. These are the features that support the vendor's claim that a FlashArray has a 10+ year life cycle, enabling the nondisruptive integration of the latest technologies to ward off the need to move to a completely new system. And the vendor claims to have customers that have done just that — continuously upgraded their arrays first purchased 10+ years ago, keeping them performant and in service through today. Pure claims that 97% of its arrays that have ever been sold are still in service today, looking and running like new arrays because of the ease of upgrading them. It also claims to have upgraded its customers' storage arrays well over 10,000 times.

The extended life cycle has very positive impacts. Storage administrators know the pain involved in data migration during technology refresh, and Pure Storage's approach means that over the life cycle of a single array, a relationship with the vendor could cut the need for that at least in half, allowing customers to avoid rebuying entirely new storage systems and replacing storage capacity every 4-5 years. IDC would encourage prospects that understand the technology behind Pure's Evergreen architecture to compare the total cost of ownership and the management impact of this life-cycle extension with that of the traditional 4- to 5-year enterprise storage life cycle.

AFA CHALLENGES AND OPPORTUNITIES

Prospective storage buyers may consider Pure Storage's claims about the value driven by the Evergreen Storage program as "just another program." Several competitors have introduced their own as-a-service and on-demand programs that attempt to address the needs of buyers looking for more IT agility and a shift to opex (operating expenditure)-based procurement, but few have the program maturity, install base and experience, or depth of program guarantees and SLAs that Pure offers. Communicating these differentiators is a challenge that the vendor must overcome in the marketing of its Evergreen architectural design choices (and the Evergreen Subscriptions portfolio). Like many unique offerings, the differences between Evergreen and those of Pure's competitors present both a challenge and an opportunity for Pure Storage as a vendor.

Buyers are reminded that the selection of both a vendor and an AFA architecture for critical workloads should be approached methodologically, as poor choices can lock IT into a long-term relationship with technology that poses several challenges:

- Inability to scale and expand to keep up with modern workloads and applications and lack of access to new software features and functionality caused by out-of-date hardware
- Expensive, time-consuming, and complicated processes for IT staff to upgrade legacy storage (over and over, usually through complete rebuys of existing storage and capacity) to meet the needs of the business
- Interruption to operations and services from downtime caused by legacy storage maintenance and upgrade processes
- Significant financial impact/costs associated with continual legacy storage upgrades
- Difficulty in supporting business ESG goals or sustainability metrics due to the need for additional space, power, and cooling for less dense, less efficient storage systems

Pure has focused on addressing these key challenges through its architectural choices — there is an opportunity for it to showcase CX success stories with FlashArray and FlashBlade to educate prospective buyers and gain market share.

ARCHITECTURE QUESTIONS TO ASK AND WHAT TO LOOK FOR

- Can data remain in place during all upgrades, or does it need to be moved off the affected storage array (whether onto another array, backup device, or cloud-based service)?
- If the system can't upgrade and expand nondisruptively, what is the recommended procedure? How long does it typically take to perform, and what are the performance or data availability impacts?

- Can software versions be upgraded without taking the system offline and without any reduction in performance?
- Are any optional or third-party products required to accomplish software upgrades?
- Can data remain in place during all controller upgrades, or should it be moved off whichever storage array is affected?
- Can this nondisruptive controller upgrade process be used when moving to different generations of a controller? If not, what is the recommended procedure? How long does it typically take to perform, and what are the performance or data availability impacts?
- Can flash storage media be expanded without taking the system offline and without any reduction in performance?
- Can older-generation flash storage media be used with newer controllers after upgrades, or does the controller upgrade require an upgrade to newer flash media as well?
- Are there any performance impacts or feature limitations when using older flash media with newer controllers?
- Can flash media of different generations, sizes, and geometries be mixed in a single array so that system capacity can be expanded over time without rebuying capacity?
- Can array capacity be consolidated onto higher-density media within the array later, without taking the system offline and without any reduction in performance?
- Are controller upgrades included in an optional subscription, maintenance, or other ownership programs, based on upgrading at regular intervals, to protect customer investment while modernizing the array?
- Are controller upgrades available on demand and as part of an optional subscription, maintenance, or other ownership programs, based on upgrading at a time of customer's choosing, to increase IT agility and protect customer investment while upgrading array?

CONCLUSION

The benefits to organizations in selecting all-flash arrays rather than spinning disks for many of their critical application and workload storage needs are well established: better performance, reliability, lower operational complexity and costs, and greater sustainability. And for those same reasons, the underlying storage architectures of AFA systems must also be considered.

Pure's Evergreen architecture and design philosophy are unique. The capabilities outlined in this white paper enable simple, nondisruptive technology refresh at reduced cost throughout the life of a Pure Storage FlashArray or FlashBlade system and can extend the life of an array to 10 years without imposing the performance, availability, scalability, or functional limitations that typically drive the requirement for a forklift upgrade.

The end result of Pure's architectural design decisions is that there are specific unique features that support Pure Storage's Evergreen claims and drive a differentiating CX around FlashArray and FlashBlade ownership. These features all come together to define Pure's value proposition — that customers will not have to give up performance to get extended storage life cycles, simplicity of management and growth, efficient usage of storage resources, or a compelling total cost of ownership for their enterprise storage platforms.

MESSAGE FROM THE SPONSOR

Interested in finding out how your organization can benefit from Pure's unique storage solutions? More detailed information can be found here:

https://www.purestorage.com/content/purestorage-com/do-storage-differently.html

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. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology, IT benchmarking and sourcing, and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives. Founded in 1964, IDC is a wholly owned subsidiary of International Data Group (IDG, Inc.).

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com