White Paper

Pure Storage Introduces a New Technology Upgrade Model with Evergreen Storage

Sponsored by: Pure Storage

Eric Burgener
June 2015

IDC OPINION

Managing technology refreshes is not a popular task with enterprise storage administrators, although it is a necessary one for successful businesses. As a business evolves, managing more data and adding new applications in the process, enterprise storage infrastructure inevitably needs to grow in performance and capacity. Enterprise storage solutions have traditionally imposed limitations in terms of their ability to be upgraded in-place to incorporate newer storage technologies that improve performance, increase storage densities, improve efficiencies, and lower overall costs. With its Forever Flash Program and associated set of well-targeted architectural enhancements to its All Flash Arrays (AFAs), Pure Storage aims to challenge customer preconceptions about the risk, expense and waste associated with forklift storage upgrades that generally occur every three to five years for most enterprises. Pure Storage's actions establish a new Evergreen Storage model for its customers to manage major technology refreshes; the model is designed to provide the following benefits:

- Allow customers to upgrade to the latest technologies across the board in controllers, external host and internal array connectivity, and solid state disks (SSDs) without any disruption or performance degradation
- Enable in-place upgrades that remove traditional risks that can be associated with forklift upgrades and data migration since the model requires neither
- Preserve customer investment since upgrades to next generation technology do not require re-purchasing hardware, re-licensing storage software products, or increases in maintenance and support costs

The Evergreen Storage model from Pure Storage provides a significant added incentive for AFA prospects to look at the company's FlashArray offerings. According to IDC's research, Pure Storage is among the leaders in market share by revenue among AFA vendors (see IDC #252304), and this new approach is designed to make the company's offerings that much more compelling.

IN THIS WHITE PAPER

Purchasers of enterprise storage have historically dealt with an upgrade cycle that was expensive, disruptive, inherently risky, time-consuming and occurred roughly every three-to-five years. In 2015, Pure Storage challenged customer preconceptions about the enterprise storage upgrade cycle with its Forever Flash Program which, combined with the company's technology architecture, creates what is referred to as Evergreen Storage. This IDC White Paper assesses Pure Storage's Evergreen Storage model, reviewing both existing approaches and this new approach in terms of the impacts on customers and their business and financial implications.

SITUATION OVERVIEW

The legacy enterprise storage upgrade cycle is one familiar to most storage administrators. A new storage array is purchased that includes a given storage capacity which may be expanded over the life of the product, but the maximum storage performance achievable by the system is fixed based on the capabilities of the controllers and the internal array bandwidth at the time the product is shipped. Regardless of how much capacity may be added over time, the maximum performance potential in terms of storage latency, throughput, and bandwidth does not increase. Successful enterprises tend to grow their businesses over time, requiring additional performance from their IT infrastructure. A typical enterprise storage life cycle varies, but is generally somewhere in the 3-5 year range. Ultimately the fixed storage performance of this system no longer meets requirements, and the business is forced to perform a forklift upgrade and repurchase the system, and deploy the newer technologies in storage controllers and storage media necessary to meet its requirements. The supplier may also increase maintenance renewal rates to force this upgrade. All of its data must be migrated to the new array during the technology refresh. This cycle repeats over time.

This approach has significant impacts on the business, but because it has been the "way things are done" in enterprise storage for so long, it is instructive to call out the following limitations:

- **The model locks customers into older technology.** When a legacy enterprise storage array is designed, it can include the latest in controller and storage media technology. Firmware and software upgrades over the life cycle of the product can provide incremental performance improvements, but customers are locked into the limitations of that technology for the life of the product. Although capacity can be added, often all the drives can only be of the type available when the system was first purchased. Customers do not have access to major advancements that provide order of magnitude improvements in performance, storage density or cost.
- **Forklift upgrades are disruptive.** Moving to the next generation of controller and storage media technology requires a completely redesigned array with typically much higher internal bandwidth to take full advantage of performance and density advancements. This means that a completely new array must be brought in to replace the existing array, and that means downtime. In addition, snapshot trees and replica libraries are lost with forklift upgrades, so such upgrades are disruptive to data services as well. How long the upgrade will take — and therefore impact application services — is a key question that must be answered as enterprises plan for the upgrade.
- **These upgrades are time-consuming and risky.** During the technology refresh, all of the data in the old array must at some point be migrated to the new array. Even the smallest enterprises today are dealing with generally at least tens of terabytes (TBs) of data, and most are dealing with hundreds of TBs and looking at managing petabytes (PBs) of data in the next several years. Even if that data is migrated over high performance local area networks like Fibre Channel (FC), migrating that much data can easily take days for many enterprises. Since this process inherently requires application downtime, enterprises typically plan it very carefully, leveraging snapshot, replication or other data services in an attempt to minimize the impact on the business. Often the new systems use a new, higher performance or more efficient on-disk format, so customers can incur conversion risk during the migration as well.
- **Upgrades are extremely costly.** First, a customer must buy the new hardware and software. None of the hardware and software from the older array can be transferred to the new one, so all of the capital expenditure must be repeated even if the customer just wants the same basic features (x amount of capacity, snapshot and replication software, etc.). Second, maintenance

and support costs will increase based on the price of the new solution. And third, to help this inherently risky process go more smoothly, many enterprises hire outside professional services firms to plan and execute the technology refresh, a decision that can easily add anywhere from tens of thousands to several hundred thousand dollars in costs to what is already considerable expense.

IDC refers to this legacy approach as Model 1. Some enterprise storage suppliers looking to minimize the impact of technology refresh have introduced an overarching software layer that allows storage to be more completely virtualized across multiple arrays. This has two impacts. First, systems of disparate types can be combined into loosely coupled groups, sometimes referred to as "clusters." This allows newer and older systems to be combined so that their resources can be more flexibly allocated in logical pools that are less restricted by hardware limitations, providing an easier way to incorporate new technology into systems. Second, storage virtualization can allow data to be migrated on-line, significantly minimizing the disruption associated with data migration. IDC refers to this approach as Model 2.

While Model 2 addresses some of Model 1's issues, it still suffers from significant problems from a customer's point of view. While new arrays can be added to the IT infrastructure, customers are still locked into old technologies with the existing systems. Many customers use the on-line data migration to move the newer array into production and, once the new array is serving production data, retire the older array. Data migrations still take time, and often impose performance degradation on production applications during the migration process, which will generally occur more slowly than it would if the data was migrated between two non-production systems. Customers will still have to face losing their snapshot trees and replica libraries, and any risks incurred by a conversion to newer on-disk formats are still present. Finally, customers still have to repeat their hardware and software spend when buying the new system, losing their capital investment on the original array while they are forced to accept the maintenance and support cost increases associated with the newer array.

Pure Storage's Evergreen Storage Model and the New FlashArray//m

Pure Storage is an All Flash Array (AFA) vendor whose arrays are based around a modular, stateless design. Seeking to provide customers with a better approach to managing enterprise storage lifecycles, Pure Storage announced in June 2015 changes to an existing maintenance program called Forever Flash. This change was made in conjunction with the release of the Pure Storage FlashArray//m, an AFA that has been entirely designed to support a new value proposition in the industry with respect to how upgrades are managed across technology generations. This maintenance program and the architectural changes to the FlashArray//m challenge customer preconceptions about the enterprise storage technology upgrade process and blur the traditional distinctions in the benefits between scale-up and scale-out architectures.

The Forever Flash maintenance program applies to all Pure Storage FlashArrays under a valid maintenance contract, and includes three basic provisions:

- **Flat and fair.** Once a customer has purchased an AFA from Pure Storage, maintenance pricing on a per/GB basis on that system will be capped, ensuring that it either stays flat or decreases over time for as long as the customer owns the system. The actual maintenance cost may increase over time as a customer builds out the capacity of a system, but the costs at a component level are capped at the prevailing rates at the time of original purchase.

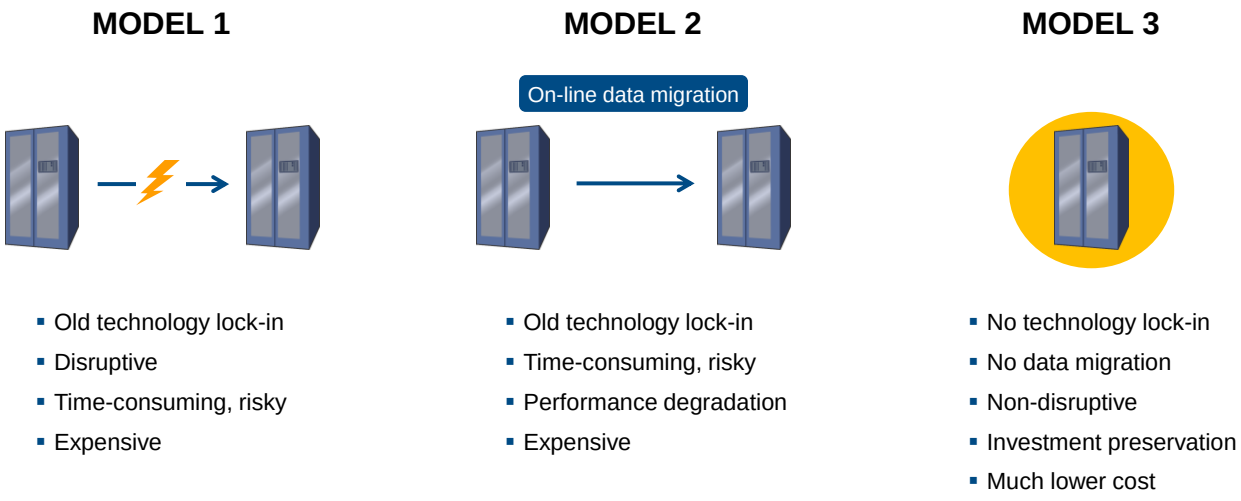
- **Free Every Three.** Upgraded controllers that support the latest in storage technologies, including new higher density drives, are provided free of charge to Pure Storage customers every three years, including controllers that cross product generations.
- **Forever Maintenance.** Any failed components are replaced at no additional charge by Pure Storage during the life of the array. This includes coverage for SSD wear-out conditions.

Forever Flash helps modernize customer's arrays using maintenance and support (i.e., opex) dollars over time. Alternatively, customers with a near-term demand for the latest in controller performance, scale, and features can purchase a controller upgrade (i.e., capex). Pure Storage provides customers a trade-in credit for their existing controllers when they purchase a controller upgrade bundled with a capacity expansion.

Changes have been made in the FlashArray//m platform to allow every internal component that impacts performance, including both the controllers themselves as well as host connections and internal busses, to be upgraded on-line without having to shut down applications or incur any performance degradation. The ability to just upgrade controllers is not sufficient to leverage all the performance of which new storage technologies are capable, since older host connections and internal busses may not support acceptable bandwidth. In the past, an upgrade to an entirely new array frame included higher performance controllers, increased bandwidth both internally and externally, and newer, faster, more dense storage devices. The Forever Flash maintenance program, combined with the enhancements in the FlashArray//m, mean that a Pure Storage system is designed to be completely upgraded across technology generations without requiring any downtime, data migration, or additional hardware and software expense. In effect, the Forever Flash maintenance program aims to establish a third model for enterprise storage technology refresh that has significant advantages over the other two models. (See Figure 1). Pure Storage refers to this new model as Evergreen Storage.

FIGURE 1

Enterprise Storage Upgrade Models



Source: IDC, 2015

Analysis and Implications

The benefits of Evergreen Storage are predicated on Pure Storage's FlashArray architecture as it is implemented in the FlashArray//m. This architecture is designed to provide flexibility that traditional scale-up architectures do not, allowing customers to increase capacity by adding more solid state disks (SSDs), mix flash geometries so that newer, more dense storage devices can be added as they become available, upgrade performance by moving to higher model or next generation controllers, enhance connectivity by either adding or upgrading ports to higher bandwidth options as they become available, and take advantage of new features through software updates. Pure Storage's FlashArray//m aims to allow all five of these areas to be expanded independently so that customers can take advantage of the annual innovation cycle driven by newer Intel processors (for faster controllers), newer flash technologies (for higher densities and lower cost/GB), and firmware and software updates to keep their array continually up-to-date in terms of the latest storage technologies.

A working knowledge of some of the FlashArray//m technical features is critical to understanding the benefits of Evergreen Storage. Although Pure Storage's FlashArrays have always been built around a dual controller architecture, they do not use the typical active/passive controller design. During normal operation, both controllers are in use, accepting I/O from the host side, while all I/O to SSDs internal to the array are handled through a single controller. The controllers are designed so that a single controller can deliver the maximum performance at which the array is rated, which means that during normal operation each controller would be no more than 50% loaded on the host side. In the event of a controller failure, all I/O on both the host and array sides are handled by the single remaining controller with no impact on performance. Whether a controller is taken off-line for maintenance purposes or fails, there is no performance degradation associated with that action.

With the FlashArray//m, Pure Storage is making changes to how it packages its flash storage capacity. SSDs are now incorporated into a flash module that includes two drives in each field replaceable unit (FRU) and an NV-RAM Module that includes a hot plug PCIe connector. Making PCIe hot plug compatible is an innovation that clearly supports the Evergreen Storage goals of enabling the integration of newer, higher performance storage technologies in a completely non-disruptive manner.

When upgrading to next generation controllers, Pure Storage customers can get the benefit of this design to perform "rolling upgrades" without impacting application services. The FlashArray//m controllers can be configured with either 8 or 16 Gb/s FC or 10GbE iSCSI host connections, include non-transparent bridging (NTB) over PCIe for inter-controller connections, and use 12Gb/s SAS for controller-to-flash device connectivity. The first FlashArray//m uses PCIe Gen3 to deliver up to 15.75 GB/sec of bandwidth between controllers, but is field upgradeable to PCIe Gen4. With the FlashArray//m design, the host-to-array, controller-to-controller, and controller-to-flash device connectivity can all be upgraded as part of a controller FRU upgrade which does not cause downtime or impact application performance. Because flash geometries can be mixed, data does not have to be migrated off older drives as newer ones are added. Newer, higher performance, more cost-effective and denser SSDs can be added in shelves that support higher bandwidth, allowing customers to take advantage of flash technology advancements as they become available. With this design, the entire system can be completely upgraded to newer technologies in-place, removing risks associated with technology lock-in, data migration, loss of snapshot trees or replica libraries, or cost increases.

With Models 1 and 2, on-disk format changes that require data conversions can introduce risk, often result in downtime, and take time and effort on the part of a customer. The Pure Storage FlashArrays include an adaptive metadata structure that is scalable, versioned and hierarchical. When major

metadata updates occur, the old metadata structures can be left intact and referenced by the new metadata. The natural background optimization processes of the array migrate the older metadata to the new format over time without the need for any explicit data migration tasks and without any application downtime or performance impact. The use of variable-sized data segments allows segment structures to be updated to add additional features transparently, without any data migration or application downtime. This is not a theoretical argument — Pure Storage has enhanced its metadata structures and data layout segments in every major release since its initial product offering without requiring downtime or data migration. Fixed metadata structures and segment sizes in competitive AFAs introduce limitations that generally preclude the ability to make these types of major changes without data migration, introducing risk, effort and downtime.

Note that Pure Storage's ability to perform an entire array upgrade in place without downtime is designed to have three key implications. First, data must never be migrated during the process, reducing risk and saving time and money. Second, customers preserve any capital investments they have made even as they migrate to newer technologies. They never need to buy another frame or re-license any snapshot, replication or other software that they have already licensed on a FlashArray, and they continue to pay the same maintenance and support costs at a component level that they were paying on the original system. And third, there is no downtime or performance degradation associated with the technology refresh process.

With the ability to completely upgrade the array in-place without downtime, Pure Storage is expecting to extend the enterprise storage life cycle, which can be as little as three years, to as much as ten years or more while still allowing customers to take advantage of newer, faster, more cost-effective technologies as they become available. Although Pure Storage will never require a frame swap, the company expects that customers will decide on their own when they want to do this based on the nature of storage technology advancements. In the interim, however, Pure Storage customers can get the benefits of non-disruptive technology refreshes that do not require data migration, constant maintenance costs at a component level, and investment protection.

The cost implications of Evergreen Storage are significant. Relative to the other two models where arrays must be replaced as often as every three years, hardware and software must be re-purchased, maintenance and support costs significantly increase, and data must be migrated in a disruptive and time-consuming manner, Pure Storage's model is geared to dispense with all of that. The actual cost savings will vary significantly based on the size of the system, but even assuming only one technology refresh over a six year life cycle, capital costs will be one-half to one-third as much while maintenance costs will be roughly half as much. The costs with Evergreen Storage will be front-loaded -— the big savings in lower costs occurs at each technology refresh — but even then those upfront costs are comparable to the initial purchase cost of other AFAs, with some being lower and some being higher.

There is one other benefit that needs to be taken into account. The non-disruptive upgrade path gives customers the option to perform technology refreshes more often without impacting application services. The Forever Flash maintenance program includes controller upgrades every three years, but over the past four years of shipments Pure Storage has released new, faster controllers once a year. It is part of Pure Storage's product strategy to continue to innovate at this rate going forward, in effect allowing customers to ride the Intel processor technology curve very closely. Customers can choose to upgrade their controllers more often by paying separately for them and receive trade-in credit for old controllers when purchasing controller plus capacity upgrade bundles. Pure Storage will pick up the upgrade cost once every three years. This is designed to give FlashArray customers the ability to incorporate newer controller technologies into their systems in the same non-disruptive manner that

scale-out storage architectures do, but does so by leveraging a mature, proven scale-up storage architecture that is also more efficient since it doesn't tightly couple capacity and performance together. Customers purchase only the capacity and/or performance they need, without any excess.

Challenges

As revolutionary as Evergreen Storage is designed to be, Pure Storage could improve upon it even more. Although the company supports on-line capacity expansion of FlashArray//m systems with newer, denser, more cost-effective and higher performance flash storage, it does not today provide a non-disruptive way to migrate data on older drives to newer drives in the event a customer wants to retire them. This option would give customers the ability to increase a system's storage density across a much greater range in a manner which is fully compatible with Pure Storage's in-place technology refresh policy.

The use of in-line storage efficiency technologies like compression, deduplication, thin provisioning and others has always been an integral part of FlashArray design, and empirical data collected by Pure Storage across its entire installed base indicates that, on average, customers enjoy data reduction ratios in the 5:1 to 6:1 range. Under that assumption, the effective usable capacity of a FlashArray//m system is in the 500TB range – a capacity that today probably meets the requirements of most enterprises. IDC is already seeing many customers that plan to migrate all of their primary storage applications to AFAs over the course of the next several years, and as a result are expressing an interest in vendors' long term plans for performance and capacity expansion. Pure Storage does not yet support clustering multiple systems together, although the company supports a unified management interface (Pure1) that offers cloud-based management of multiple FlashArrays. The ability to cluster multiple systems together and non-disruptively migrate data across systems would significantly extend the capabilities of a Pure Storage solution to handle high growth over an extended life cycle even as customers pursue a strategy around dense mixed virtual workload consolidation. This does not take anything away from the value that Evergreen Storage provides to customers today and in the foreseeable future, but would certainly improve the value that Pure Storage's technology refresh strategy provides to customers across a life cycle that includes multiple technology refreshes.

CONCLUSION

The traditional enterprise storage upgrade cycle imposes significant burdens on data centers in terms of risk, expense and waste, and is sorely in need a facelift. Pure Storage's Forever Flash maintenance program, buttressed as it is by changes in the company's system architecture, is designed to offer customers a much better choice that can keep them working with the latest technologies, lower costs significantly over a newly extended enterprise storage life cycle, and remove the risk, downtime, and hassle associated with forklift upgrades and time-consuming data migrations. This new technology refresh strategy throws down the gauntlet to other enterprise storage providers, challenging them to provide more value to their own customers, and absolutely makes Pure Storage solutions more attractive to enterprises.

About IDC

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

Global Headquarters

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