



IDC TECHNOLOGY SPOTLIGHT

Primary Flash Market Evolving to Next-Generation Architectures

April 2017

Adapted from *Worldwide All-Flash Array Market Shares, 3Q16: Top Five Market Share Holders Consolidate Gains*, by Eric Burgener, Liz Conner, and Jingwen Li, IDC #US42374817

Sponsored by Pure Storage

As the information technology (IT) industry enters the cloud era, where hybrid IT is the dominant deployment model in organizations of all sizes, the capabilities of primary all-flash arrays (AFAs) will need to evolve to handle cloud scale and agility. Datacenters of all types and sizes will be using AFAs as general-purpose storage platforms more and more and will be increasing the workload densities that these systems must support. A key challenge is that current-generation AFAs have fundamental bottlenecks that both limit their consolidation and cloud-scale capabilities and create architectural risk as future memory technologies emerge. A next generation of AFAs — what could be called cloud-era flash arrays — will be best positioned to meet these requirements cost-effectively, will employ flash-driven rather than flash-optimized architectures, and will have NVMe rather than SCSI technology at the heart of their designs. This Technology Spotlight examines the evolving primary flash array market with a particular emphasis on what next-generation flash-driven enterprise storage architectures will look like. It also looks at the role Pure Storage, with its FlashArray//X, plays in this strategically important market.

Introduction

As flash storage has permeated mainstream computing, enterprises are coming to better understand not only its performance benefits but also the secondary economic benefits of flash deployment at scale. This combination of benefits — lower latencies, higher throughput and bandwidth, higher storage densities, much lower energy and floor space consumption, higher CPU utilization, the need for fewer servers and their associated lower software licensing costs, lower administration costs, and higher device-level reliability — has made the use of AFAs an economically compelling choice relative to legacy storage architectures initially developed for use with hard disk drives (HDDs). As growth rates for hybrid flash arrays (HFAs) and HDD-only arrays fall off precipitously, AFAs are experiencing one of the highest growth rates in external storage today — a compound annual growth rate (CAGR) of 26.2% through 2020.

Already, AFAs drive over 70% of primary storage spend, and over the next year, 76% of those organizations that do not already use an AFA in production have plans to evaluate and/or deploy one. AFAs are widely used for the mixed consolidation of primary workloads, with 47% of deployed AFAs hosting from 5 to 9 workloads and 36% hosting 10 or more. This number is expected to increase precipitously as organizations continue to retire aging storage equipment and move those workloads to AFAs over the next 12 months to gain the benefits of flash deployment at scale.

We are at an inflection point in the all-flash markets. As workload density increases, the baseline requirements for AFAs will evolve toward systems that offer more performance and more capacity with higher infrastructure densities. Moving to this new level of infrastructure density will require some significant changes to enterprise storage architecture. Infrastructure density can be thought of as a conceptual measure of how much business-relevant work can be done (transactions per second, page views per second, profit per hour, ability to support concurrent users at sub-millisecond response times, etc.) per U of rack space.

There are several challenges that must be addressed as AFAs evolve to become the mainstream general-purpose enterprise storage platform. Larger SSDs help improve storage density but suffer in terms of the IOPS per terabytes they produce relative to smaller devices. Older interfaces originally designed for use with HDDs, such as SAS and SATA, are relatively slow and inefficient compared with interfaces like NVMe that were specifically designed with flash in mind. Depending on controllers built into individual SSDs to handle I/O optimization and garbage collection will increasingly become a limitation on the efficiency that can potentially be achieved with systems that are architected to use "flash as flash" instead of "flash as disk."

These factors will limit the ability to consolidate workloads at the scale that will be demanded over the next several years. As organizations move to the hybrid cloud environments of the future and deploy more next-generation applications (NGAs), general-purpose storage platforms will need to offer the infrastructure density to support literally hundreds of different workloads, along with the multitenant management tools to effectively administer these next-generation environments.

The next-generation AFA architecture needed to meet the requirements for dense workload consolidation at cloud scale must truly be designed for high efficiency with flash and other emerging memory technologies and generally architected for cloud-era environments. Conceptually, this means a number of important changes, not least dispensing with much of the legacy HDD technology that today resides between the array software and the flash media to achieve much better efficiencies and get the most out of flash and other emerging memory technologies as they mature to mainstream persistent storage status.

This type of design is not just flash optimized but flash driven, with no artifacts of the prior HDD-based era limiting what can be achieved with flash. NVMe technology must be at the core of this next-generation AFA design. It offers significant benefits across multiple areas:

- **Performance.** Without the additional latencies imposed by HDD-based SCSI I/O stacks, systems will be able to further improve the number of IOPS a given amount of storage can sustain. No flash translation layer will be required since the system is designed to natively work with "flash as flash." With less I/O overhead, systems will be able to deliver lower latencies, higher throughput, and improved bandwidth since more of a system's theoretical performance maximums can be achieved. This higher performance needs to be balanced with higher-performance host connections leveraging 40/50/100GbE and NVMe over Fabric options so that applications gain a high-bandwidth, low-latency connection straight through to flash. This is an important dimension as it blurs the distinction between SAN and DAS, enabling next-generation AFAs to deliver internal storage latencies with the reliability, manageability, serviceability, and efficiencies of shared storage.
- **Higher efficiencies.** The ability to schedule I/O optimization, garbage collection, and replacement of failed flash cells at the system level with visibility down to the individual flash die will result in much better utilization of system resources, driving lower cost for a given level of performance. Managing these tasks at a global level removes the uncertainties associated with managing them at the individual device level and will further improve these systems' ability to deliver predictable, deterministic latency. Working with a single, global pool of over-provisioned flash media is more efficient than working with multiple, device-level pools — another factor that will help to drive better cost efficiencies by reducing the amount of over-provisioned capacity needed to support any given level of flash endurance.

- **Improved reliability.** Global management of I/O optimization will result in lowered write amplification and more efficient scheduling of writes in a manner that helps to improve overall flash endurance. Because the pool of over-provisioned capacity is managed globally, a smaller pool size is needed for any endurance level, lowering overall flash cost per gigabyte as viewed from the system level.
- **Lower cost.** Taken together, all these features mean that a higher level of performance can be delivered from any given definition of system resources in terms of CPU cycles and raw flash capacity. Flash media costs are expected to decline at a CAGR of 26.0% through 2020, but flash-driven designs will help to further lower the realized cost of flash performance and capacity at the system level.

Trends

Since their initial introduction, we have seen AFAs move from a dedicated application deployment model to one that must support mixed workload consolidation. The next generation of AFAs will take that to the next level of performance, scalability, and infrastructure density. Requirements for this next-generation class of AFA include:

- **Performance and scalability.** These systems will need to offer easy, nondisruptive performance scalability that accommodates roughly tens of gigabytes of bandwidth, consistent microsecond (instead of millisecond) latencies as well as fast host connectivity, and capacity scalability into petabytes (of effective capacity, which takes storage efficiency technology into account).
- **Comprehensive, mature enterprise-class data services.** These must include in-line storage efficiency technologies, such as compression, deduplication, thin provisioning, pattern recognition, write minimization, and space-efficient snapshots and replication, which do not impact the systems' ability to deliver microsecond latencies, as well as encryption, quality of service, dense multitenant management capabilities, and cloud-based predictive analytics for telemetrics purposes.
- **Flash-driven architecture.** These systems must be specifically designed for flash and other emerging memory technologies without any limitations imposed by the vestiges of HDD-based designs. This will require the elimination of any remaining disk-era technologies — SAS, SATA and SCSI technology in general, flash translation layers, SSDs ("flash as disk" instead of "flash as flash"), and device-level garbage collection, write optimization, and over-provisioning. It means an architecture designed with nothing between the flash media and an array software layer that globally manages issues like I/O optimization, garbage collection and over-provisioning at the system level. NVMe technology should be at their core, and they should employ an architecture that is ready to accommodate future advances in persistent memory technologies.
- **Simplicity.** Given the continued migration of storage management tasks away from dedicated storage administration groups more toward IT generalists, systems need to be extremely easy to use with comprehensive wizards, extensive use of automation, and an ability to integrate through documented APIs with datacenter workflows (including cloud-based workflows). The systems also should leverage cloud-based predictive analytics to identify and resolve a high percentage of events and scenarios that in the past required manual intervention.

As customers move to consolidate more workloads onto their AFAs, the poor efficiency of array designs that treat "flash as disk" will become a limiting factor across several different metrics — top-end performance and scalability, infrastructure density, floor space and power consumption, reliability, and cost. In the long run, it is clear that NVMe technology will replace SCSI technology in enterprise-class arrays. This transition will need to occur sooner for certain customers, depending on their strategic plans for increasing storage infrastructure density.

Considering Pure Storage

Pure Storage was one of the first vendors to enter the AFA space back in 2011 with its FlashArray, a primary storage flash array targeted for the consolidation of block-based workloads. In 2016, Pure Storage announced FlashBlade, a big data flash array supporting file- and object-based environments and targeted for use with big data analytics, cloud-native applications, digital science, engineering and design workloads, and 4K/8K media workflows. This broad portfolio gives Pure Storage the ability to realistically deliver an all-flash datacenter that covers all types of workloads — structured and unstructured. With its FlashArray//X announcement, Pure Storage is now introducing the next-generation AFA for block-based workloads, a system that is designed for the industry evolution to cloud-era flash arrays.

FlashArray//X is an ideal example of a flash-driven architecture. Its "DirectFlash" design leverages experience already achieved with this approach on their FlashBlade platform and includes no vestiges of HDD limitations — the system uses raw flash on custom flash cards (CFMs) instead of traditional SSDs, as well as NVMe for both the I/O interfaces and the array backplane. In addition, the design employs the Purity Operating Environment array software that does flash management, garbage collection, and all I/O scheduling and optimization at a global level for very high efficiency. Over-provisioned flash is managed from a single global pool. There is no flash translation layer, no write cache (the system provides uniform access to all flash in the array with its NVMe-centric design), and no SCSI. The system is software defined, leveraging the mature Purity Operating Environment, with its full complement of enterprise-class data services that Pure Storage uses on both the FlashArray and FlashBlade and has been hardening over the past five years. FlashArray//X can deliver maximum performance (gigabytes of throughput) with just 10 DirectFlash Modules, enabling organizations to maximize the amount of business-relevant work that can be done per U of rack space. FlashArray//X also scales to roughly 1PB of effective capacity in just 3U (assuming a proven 5:1 data reduction ratio against block-based workloads over the past five years), providing an attractive combination of infrastructure density and scalability. When deployed as a full rack of 13 arrays (plus top-of-rack networking) managed seamlessly by Pure1, FlashArray//X delivers ~15PB capacity and over 100GB/s of bandwidth, making it a leader in tier 1 application consolidation. Several high-performance host connectivity options exist, including 40/50/100GbE and NVMe over Fabric coming in the future.

Like Pure Storage's other enterprise-class arrays, FlashArray//X is covered by its Evergreen Storage subscriptions so existing FlashArray//m customers can upgrade to it without disruption, data migration, or repurchase of their existing storage capacity. It offers SaaS-based monitoring and reporting through Pure1, the company's cloud-based management and support platform; provides predictive analytics; and offers flexible integration and automation via REST APIs. FlashArray//X goes further and guarantees effective capacity to remove any data reduction concerns, provides nondisruptive upgrades to next-generation controllers and media, and offers investment preservation for both controllers and flash media. According to the company, Evergreen Storage is a key contributor to the high level of customer satisfaction that Pure Storage generates with its customers — 70% of Pure Storage's business comes from repeat customers, and the company has a Net Promoter Score (NPS) of 83 (on a scale of -100 to +100) — a score in the top 1% of business-to-business (B2B) providers. NPS is a standardized measure of customer experience that is used across 220 different industries, with a higher NPS score meaning happier customers.

Challenges

As these types of next-generation systems emerge, vendors that have chosen to stay with traditional SSD-based designs will continue to compete. While the efficiency benefits of designs that treat "flash as flash" rather than "flash as disk" seem undeniable, the concept of "good enough" does not necessarily mean that the best technology implementation wins. IT organizations will need to define

their primary purchase criteria (top-end performance and scalability, infrastructure density, floor space and power consumption, reliability, cost, etc.) and evaluate which designs best meet their mix of requirements. Next-generation AFAs like Pure Storage's two entries in this space, FlashBlade and FlashArray//X, are new, and customers will need to see how they perform in actual use and whether they deliver on their promise of significantly improved efficiencies at lower cost. These new "cloud-era flash arrays," however, do appear to offer significant efficiency advantages that translate to lower cost than legacy, SCSI-based designs can deliver.

Conclusion

As flash costs continue to drop and new, flash-driven designs help to magnify the compelling economic advantages AFAs offer relative to HDD-based designs, mainstream adoption of AFAs — first for primary storage workloads and then ultimately for secondary storage workloads — will accelerate. Well-designed AFAs that still leverage legacy interfaces like SAS will be able to meet many performance requirements over the next year or two. Those IT organizations that aim to best position themselves to handle future growth will want to look at next-generation AFA offerings, as the future is no longer flash-optimized architectures (implying that HDD design tenets had to be optimized around) — it is flash-driven architectures.

IDC believes that we will start to see NVMe technologies appear more widely in all-flash, general-purpose enterprise storage platforms in 2017. Already we are seeing many vendors talk about how they are making their platforms "NVMe ready" to better accommodate increasing performance requirements. Those AFAs that will be best positioned to deliver highly efficient, low-cost operations will be those next-generation designs that employ a flash-driven architecture, completely leaving the limitations of SCSI technologies behind. To the extent that Pure Storage can deliver on the promise of increasingly efficient and performant flash-driven systems with FlashArray//X, the company has a significant opportunity to capitalize on its past successes and grow market share even further.

ABOUT THIS PUBLICATION

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

COPYRIGHT AND RESTRICTIONS

Any IDC information or reference to IDC that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For permission requests, contact the IDC Custom Solutions information line at 508-988-7610 or gms@idc.com. Translation and/or localization of this document require an additional license from IDC.

For more information on IDC, visit www.idc.com. For more information on IDC Custom Solutions, visit http://www.idc.com/prodserv/custom_solutions/index.jsp.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com