IDC ANALYZE THE FUTURE

IDC MARKET SPOTLIGHT

Sponsored by: SHI International Corp.

Advanced infrastructure technologies and cloud-based services let customers consolidate and modernize storage workloads in new ways. Organizations must now continuously evaluate their storage strategies in the context of security requirements, digital business initiatives, and data management needs.

Advanced Infrastructure Technology Redefines the Value of Storage Workloads

April 2019

Written by: Andrew Smith, Research Manager

Introduction

Limited budgets and high data growth continue to push IT organizations to do more with less. This can mean a range of things: better integration of on-premises and cloud assets, reduced energy consumption, improved storage efficiency, storage automation, or increased use of cloud-based storage services and outsourcing. The fact is that IT leaders have more choices than ever when formulating their infrastructure strategies.

This IDC Market Spotlight highlights key market forces influencing IT infrastructure and datacenter purchases and explores how the adoption of advanced infrastructure technology, including cloud services, flash storage, and software-defined storage (SDS), is changing the way organizations think about their storage workloads.

AT A GLANCE

KEY STATS

- By 2025, the global datasphere will grow to 163ZB (i.e., a trillion gigabytes). That's 10 times the 16.1ZB of data generated in 2016.
- This exponential data growth will continue to put pressure on all organizations to run their primary and secondary storage environments at peak efficiency.

Definitions

This section provides brief definitions and examples of how IDC classifies storage workloads across three tiers: primary, secondary, and emerging. This terminology is used extensively throughout the analysis.

- **Primary storage workloads:** Primary workloads are latency sensitive, often considered "mission critical" to the business, and store structured data (usually block-based storage). Online transaction processing (OLTP) databases, enterprise resource management applications, and collaboration platforms are examples of primary storage workloads. These applications also require high availability and rapid recovery.
- **Secondary storage workloads:** Secondary workloads are less latency sensitive and are geared toward cost effectively supporting large data sets over extended periods of time. These systems store unstructured data (file or object based). Backup, disaster recovery, archive, and other data sets that need to be retained for business or operational purposes are typical secondary storage workloads.
- **Emerging storage workloads:** As big data/analytics become more important in IT organizations of all types, CIOs are seeing a third storage requirement profile arise. Big data platforms, often based on open source NoSQL databases such as Cassandra, Hadoop, MongoDB, and Redis, allow for the distributed processing of large data sets across clusters of compute. Many modern applications for Internet of Things (IoT), artificial intelligence, and machine learning are designed with this architecture in mind. Throughput, bandwidth, and multipetabyte scalability are critical for these workloads.

Benefits

Storage workload modernization is a strategy that, in conjunction with virtualization, can improve the efficiency of IT operations. Modernization typically takes place within primary or secondary storage environments, but not across both. Because primary and secondary storage workloads have very different requirements in terms of performance, scalability, recovery, and cost per gigabyte, it has not been feasible to consolidate them on the same platform and at the same time meet individual application requirements. New technologies — namely, persistent flash storage, SDS, and cloud — are changing this dynamic. With enterprises of all sizes undergoing digital transformation and modernizing infrastructure, it is an opportune time to reconsider storage consolidation strategies to improve efficiencies and lower cost.

In just the past three years (perhaps since an enterprise's last storage technology refresh cycle), several of the following technologies have matured, significantly changing the landscape with regard to consolidating multiple storage workloads onto a single platform:

- **Affordable, persistent flash storage.** Continued price drops are narrowing the difference between flash and spinning disk on a price-per-gigabyte basis. Certain workloads that can significantly benefit from inline data reduction already drive a lower total cost of ownership (TCO) for all-flash arrays (AFAs) versus hybrid flash arrays (HFAs) or hard disk drive (HDD)–only arrays. AFAs already dominate latency-sensitive environments — in 2017, over 70% of primary storage spend was driven by AFAs, with the remainder taken up by HFA and HDD-only options.
- **Software-defined storage.** Hardware-defined storage solutions with specialized software bundled with and running only on a single proprietary hardware platform are a thing of the past. The future is software defined, meaning most of the storage functionality is abstracted in software designed to run on commodity off-the-shelf (COTS) hardware, such as x86 servers, running standardized operating system distributions. SDS architectures enable valuable capabilities such as the ability to apply data services at the application level. In addition, they offer the ability to deploy self-driving storage that automatically adapts based on workload or to meet defined service-level agreements (SLAs). SDS designs can also offer nondisruptive performance and capacity expansion.
- **Cloud-based services.** Public cloud infrastructures offer the opportunity for enterprises to effectively outsource part of their IT infrastructure, shifting from a capital expenditure model to an operational expenditure model. Cloud-based infrastructure services not only reduce the equipment an enterprise must buy and own but also lower energy and direct administrative costs (because the cloud provider houses and manages the physical infrastructure). Public cloud infrastructure also allows for IT elasticity (easy expansion/contraction of resources) that a direct ownership model can't match. The ability to outsource select storage workloads (e.g., backup, archive, management) is a powerful tool to provide IT agility. Hybrid cloud environments where an enterprise's IT operations include both on-premises, owned equipment and cloud-based infrastructure are already the norm, although the split between these environments varies depending on many variables (e.g., organization size, industry vertical, workload type).

While this list isn't exhaustive, IDC believes these three technology pillars will continue to redefine the value of storage workloads in the modern, digital enterprise. These technologies will allow enterprises to decrease the number of storage silos needed to meet workload requirements, but they will not allow all organizations to simply consolidate primary and secondary storage workloads onto a single platform. However, in most real-world cases, these new technologies and deployment models will expand the range of workload consolidation that can be handled by a single platform.

Considerations

IDC forecasts that by 2025, the global datasphere will grow to 163ZB (i.e., a trillion gigabytes). That's 10 times the 16.1ZB of data generated in 2016. This exponential data growth will continue to put pressure on all organizations to run their primary and secondary storage environments at peak efficiency. The challenge for many buyers is not simply an "either-or" adoption choice regarding public cloud, flash storage, or software-defined storage; rather, it is also about how to execute several of these initiatives in tandem while remaining secure, compliant, and cost efficient. IT and line-of-business (LOB) decision makers involved in infrastructure and storage-related purchases associated with the three technology pillars outlined in the previous section should consider the following factors:

- **New infrastructure solutions must not compromise the organization's digital trust and digital resilience:** Organizations providing modern, digital products, services, and experiences must ensure their platforms are reliable and trustworthy. Events such as the Marriott breach in September 2018 that exposed the personal data of over 500 million customers or the 16-hour service outage in November 2018 that disrupted Microsoft Office 365 users worldwide because of issues with multifactor authentication are examples of untrustworthy and unreliable experiences. IDC believes that to deliver consistent, trustworthy experiences, organizations must continuously improve their digital resilience. We define digital resilience as the rate at which a system regains structure and function following a threat or disaster through a series of planned proactive, preventive, and reactive measures involving people, process, and technology. Any modern infrastructure initiative should be designed and architected with the preservation of digital resilience as a central tenet.
- **Organizations should be consciously aware of data sprawl and decision paralysis:** The exponential growth of enterprise data not only will generate new opportunities but also will bring a host of challenges. The growing volume of enterprise data means information can be spread more widely and unpredictably throughout the organization, creating more storage silos than ever before. Organizations must reduce the complexity associated with enterprise data growth to ensure decisions can be executed in a timely, cost-effective manner. At the infrastructure level, this means converging storage and compute into a single layer to create scalability. Organizations should also seek to adopt consumption-based pricing, allowing infrastructure resources to be applied to the data that is most relevant to their needs, at the point in time it is necessary. Irrelevant or extraneous data can be quickly tagged and migrated to low-cost storage tiers.
- **Infrastructure and storage decisions must adhere to modern security and compliance requirements:** The value of new, data-oriented regulations such as GDPR is that they can help drive holistic change and serve as a catalyst for consolidation of inefficient primary and secondary storage technologies. IDC believes the implementation and enforcement of these new regulations create long-term opportunity for customers to reevaluate their infrastructure and data management needs and build a business case for investment in new solutions or services.

» **Adoption of advanced infrastructure solutions may be complicated by evolving IT responsibilities:** The lines between previously discrete IT activities are slowly eroding. Data protection is a clear case in point, where backup, disaster recovery, and high availability are evolving from disciplines to points on a continuum. This situation challenges organizations to develop an infrastructure strategy that increasingly converges storage, networking, and compute resources but also can be maintained by as few full-time employees as possible. Increasingly, these individuals are IT generalists who have experience across all parts of enterprise IT infrastructure (e.g., virtualization, networking, compute, storage). Furthermore, IT generalists are also tasked with elevating the role from discrete management and break/fix maintenance to value generation, made possible by advanced platforms and tools that provide cloud data management, migration, analysis, and protection. In addition, IT generalists are expected to develop and manage relationships with a number of service provider partners. This evolution toward IT generalists, and the subsequent demands placed on IT managers and administrators, may hinder adoption of complex infrastructure technology as implementation is executed by IT generalists with the help of third-party resources and partners.

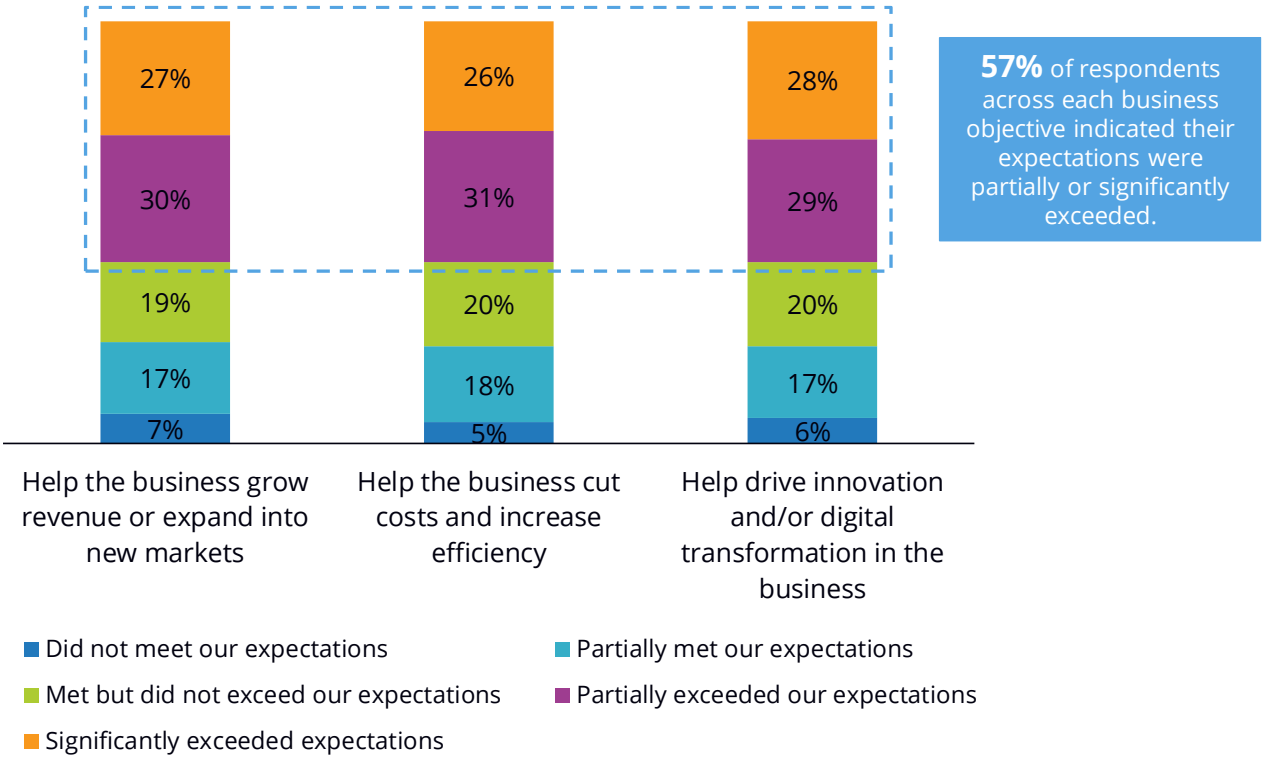
Key Trends

Despite the complexity of the factors outlined previously, IDC research shows that adoption of all-flash storage, software-defined storage, and cloud-based storage services is already well underway. The implementation of these capabilities acts as an accelerant for the modernization of storage workloads. The following market trends provide additional detail on the evolution of storage workloads due to advanced technology and modern business initiatives:

» A key enabling component of primary and secondary storage workloads is cloud infrastructure. IDC research shows that enterprises continue to purchase cloud storage capacity. IDC expects the infrastructure-as-a-service (IaaS) storage market to grow at a 34% CAGR to $19 billion from 2017 to 2022, significantly outpacing the growth of traditional storage segments. This growth of public cloud storage providers (e.g., AWS, IBM, Microsoft) continues to create opportunity for primary and secondary storage services providers to add value. In fact, IDC research shows that 40% of independent software vendors (ISVs) intend to partner with a public cloud IaaS or platform-as-a-service (PaaS) provider to deliver their SaaS offering in the next five years. Several IDC studies also show that end users are embracing a cloud-first (i.e., any mix of private/public, hybrid/multicloud) approach to infrastructure. According to IDC's 2018 *CloudView Survey,* the top 3 reasons to adopt public cloud are improved agility, improved security, and simplification and standardization of IT infrastructure and application platforms. The survey also found that cloud-based storage and compute workloads partially or significantly exceeded expectations for 57% of respondents across business objectives for costs, revenue growth in new markets, and innovation and/or digital transformation (see Figure 1).

Figure 1: ***Cloud-Based Storage and Compute Workloads Exceed Expectations for Users Across Multiple Business Objectives***

Q ***For those infrastructure storage and compute workloads that you now have on cloud, to what degree did you meet your business objectives?***



Unweighted n = 2,418 worldwide respondents

Source: IDC's CloudView Survey, May 2018

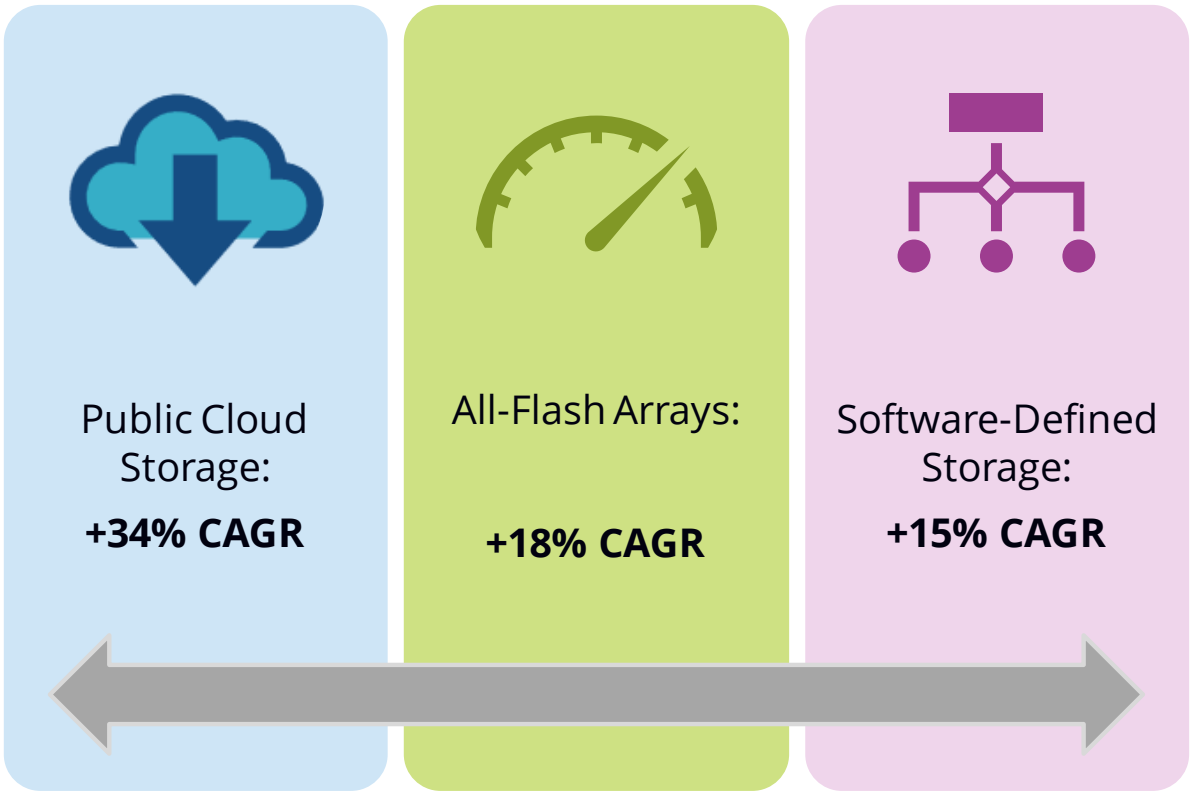
» The adoption of AFAs continues to expand into a broad range of primary and secondary storage workloads, including big data and analytics, technical computing, content and media streaming, backup, and disaster recovery. IDC forecasts that the AFA market will grow at a 17.6% CAGR to $14.9 billion from 2017 to 2022, much faster than hybrid flash and HDD-only segments of IDC's external enterprise storage systems market. As discussed previously, the adoption of AFAs is an important enabling factor for primary and secondary storage workload consolidation.

» Cloud repatriation is an increasingly important trend that will drive a mix of on-premises and cloud storage workload consolidation and modernization. While IT purchasers continue to deploy public cloud infrastructure at a rapid rate, they are also significantly increasing investments in on- and off-premises private cloud solutions where dedicated systems better address specific security, performance/latency, cost, and control requirements. According to IDC's 2018 *Cloud and AI Adoption Survey,* security concerns were the number 1 driver of cloud repatriation activities, ahead of performance, cost, control, and IT centralization. We believe this result is driven by security and compliance legislation associated with modern, data-driven workloads. Video surveillance is a prime example. Video surveillance and video analytics tools are increasingly used to improve services in manufacturing, law enforcement, and local government, but demanding analytics tools and

increasingly high-resolution images have an impact on the underlying storage infrastructure. Furthermore, retention requirements for video data continue to expand because of government legislation. This creates a security dilemma for many infrastructure administrators who are required to store increasingly high volumes of data for both primary and secondary workloads. Depending on the organization's needs, some data can be cost effectively stored in the cloud, while some must be kept strictly on-premises for security purposes. In many cases, this complexity will require flexible migration of stored data between environments, leading to ongoing repatriation and reallocation of storage resources.

» IDC forecasts that the software-defined storage market will grow at a 14.7% CAGR to $21.3 billion from 2017 to 2022. Within the SDS market, hyperconverged infrastructure (HCI) solutions are forecast to have the highest growth rate — achieving a 25.2% CAGR over the same forecast period. The growth of the SDS market in general and the HCI submarket in particular is an important indicator of expanding opportunity and customer readiness for primary and secondary storage workload consolidation. For many enterprise customers, HCI is relatively easy to expand — you just buy another node that already includes compute, memory, storage, networking, and software — and most block-, file-, and object-based systems offer nondisruptive expandability through the addition of more nodes. SDS also allows hardware choice, enabling customers to enjoy lower hardware costs when vendors must openly compete on price for COTS hardware. Furthermore, SDS makes storage management tasks much faster and more intuitive. Storage provisioning, a common operation that with many legacy storage infrastructures took several days and required the involvement of multiple groups, can now often be performed within hours by the same resource managing an enterprise's virtual infrastructure. All these factors contribute to storage workload modernization and consolidation (see Figure 2).

Figure 2: ***Key Technology Trends Driving Storage Workload Evolution***



Source: IDC, 2019

Conclusion

It is more vital than ever for organizations to reassess their legacy storage workloads in the context of modern business objectives — whether multicloud enablement, digital transformation, or cyber-security/resilience. Today, the combination of data and intelligence represents a unique opportunity to create value. IoT, mobile devices, and big data — combined with historical data, systems of record, and global information — continually sense an environment and put it into new contexts. Leading organizations are differentiated by the ways they leverage data to deliver value-added predictions and actions for personalized experiences, improved industrial processes, customer engagement, data monetization, and many other forms of enterprise decision making.

The increasingly important role of data within the modern enterprise means every organization needs to reassess the value of its primary and secondary storage workloads.

The increasingly important role of data within the modern enterprise means every organization needs to reassess the value of its primary and secondary storage workloads. Solutions for backup, disaster recovery, archiving, and cold storage should no longer be viewed as noncritical insurance policies or low-cost storage options. With the help of advanced infrastructure technologies such as cloud, SDS, and flash, the functionality of these secondary storage solutions is rapidly evolving. As a result, the lines between primary and secondary storage workloads are less discrete than ever before. This may create complications in the short term as organizations work to rationalize or modernize their storage infrastructure; however, over the long term, these new technologies and deployment models will be critical for organizations to capture, analyze, and retain growing volumes of data and successfully deliver modern data services to market.

About the Analyst

Andrew Smith, *Research Manager*

Andrew Smith is a Research Manager within IDC's Enterprise Infrastructure Practice, covering a broad range of storage research with a primary focus on archiving and data management software and services. Andrew contributes to market sizing efforts across IDC's storage software segments and is responsible for tracking the revenues of leading software and cloud services providers within IDC's Worldwide Storage Software and Cloud Services QView and Worldwide Semiannual Software Tracker.

MESSAGE FROM THE SPONSOR

Data Management by SHI International Corp.

SHI is a $10 billion global provider of technology products and services. With a dedicated, customer-centric approach, SHI helps customers select, deploy, and manage the technologies they need to meet their business objectives.

SHI has grown in the technology sector since 1989, and currently offers products, services, and solutions to assist in the strategy, design, deployment, and management of modern application platforms and infrastructure.

Today, over 15,000 customers rely on SHI as a trusted advisor to enhance their data management platforms. For more information on modernizing data protection with an industry-leading IT solution provider, visit ***www.shi.com/hybrid_cloud_data_protection***.

IDC Custom Solutions

IDC Corporate USA

5 Speen Street
Framingham, MA 01701, USA

T 508.872.8200

F 508.935.4015

Twitter @IDC

idc-insights-community.com

www.idc.com