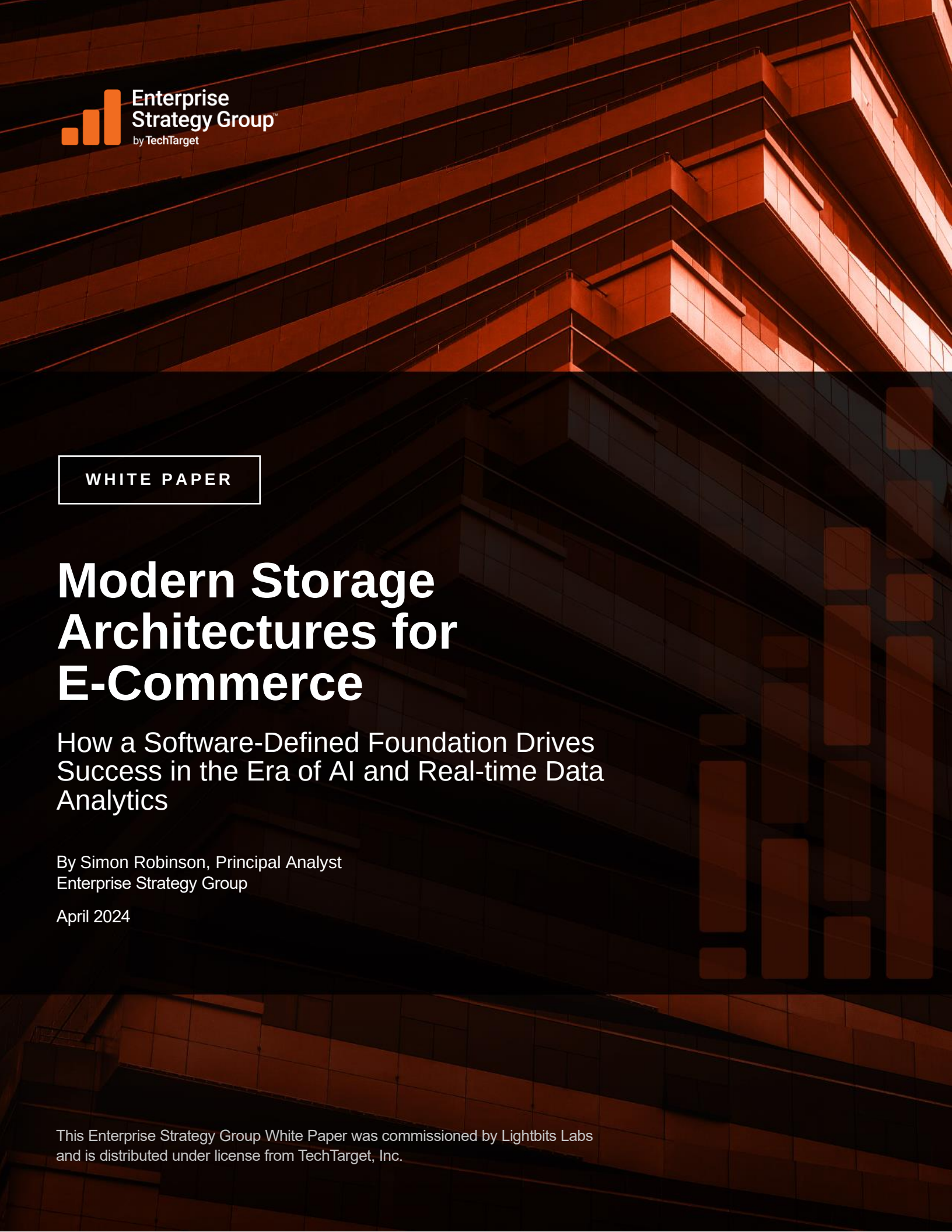


WHITE PAPER

Modern Storage Architectures for E-Commerce

How a Software-Defined Foundation Drives Success in the Era of AI and Real-time Data Analytics

By Simon Robinson, Principal Analyst
Enterprise Strategy Group

April 2024

Contents

Introduction

Market Overview – E-Commerce Companies at the Forefront of Innovation

There is no part of the shopping experience that has not been transformed by the tremendous growth of e-commerce since the turn of the millennium. From consumer electronics to fashion, furniture, health, media, and food, we now buy online first, chiefly from our smartphones. The reasons behind the astonishing growth of e-commerce are straightforward: Shopping online is easy, quick, convenient, and increasingly personalized to the needs of the individual. It has also led to a transformation in shopping and consumption: At-home meal delivery and ultra-fast grocery delivery are now common in urban areas, and innovation in new areas, such as AI and autonomous delivery, open up a whole new world of possibilities.

With such a large, innovative, and growing market, the opportunities for e-commerce providers are huge. Yet, as ever, retail competition remains fierce, and providers—whether online only or omnichannel—continue to run on often razor-thin margins.

This means that in the age of digital retail, technology choice becomes a—if not the—key determinant of success. A provider may live or die on its ability to pivot to tap into the latest trend, meet demand surges, and be first to differentiate. None of this is possible without a capable and agile underlying technology infrastructure, as well as a detailed understanding of and ability to act upon huge amounts of data. In such an environment, the maxim "innovate or die" has never been more apt.

All of this requires an innovative approach to infrastructure and data that is simultaneously agile, extensible, scalable, and cost-effective. In other words, a modern, future-ready infrastructure is simply a prerequisite for retail success.

Situation Overview – Why Data and Storage Infrastructure Matter

In such a competitive and dynamic space as e-commerce, there is one aspect that distinguishes the most successful businesses from the rest: *their ability to take advantage of data in real time or near-real time*. Harnessing data effectively enables merchants to understand both market trends and customers more quickly, allowing them to adjust strategy to offer products and experiences that are in tune with an ever-changing landscape.

This latter aspect is particularly relevant to e-commerce businesses because of the cyclical nature of retail. In addition to the enormous spikes in demand that are expected around holiday periods, including Black Friday, more unpredictable surges or changes in demand occur when disaster strikes (e.g., the COVID-19 pandemic, hurricanes) or when a new TikTok trend has every teen in the world chasing a new type of beverage.

This presents challenges to e-commerce providers because the task of right-sizing the environment is much easier said than done. Right-sizing for peak demand could mean massively overprovisioning the infrastructure for much of the year. Right-sizing for business as usual might mean missing out on business in key periods—and that may make all the difference.

Similarly, for many businesses, being data driven is good in theory but difficult in practice. Modern e-commerce businesses deal with enormous amounts of data—typically petabytes in scale. As well as the sheer volume, the variety and variability of data across the organization compounds this complexity. To drive insight effectively, businesses need to be able to combine the right data sets from multiple data sources (e.g., supply chain data, customer data, and logistics data) in one place, often in near-real time. The emergence of cloud-native, composite applications built on container technology such as Kubernetes often harvest data from multiple databases (a practice known as *polyglot persistence*), further compounding this challenge.

So, what of storage? As the retail and e-commerce world continues to modernize the application layer and user experience, the underlying storage can be a major bottleneck to progress. The core issue is that many legacy storage architectures were not built for the 21st century world of lightning-fast, turn-on-a-dime e-commerce, where data needs to be analyzed in real time or near-real time.

Challenges include:

- **Slow performance.** Storage systems often cannot meet the high performance (IOPS, throughput, and latency) requirements of modern e-commerce applications. This can be particularly challenging at peak times or when the business needs to rapidly analyze multiple large data sets to understand an emerging opportunity.
- **Difficult, complex, and expensive scaling.** Legacy storage architectures are often not built with "scale out" principles. The result can either be lengthy and costly controller upgrades to meet changing requirements, which may involve downtime, or massive overprovisioning in order to provide longevity and/or account for spikes in demand at certain times of year. Although some organizations have deployed direct-attached storage (DAS) or just a bunch of flash (JBOF) to cost-effectively boost performance, at scale, this becomes problematic in terms of resiliency as well as efficiency.
- **Data marooned in silos.** The fragmentation we see at the data level is replicated at the storage level, and organizations grapple with a multitude of on-premises systems, including storage area network (SAN), DAS, and network-attached storage (NAS). This may be further compounded when used in conjunction with cloud-based storage services potentially fragmented across multiple hyperscale and SaaS clouds.

Legacy storage architectures have not kept up with the scale and speed to process vast amounts of data in real time. What is required is a modern, cloud-like architecture that delivers the efficiency of hyperscale clouds and the high performance of direct-attached, flash-based storage systems.

Enterprise Strategy Group Research Findings

Modernizing the Infrastructure Stack for a New Era of Applications and Digital Services

Retail and e-commerce organizations currently utilize a range of infrastructure technologies and architectures that underpin their digital services. While some providers may utilize a range of on-premises/private cloud infrastructures, other, often newer, "born on the cloud" retailers run their infrastructure entirely in public cloud and SaaS environments. Additionally, companies are increasingly looking to deploy a software-defined hybrid cloud strategy that utilizes an optimal mix of on-premises and cloud-based infrastructure.

Regardless of their posture, many organizations are looking to modernize their infrastructure to stay ahead. According to research from TechTarget's Enterprise Strategy Group, 91% of retail organizations agree that data center modernization is strategic and can lead to a competitive advantage.[1] Modernization at the infrastructure level takes many forms, and research suggests many organizations are investing to modernize multiple parts of the estate.

At the storage infrastructure level, retail organizations are investing in three core areas:

1. Storage to support modern, cloud-native applications built on containers (40% of retail organizations).
2. Modernizing storage consumption models, such as "as-a-service" (26%).
3. Modernizing by consolidating storage to improve operational efficiency, total cost of ownership, and ROI; reduce technical debt; and consolidate sprawl; among other goals (46%).

[1] Source: Enterprise Strategy Group Complete Survey Results, *Navigating the Cloud and AI Revolution: The State of Enterprise Storage and HCI*, February 2024. All Enterprise Strategy Group research references are from this survey results set.

At the same time, organizations are aggressively embracing infrastructure for net-new apps in areas such as AI: Around half of organizations in the same study said they were actively evaluating or currently implementing initiatives focused on both generative and predictive AI.

Yet, there are challenges to delivering this vision: 68% of retail organizations say storage infrastructure demands and spending are hard to predict, 74% say infrastructure complexity is slowing down IT operations and digital initiatives, and 53% say their organization regularly encounters issues with visibility across all data.

Key IT Challenges for Retail Organizations

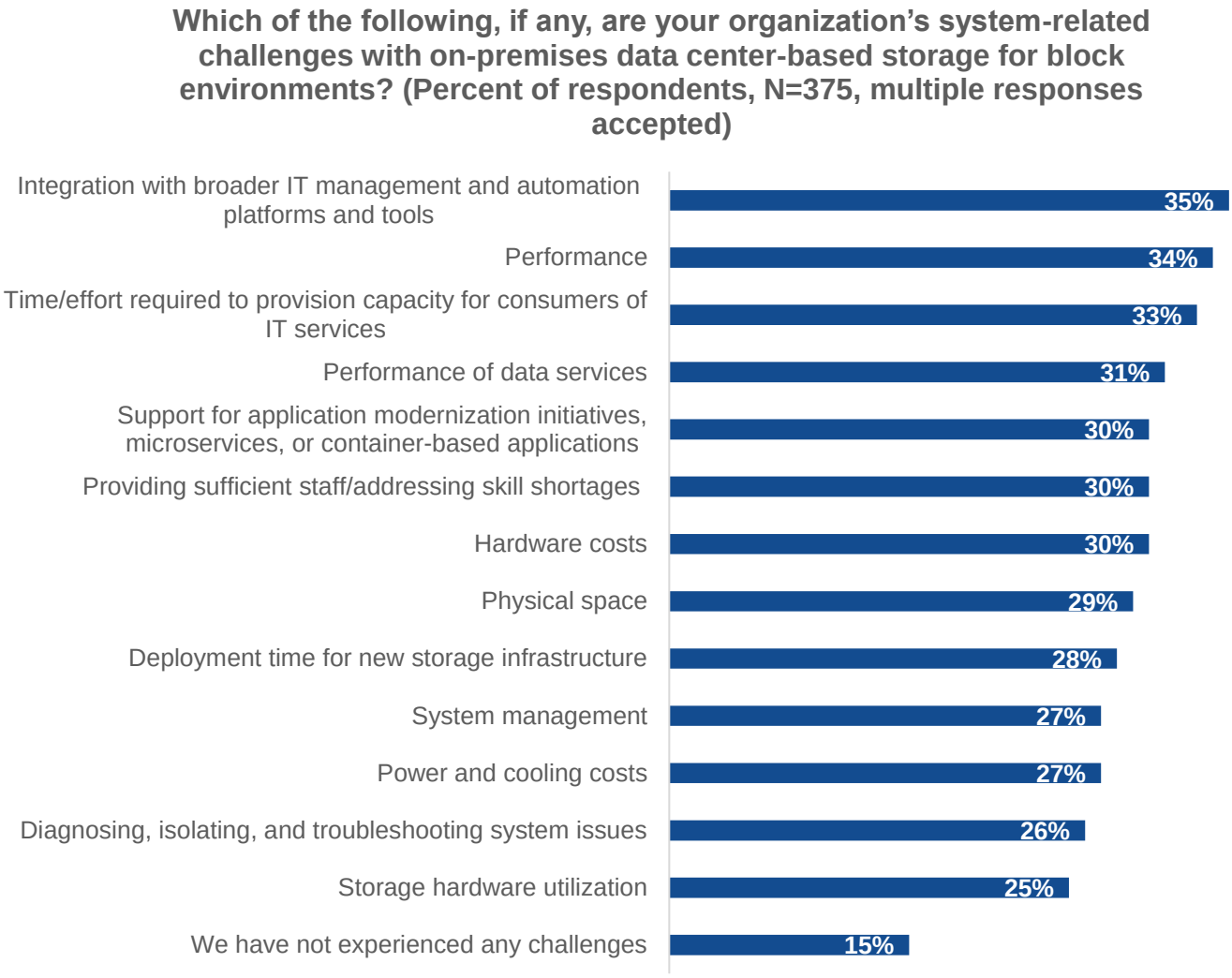
- Storage infrastructure demands and spending are hard to predict (68%).
- Infrastructure complexity is slowing down IT operations and digital initiatives (74%).
- Data visibility issues (53%).

On-Premises Storage/Private Cloud – Challenges and Opportunities

For retailers operating some or all of their IT environment on premises, legacy storage architectures can often be an inhibitor to business. Enterprise Strategy Group research suggests that the top three challenges with block storage at the systems level are integration with the broader IT management and automation platforms and tools (35%), performance (34%), and time/effort around provisioning (33%, see Figure 1). All of these factors combine to slow down the infrastructure, which could translate into poor customer experience, delay time to insight when running analytic workloads, and act as a hindrance to the e-commerce organization's ability to respond quickly and cost-effectively to meet demand surges.

These are not the only challenges, however. Many organizations are struggling to optimize their storage for modern, cloud-native applications built on containers/microservices, while hardware costs, poor storage utilization, and skills shortages are all problematic areas for many as well. In other words, the barriers to storage modernization are substantial. Figure 1 illustrates the most common block storage challenges that organization's face today.

Figure 1. Most Common Block Storage Challenges

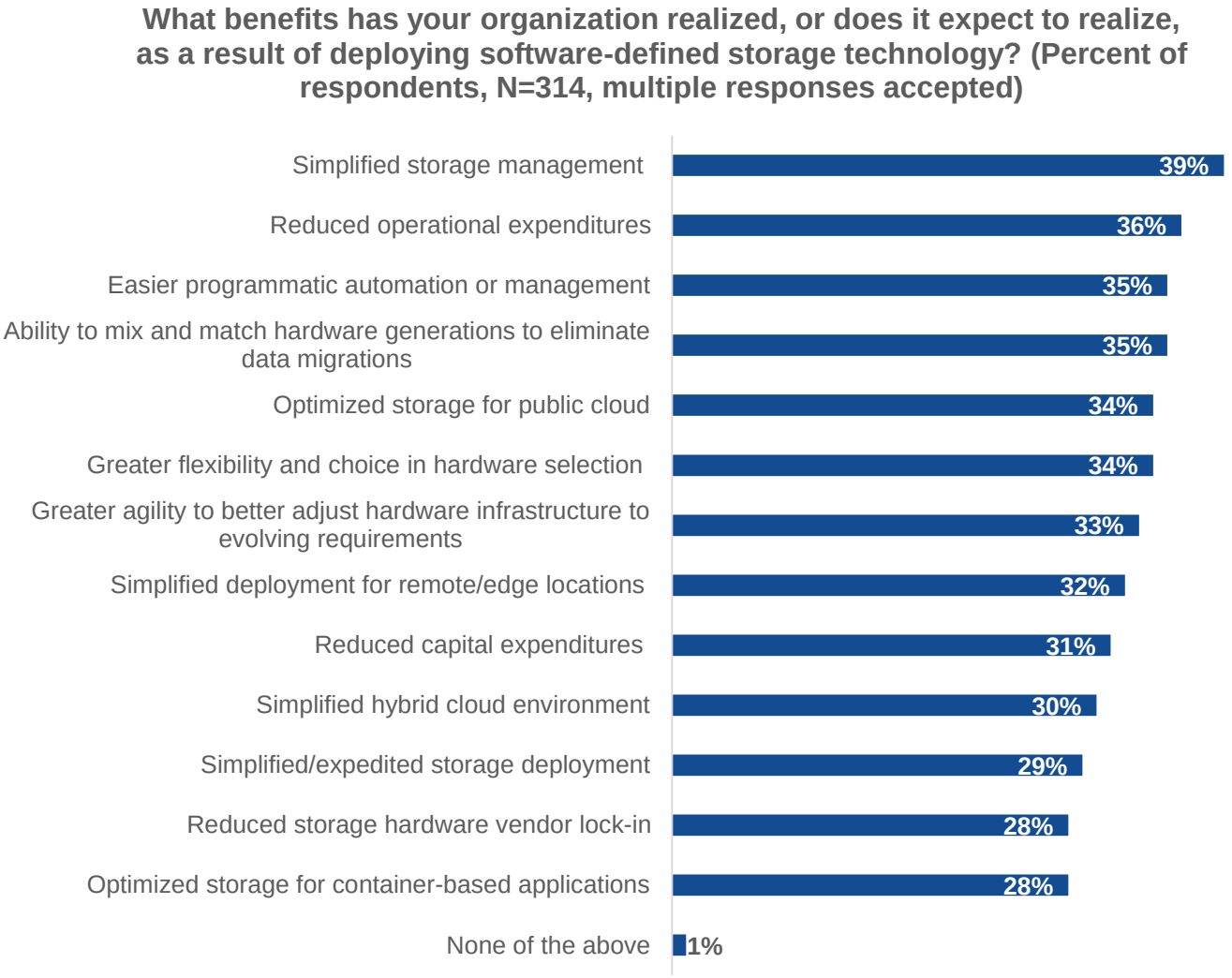


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

With these challenges in mind, many organizations are turning to alternative architectural approaches for their storage environments that mitigate some or all of these challenges. One increasingly popular alternative is to deploy software-defined storage (SDS). Though this is a broad category that incorporates many different approaches, the essential purpose of SDS is to divorce the storage control plane from the underlying hardware, moving it into a software layer that supports data portability and promotes more freedom of choice around underlying hardware and architecture, including more cost-effective options that don't hinder performance. This also acts as a basis for "scale out" approaches that enable organizations to add performance or capacity as required.

Enterprise Strategy Group research suggests that organizations are benefiting in multiple ways from deploying SDS (see Figure 2). The top three cited benefits of SDS are simplified storage management (39% of organizations), reduced Opex (36%), and greater levels of automation (35%). These benefits result in cost savings for organizations, as fewer or less-skilled administrators are required to manage the storage environment and, thus, are free to focus on more value-added tasks. Organizations deploying SDS also benefit from greater flexibility around underlying hardware choice, simplified integration with cloud environments, and greater flexibility in response to evolving requirements.

Figure 2. Benefits of Software-defined Storage



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Public Cloud Storage – Challenges and Opportunities

The emergence of the public cloud model has been a substantial enabler of innovation and transformation in retail and e-commerce. Indeed, the pay-as-you-go and "scale on demand" aspects of public cloud services such as IaaS are, in many ways, ideally suited to the highly fluctuating demands of retail organizations.

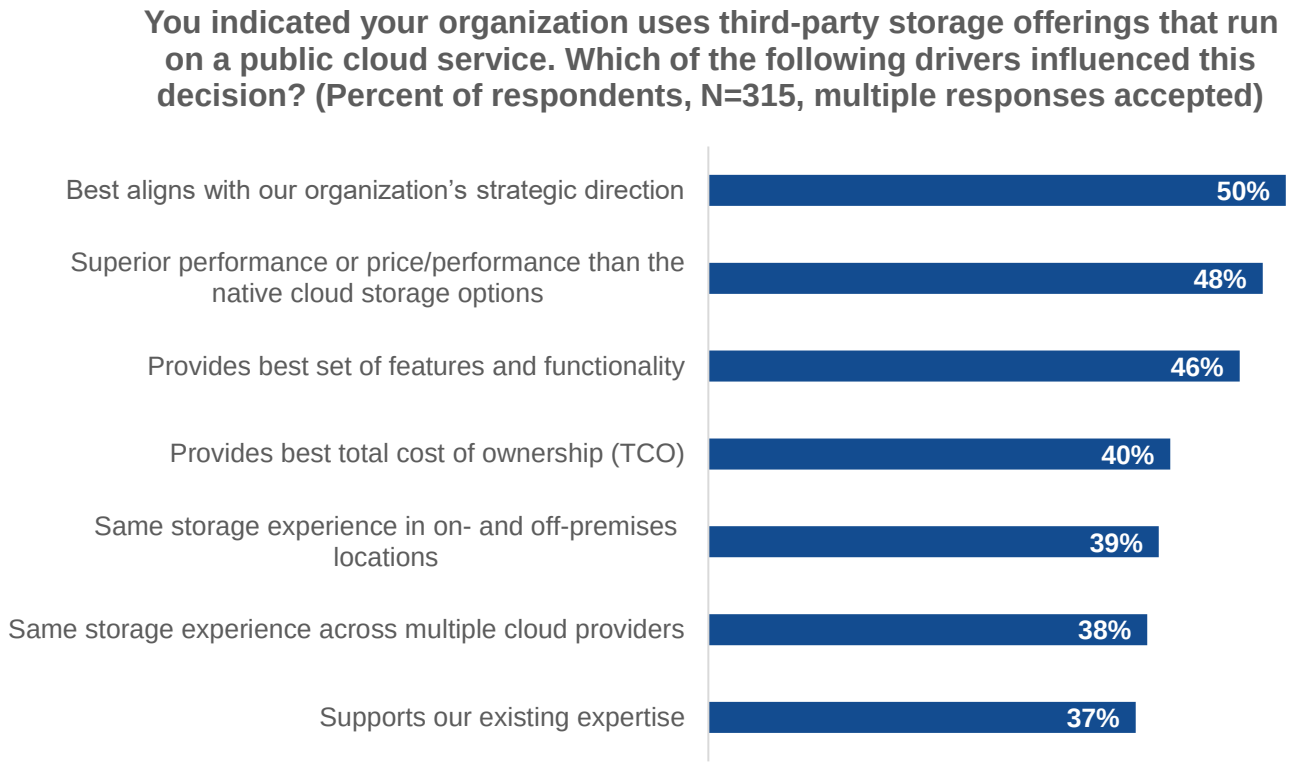
Nonetheless, many organizations are running into challenges as they continue to scale their public cloud investments. As shown in Figure 3, key challenges associated with cloud native block storage services include data security issues (cited by 34% of organizations), performance issues (31%), and limited data services (28%). Cost issues were cited by a quarter of respondents, while data ingress/egress were also considered a concern.

Figure 3. Challenges for Cloud-native Block Storage

Source: Enterprise Strategy Group, a division of TechTarget, Inc.

One way that organizations can address these issues is by implementing a third-party SDS solution that can run within or alongside a public cloud service. Such configurations enable them to utilize the cloud for certain compute services while tapping into storage capabilities that can scale independently. Research suggests that organizations utilizing such services have a range of benefits in mind. Some of these benefits include gaining performance and price-performance advantages, as well as tapping into an optimal set of features and functionality (see Figure 4).

Figure 4. Top Rationale for Using Third-party Storage in the Cloud



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

A further important benefit of SDS solutions attached to public clouds is that they enable organizations to build true hybrid cloud environments. In this way, e-commerce organizations can dynamically and optimally manage their infrastructure, both on premises and in the cloud, taking advantage of both as circumstances change.

Lightbits – A Software-Defined, Hybrid/Multi-Cloud Approach to Storage for Modern E-Commerce Organizations

Overview

Lightbits has developed a proven, innovative SDS solution for block-based storage that enables customers to build high-performance, agile, and feature-rich data environments at low cost. What's more, it supports any OS—both virtualized and cloud-native—and can run on premises, in the cloud, at the edge, or across a hybrid environment.

The core enabler of this is NVMe/TCP (non-volatile memory express over Transmission Control Protocol), a transport implementation of the NVMe over fabrics specification that the engineers at Lightbits created. Built on top of the standard TCP/IP software stack, Lightbits' SDS solution creates clusters of server-based SSDs in a highly parallel, highly scalable, low-latency fashion.

Lightbits as an Infrastructure Foundation for E-Commerce

Lightbits' innovative SDS approach is proving popular in a number of use cases. In particular, it has strong appeal among e-commerce merchants—those that have grown quickly and now have massive data environments that their legacy architectures, such as SAN and DAS, are struggling to support effectively and those running primarily in the public cloud that are looking to drive cost efficiencies.

The Lightbits SDS solution offers multiple benefits to such organizations, including:

- **High performance at scale.** Whether meeting demand spikes, running time-sensitive customer data analytics, or simply supporting a fast-growing business, delivering performance for e-commerce businesses is non-negotiable. Lightbits supports the creation of disaggregated storage clusters to deliver high performance and low latency—up to 75 million IOPS with sub-millisecond latency—making it an ideal solution for data-intensive SQL and NoSQL databases as well as large-scale virtualized, containerized, and latency-sensitive workloads.
- **Cost-effective infrastructure.** With e-commerce customers often running on thin margins, cost-effective infrastructure is a key requirement, and Lightbits offers 50% lower TCO as compared with public cloud, DAS, SAN, and HCI storage. It delivers this in multiple ways: For customers running on-premises infrastructure, Lightbits utilizes more cost-effective standard servers running SSDs, rather than more expensive SANs. Its approach is also more cost-effective to scale, and running on the TCP/IP stack doesn't require any specialized networking. Disaggregating storage from compute means no overprovisioning. For public cloud customers, there's no need to pay for unused resources (storage or compute), and reduced application licensing costs associated are with a more efficient environment. All customers can enjoy the benefit of space-saving features such as thin provisioning/compression to reduce costs further.
- **Enterprise-grade high availability (HA) and resiliency.** With e-commerce systems being critical to the business, it's essential that they are protected against downtime. Lightbits does this in multiple ways, whether running on premises or in the public cloud. Its clustered architecture offers built-in HA that can tolerate node or drive loss with no service disruption, as well as support nondisruptive software upgrades. It offers synchronous replication within or across zones as well as snapshots and clones. Support for these in Kubernetes environments brings enterprise resiliency to emerging, cloud-native workloads.
- **Fully integrated, compatible architecture.** E-commerce providers run a range of systems and utilize a number of public cloud services. Lightbits offers compatibility with almost any environment. For public cloud customers, Lightbits is available on the AWS and Azure marketplaces, is certified for the Azure VMware Solution (AVS), and supports all major databases, including Oracle, MySQL, PostgreSQL, MongoDB, and ElasticSearch. There's full compatibility for a range of on-premises platforms, including VMware, OpenStack, Kubernetes, and—crucially—Lightbits is able to support multiple environments in a single cluster. Plus, its use of standard TCP/IP networks and Ethernet network interface cards enables Lightbits to be installed with ease, with no proprietary client software. Its API-driven nature also enables Lightbits to be integrated seamlessly with a range of automation and other tools.

Conclusion

The performance and efficiency of underlying infrastructure has never been more important to e-commerce businesses in the fast-moving and competitive world of online retail. Infrastructure has to be agile and performant enough to cope with a dynamic and demanding range of data-intensive applications, and, for e-commerce providers, it needs to do this as cost-effectively as possible. It also has to be extensible enough to effectively handle a range of cloud-native application architectures as well as the coming generation of AI workloads and services.

As we have seen, many legacy storage approaches carry flaws—in performance, scale, cost-efficiency, and sometimes all three—that combine to affect online merchants' ability to capitalize on opportunities. In addition, organizations have struggled to capture benefits that arise from running a true hybrid cloud environment.

SDS solutions offer a way for organizations to optimize their storage without some historical hardware lock-in, and they can scale effectively and independently of compute. Lightbits' approach to SDS does this in a way that also delivers performance and support for key cloud-native platforms, either on premises or in the public cloud.

Historically, IT decision-makers have had to make too many compromises when developing their storage architectures. Performance and agility have too often come at a cost. Lightbits has designed its innovative approach to offer the performance and agility that e-commerce providers require in a cost-effective architecture. This is a powerful combination that makes it worthy of consideration for any growing e-commerce organization today.

About Enterprise Strategy Group
TechTarget's Enterprise Strategy Group provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

contact@esg-global.com
www.esg-global.com