The State of Software-Defined Storage, Hyperconverged and Cloud Storage

SIXTH ANNUAL MARKET SURVEY

DataCore™ SOFTWARE

TABLE OF CONTENTS

Click the section to jump ahead.

Introduction

For the sixth consecutive year, DataCore Software explored the impact of software-defined storage (SDS) on organizations across the globe.

In the 2017 ***Strategic Roadmap for Compute Infrastructure***, analyst firm Gartner states that "compute infrastructure is rapidly evolving from hardware-centric silos to software-driven ecosystems where application agility drives infrastructure architectures. I&O (infrastructure and operations) leaders must plan now to address these trends impacting computing to evolve and modernize their infrastructure for digital business.[1]"

For the sixth consecutive year, the DataCore Software annual survey explores the impact of major software-driven storage deployments within organizations across the globe. The survey distills the expectations and experiences of 426 IT professionals who are currently using or evaluating software-defined storage (SDS), hyperconverged and cloud storage to solve critical data storage challenges. The results yield surprising insights from a cross-section of industries over a wide range of workloads.

Please refer to the section on Survey Demographics for details on the size of companies, geographies, and vertical markets represented.

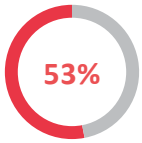
The Current State of Software-Defined Storage

Gartner predicts that by 2019, approximately 30% of the global storage array capacity installed in enterprise data centers will be deployed with software-defined storage or hyperconverged integrated system architectures based on x86 hardware systems, up from less than 5% today[2].

Research firm Neuralytix believes that "all organizations need to include software-defined storage (SDS) as part of their forward-looking IT planning." Additionally, Neuralytix estimates that an SDS investment is likely to last around 10 years[3].

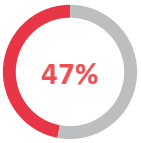
We asked respondents, ***"What are the business drivers for implementing software-defined storage?"*** The top responses were:

To **simplify management** of different models of storage.



To **future-proof** your infrastructure.

To **avoid hardware lock-in** from storage manufacturers.



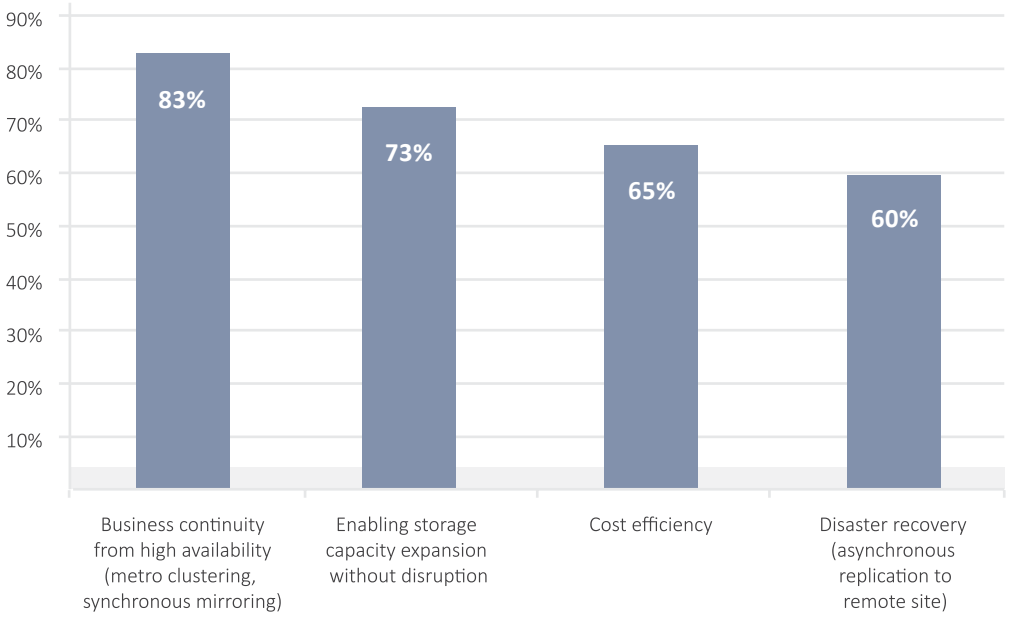
To **extend the life** of existing storage assets.

Interestingly, only 6% of those surveyed said they were not considering a move to software-defined storage. By comparison, the top reasons from our survey in 2015 for the same question were extending the life of existing storage assets and future-proofing the storage infrastructure, which topped the list with 52% of respondents naming this as a key factor; 49% stated that they looked for SDS to avoid hardware lock-in from storage manufacturers, while 45% chose simplifying management of different classes by automating frequent or complex storage operations.

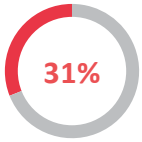
This year's survey is similar to last year's in that when compared with previous surveys, the results portray a major shift in the recognition of the economic advantages in acquisition and lifetime ownership cost (reduced CAPEX and greater flexibility in purchasing power) generated by SDS, versus the focus mainly on OPEX savings referenced in the earlier surveys.

Gartner's *Top Five Use Cases and Benefits of Software-Defined Storage* report stated similar business cases in its research with the statement that "I&O leaders are looking for software-defined storage (SDS) products that offer the potential for better total cost of ownership (TCO), efficiency and scalability to address exponential-data growth needs, and to benefit from innovations from hardware and software players independently.[4]"

We also asked our respondents ***"What are the primary capabilities that you would like from your storage infrastructure when virtualizing storage?"*** The majority of respondents (83%) replied that business continuity from high availability was the top concern (metro clustering, synchronous mirroring), while more than two thirds (73%) replied that enabling storage capacity expansion without disruption was a primary capability of importance. Cost efficiency and disaster recovery (asynchronous replication to remote site) also ranked high, coming in at 65% and 60% respectively. These top two priorities were also the same in past surveys, with cost reductions for greater infrastructure lifetime savings coming in third in past years.



However, respondents experienced some surprises after virtualizing mission-critical applications, with the top three surprises being:



Needed **shared storage** to make clusters highly available.

Application response time was **slower** than before virtualizing.

Determining storage requirements became **more difficult**.

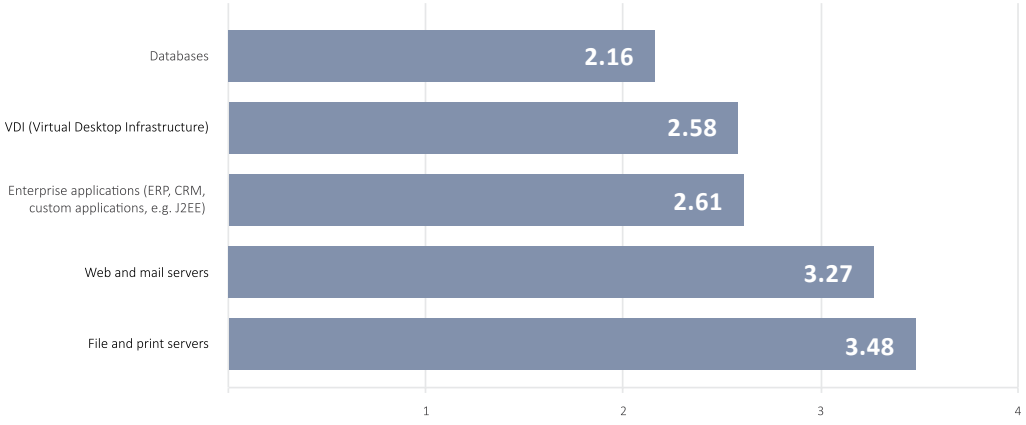
This year, 29% indicated application response time slowdowns after virtualization which compared to 31% in the prior survey. So while application response time has slightly improved, it's still in the same ballpark. In both cases, this impact is in large part due to the fact that hypervisors, operating systems and container virtualization treat I/O serially, making I/O the bottleneck on performance, even though application workloads are scheduled to run in parallel across several CPUs.

Technologies such as DataCore™ Parallel I/O are now helping to combat this issue. Rather than processing I/O requests serially, Parallel I/O performs multiple input/output operations simultaneously by leveraging available and largely idle multi-cores to drive and process I/O in parallel — resulting in faster response times and the ability to get more work done in the same time. According to Wikibon's Server SAN Readies for Enterprise and Cloud Domination report, "[Parallel I/O] is a good illustration of the new levels of performance that is being achieved by moving the storage much closer to the server.[5]"

A major trend in the industry is that 'servers are the new storage,' meaning that hyperconverged, hyper-scale, server SANs and conventional storage arrays are all becoming essentially off-the-shelf x86 servers doing the job of storage. The Wikibon report continues, "ISVs now have every incentive to write applications for server SAN, knowing that this technology can be brought to bear to solve many performance problems. They will be able to focus on creating more value within their applications for the cloud and for on premise systems, rather than constraining design to meet old storage array architectures[6]."

[*Note*: A server SAN is software-led storage built on commodity multi-core servers with direct attached storage (DAS)].

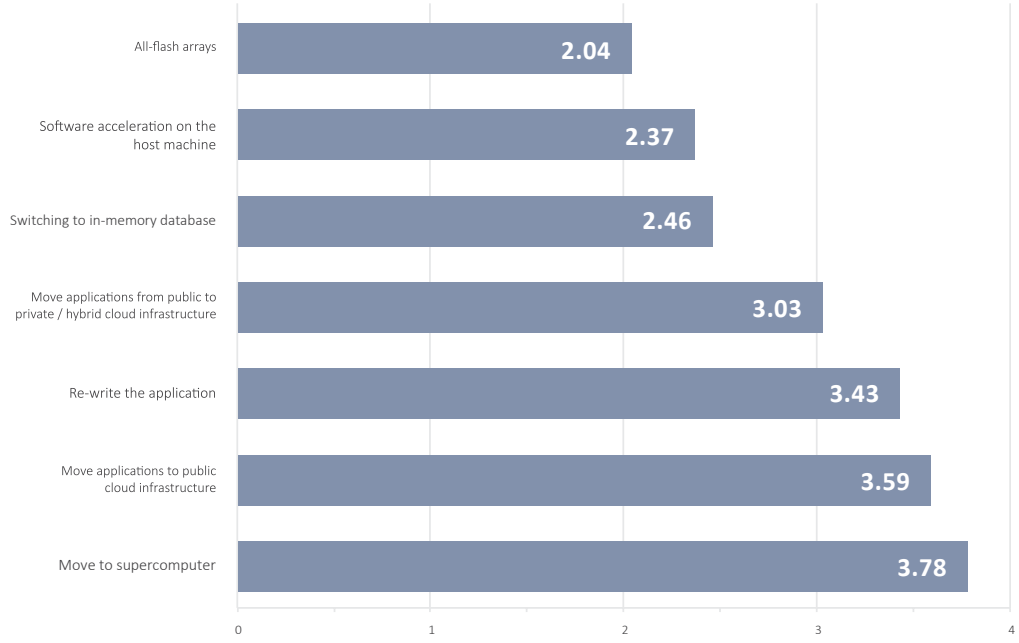
The following lists the environments that respondents ***believe experience the most severe performance challenges where storage is suspected to be the root cause*** (on a scale of 1-5, with 1 having the most problems, 5 the least problems):



The need for faster databases and data analytics is driving new requirements for technologies to optimize performance and meet demand for real-time responses. This is critical for business insights and to power technologies such as the Internet of Things. However, current technologies that accelerate performance and decrease latency also bring along significant disruptions to existing applications, greater complexity and higher costs.

To address these performance problems, the top approaches that respondents identified as preferred methods to overcome performance problems include using all-flash arrays; software acceleration on the host machine and switching to in-memory databases. The data for this was collected by asking respondents, ***"Which of the following approaches do you prefer to overcome performance problems, ranked in preferred order with 1 as most desirable, 5 least desirable?"*** The weighted average scores for each option appear below.

Which of the following approaches do you prefer to overcome performance problems?

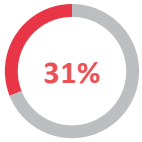


It is clear that flash technologies are now more prevalent and have helped solve many performance issues by replacing spinning disks with electronic media. From a media standpoint, this is goodness, and also brings along other cost-saving aspects such as smaller footprints and power requirements. However, the primary performance problem of processing I/O for high-velocity transactions, write traffic and latency-sensitive applications like databases has not been adequately addressed by technologies such as all-flash arrays and NVMe protocols. They have helped, but at the same time have added additional expenses and still don't fully take advantage of multiprocessing capabilities. This again ties to the problem that most computing processes work in serial, underutilizing the multi-core capabilities inherent in today's off-the-shelf server platforms. While in-memory technology has become "hot" for database applications (and while it's important to take advantage of faster memory speeds) once again, I/O processing and faster response aren't adequately addressed by this.

The industry hype would have us believe that customers will shift 100% to all flash. Qualitative and quantitative observations from surveys like this one suggest that's unlikely, due to the costs and transitional steps involved. There are other factors preventing organizations from making that move — including the realization that not all applications benefit from flash devices. See more detail on flash in other sections of this survey report, including 'False Starts' and 'The State of Flash Storage Adoption.'

False Starts and Technology Disappointments

We were also interested in uncovering where things had not gone as planned with new technologies. We asked respondents, ***"What technology disappointments or false starts have you encountered in your storage infrastructure?"*** The top three answers included:



Cloud storage failed to **reduce costs**

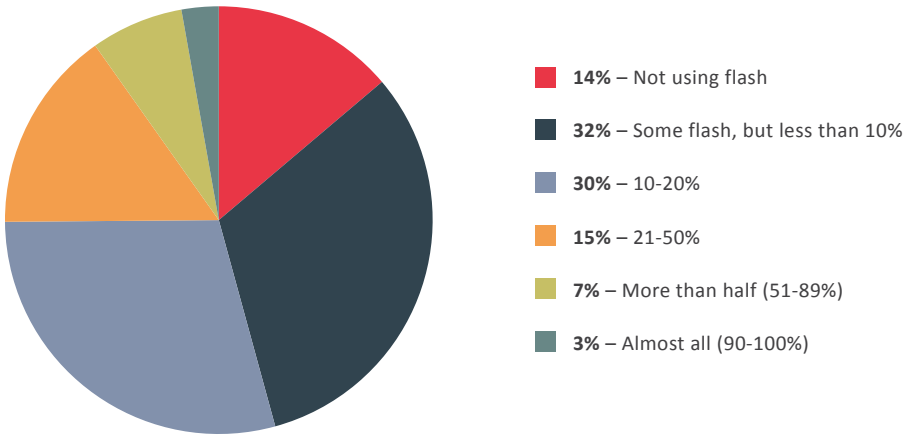
Managing object storage is **difficult**

Flash **failed to accelerate** applications

The DataCore 2015 State of the Industry survey uncovered the identical figure that 16% of participants felt that flash had failed to accelerate applications. Similarly, cloud storage failed to reduce cost for 24% of the respondents.

The State of Flash Storage Adoption

As one of the triggers to introduce SDS seems to be a direct result of difficulties encountered when integrating solid state tiers into the data center, DataCore continues to observe how flash is being adopted within the market. We asked participants, ***"Flash storage will account for roughly what percentage of your storage capacity in 2017?"*** The answers were:



Interestingly, as prevalent as flash is becoming, year after year we are seeing that only a small number of respondents have a large amount of storage capacity assigned to flash. About 60% of our audience falls within the less than 10% to 20% range of total capacity assigned to flash per the figures above – but we have seen definite growth patterns emerge over the years. In 2014, 63% of respondents had less than 10% of their storage capacity assigned to flash storage. 2015 saw the number drop to 53%, while 2016's results were that 32% had less than 10% of their storage capacity assigned to flash storage. This illustrates that flash devices are beginning to expand from a limited role inside servers to wider use within the storage infrastructure.

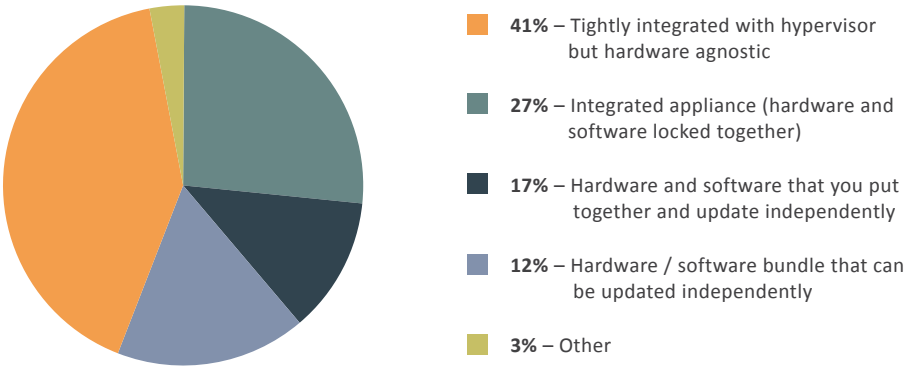
Customers have clearly not shifted 100% to all flash. Qualitative and quantitative observations from surveys like this one suggest that's unlikely due to the costs involved and the concern over their rapid wear. There are other factors preventing organizations from making that move — including the realization that not all applications benefit from flash. Flash is excellent for specialized 'hot data' workloads that require high-speed reads, such as databases. However, it is not a cost-effective solution for all workloads, and still accounts for only a very small fraction of the overall storage space.

The number of participants who answered that flash makes up more than half of their storage capacity was 10%. Nevertheless, the industry is beginning to realize a smart balance between flash and spinning disk technologies. Counter to the industry hype to go all flash, even companies like Google see the need for hybrid solutions. This is where SDS technologies such as auto-tiering optimize the cost and performance trade-offs by migrating workloads to the right mix of magnetic spinning disks or solid state storage.

Current State of Hyperconverged Infrastructure

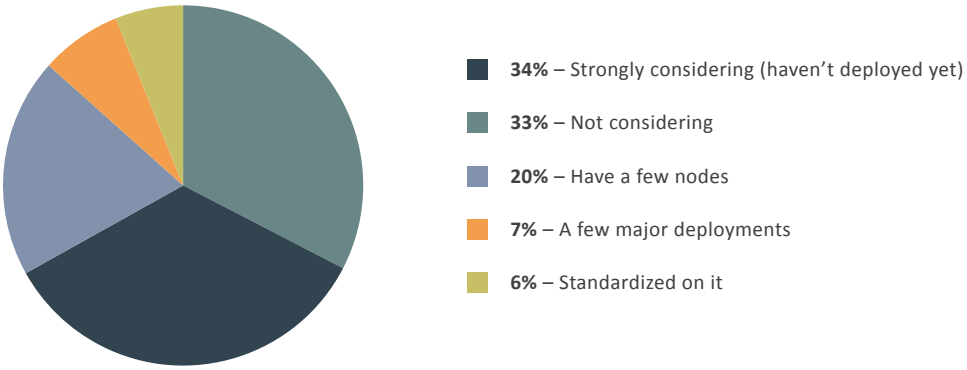
Many leading analyst firms, vendors, and other industry observers have slightly different views about what they classify as hyperconverged storage. For example, Gartner's definition includes "software-centric architectures that integrate compute, storage and networking on commodity hardware[7]"; IDC's is "hyperconverged systems collapse core storage and compute functionality into a single, highly virtualized solution. A key characteristic of hyperconverged systems that differentiate these solutions from other integrated systems is their ability to provide all compute and storage functions through the same server-based resources[8]"; and Forrester's: "An approach to technology infrastructure that packages server, storage, and network functions into a modular unit and adds a software layer to discover, pool, and reconfigure assets across multiple units quickly and easily without the need for deep technology skills. These systems can be implemented either as software plus modular physical units or as a software overlay on top of existing infrastructure[9]."

We asked our respondents what they thought hyperconverged means, and the results were:



What is most interesting from the above is that most analysts and vendor positioning assumes that the concepts of hyperconverged and appliance are synonymous, meaning that hyperconverged is not a pure software solution, agnostic to hardware, but that it must be sold as an integrated appliance. This survey suggests otherwise, with 41% of respondents believing that hyperconverged is software that is tightly integrated with the hypervisor but is hardware agnostic. In fact, only 27% of those surveyed felt hyper-converged required an integrated appliance, whereas in total, 70% saw hyperconverged as either hardware agnostic or offering the ability to update software independently of the hardware.

We have also all heard time and again that hyperconverged is a rapidly growing market, and it is indisputably so. IDC states that hyperconverged sales grew 148.0% year over year during the first quarter of 2016, generating $371.88 million worth of sales[10]. Gartner said the market for hyperconverged integrated systems (HCIS) will grow 79 percent to reach almost $2 billion in 2016, propelling it toward mainstream use in the next five years[11]. However, we wanted to drill down into what level of adoption people are currently at in their hyperconverged deployments, so we asked our respondents: "***Where are you with hyperconverged?"***

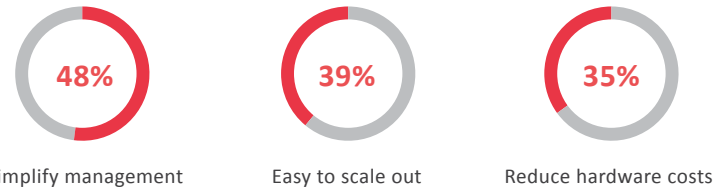


The results were somewhat surprising, with a majority of respondents either not considering hyperconverged at all (33%) or strongly considering it but haven't deployed it yet (34%). 20% of respondents said that they have a few nodes; 7% are in major deployment(s) while only 6% are standardized on it.

From additional customer discussions on the subject of hyperconverged adoption, while many hyperconverged systems work well for simple use cases such as a remote office or with virtual desktop infrastructures — where only a few nodes are needed — it's not always suitable for more challenging cases that require "real-time" response or depend on input-output (I/O)-intensive workloads like databases. Also, the same discussions revealed that while hyperconverged was a good solution for these simpler use cases, the lack of flexibility and the inability to work with existing storage and infrastructure also created hyperconverged 'silos' which could not be easily managed or integrated within the overall company infrastructure.

When high performance and fast response times are critical, many more nodes are typically required. In large enterprises with a mix of applications, that means more nodes to manage, which adds cost and complexity. In addition, some customers reported that the lack of capabilities such as fibre channel connectivity and performance make these systems unsuitable for database or more mission-critical applications.

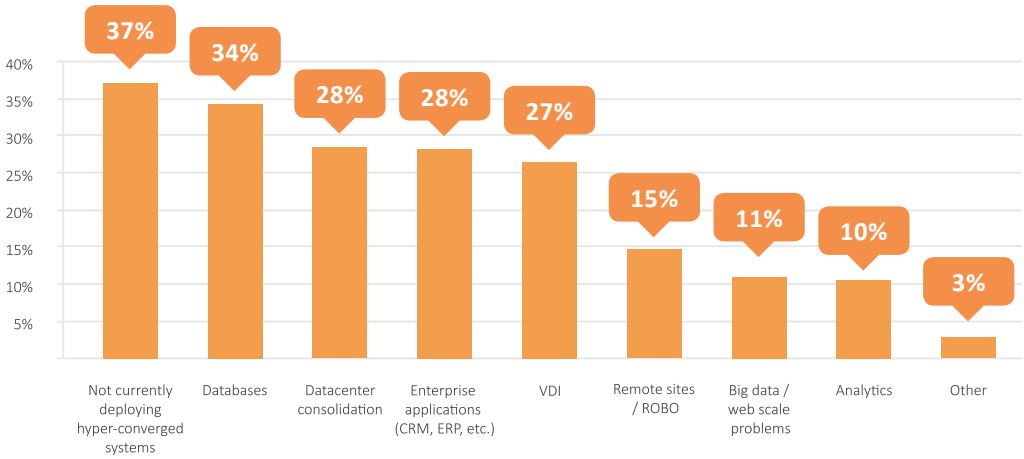
With any growing technology, we wanted to delve into the key business drivers that respondents gave for evaluating or deploying hyperconverged systems. We asked: ***"If applicable, what is the number one reason you are evaluating or currently deploying hyperconverged systems?"*** The top three reasons were as follows:



Note: Not currently evaluating or deploying hyperconverged systems - 27%

Clearly, simplification stands out as the main driver for users to choose to go hyperconverged. Additionally, we inquired into the use cases / applications that the market is primarily using for hyperconverged. Of the 33% who reported that they are using hyperconverged solutions, the summary of application use cases appears in the table below.

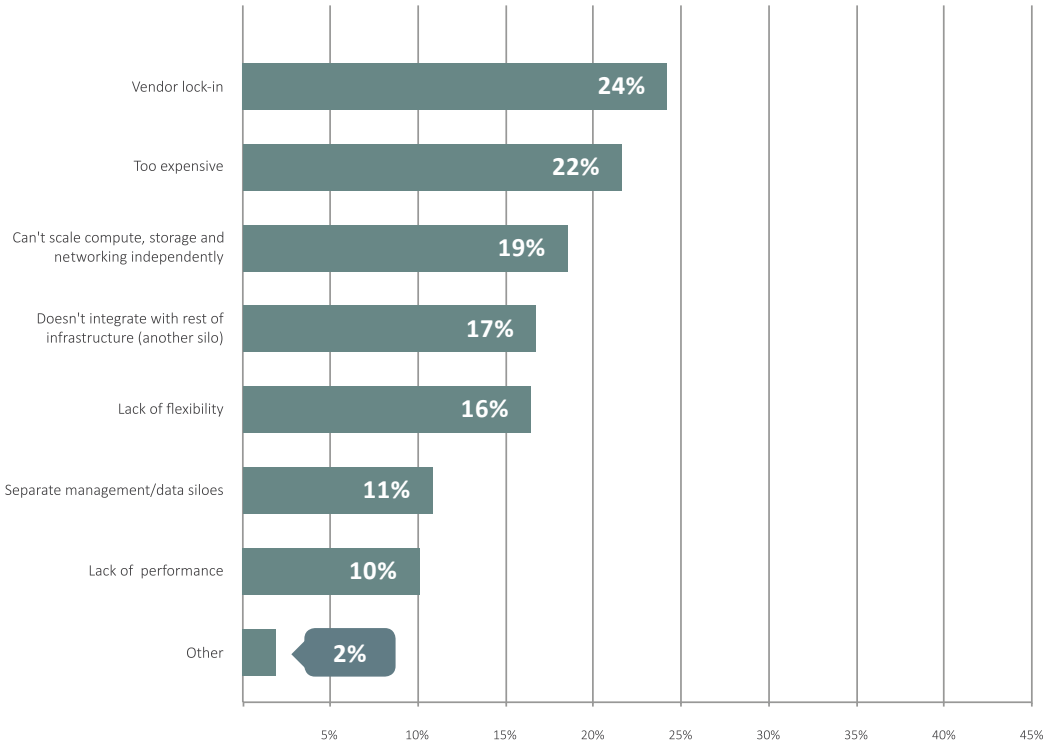
What Use Cases Applications Are You Using for Hyperconverged?



DataCore SOFTWARE

On the other hand, as noted, not everyone is on the hyperconverged bandwagon. 67% of our respondents said they were not even deploying it yet. We asked respondents ***"What is the number one reason you are ruling out hyperconverged?"*** The top reasons identified include vendor lock-in (24%), too expensive (22%), and can't scale compute, storage and networking independently (19%). The complete summary of responses appears in the table below.

What is the Number One Reason you are Ruling Out Hyperconverged?



Public Cloud / Hybrid Cloud Adoption and Use

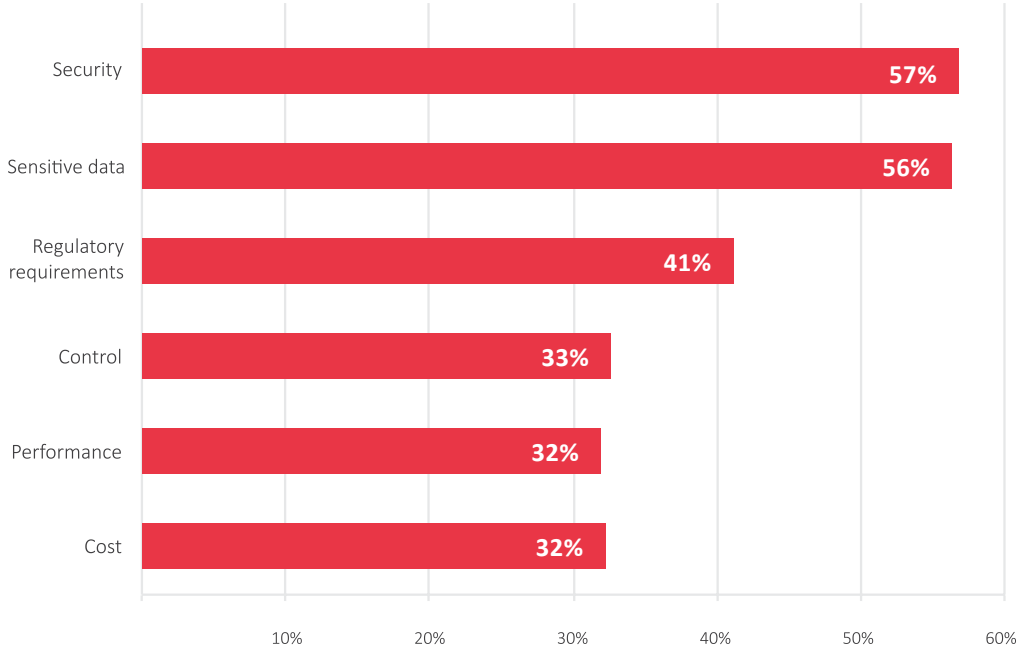
Similarly, the industry has witnessed steady growth in the cloud over the years. There are many reasons that enterprises are moving to the cloud, and we examined the types of use cases that organizations are currently considering for both public and hybrid cloud storage. First, we looked specifically at the public cloud, where there has been more hesitancy by enterprises, by asking: "***Which types of uses are you considering for public cloud for storage?"***

Not surprisingly, the top answer received was "not currently evaluating or using the cloud for storage" at 40%. But with any emerging technology, there will always be enthusiastic early adopters. The top three use cases identified by those willing to use the public cloud

for storage were: long-term archive (35%); back up to cloud and restore on premise (33%); and disaster recovery in cloud also at 33%. 11% also noted they were using public cloud storage for primary data storage.

We wanted to dig deeper into the reasons keeping people from migrating to a public cloud and asked, ***"For any applications you are unwilling to move to a public cloud, why?"*** Again, not surprisingly, security was the number one concern (57%) followed very closely by sensitive data (56%) and regulatory requirements (41%). The complete summary of responses appears in the table below.

For Any Applications you are Unwilling to Move to a Public Cloud, Why?



In looking at the application level for a hybrid cloud infrastructure, we explored what types of applications users were willing move to a public cloud or hybrid cloud infrastructure. Again the top answer was that for 47% of them, none of the applications are going to a public cloud or hybrid cloud, but 33% responded with select enterprise applications (i.e. Salesforce), followed by data analytics – 22%; databases – 21% and VDI at 16%. Many pointed out that they are being very selective on which workloads they would run on public cloud infrastructures based on security, cost and performance concerns.

According to Gartner, while cloud and software-defined technologies are mutually exclusive today, this will likely not be the case over the next few years. One big reason is the increase in hybrid data center deployments will require improved IT operations management interoperability and integration between software-defined resources within an enterprise and their counterparts within a public cloud.

Regardless of where specific applications are deployed and managed, Gartner states that the coexistence of private and public clouds as an infrastructure foundation for hybrid data centers is a highly likely scenario for many organizations[12].

A Look at 2017 Planned Storage Infrastructure Spending: Software-Defined Storage Tops the Charts

We asked respondents, ***"Which of these storage technologies will account for more than 1% of your budget in 2017?"***

Technology	1-10%	11% - 25%	More than 25%	N/A
Software-defined storage	43%	16%	13%	28%
Flash technology	49%	14%	11%	25%
Private cloud storage	33%	11%	10%	46%
Hyperconverged / VSAN	35%	13%	8%	44%
Object storage	35%	10%	6%	49%
Converged storage	32%	11%	6%	50%
OpenStack storage	21%	7%	2%	70%
Public cloud storage (i.e. Azure, Amazon)	37%	7%	5%	51%

Software-defined storage topped the charts in planned spending with 16% reporting that software-defined storage represented 11-25% of their budget, and 13% representing that it made up more than 25% of their allocated budget for storage. This was followed by flash technology with 11% responding that the technology made up more than 25% of their allocated budget for storage with 14% of survey participants putting flash in the 11-25% of total budget category. Technologies such as OpenStack storage are not making it into 2017 spending plans, with 70% of respondents marking it "not applicable" (the same exact result was seen in our 2015 research).

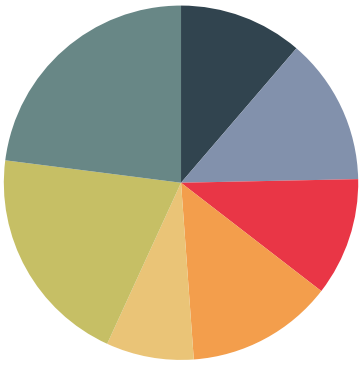
Conclusion

DataCore's State of Software-Defined Storage, Hyperconverged and Cloud Storage survey was designed to explore the impact of major software-driven storage deployments within organizations across the globe. The survey was conducted in late 2016 through April 2017 and distills the expectations and experiences of 426 IT professionals who are currently using or evaluating software-defined storage, hyperconverged and cloud storage to solve critical data storage challenges.

Respondents came from a diverse set of organizations, both in size and industry, providing statistically significant insights into the similarity in needs for software-driven storage over a wide range of IT environments. Participants were located in North America, South America, Europe, Asia, Africa, the Middle East, Australia and New Zealand in a range of vertical market segments including financial services, healthcare, government, manufacturing, education, IT services and other related industries. Respondents were from a mix of organizations including those with fewer than 500 employees, between 500 and 5,000 employees, and more than 5,000 employees.

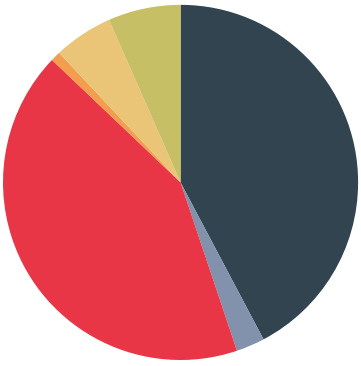
Survey Demographics

INDUSTRY



- **11%** – Financial services
- **13%** – Healthcare
- **11%** – Government (Federal, State and Local)
- **13%** – Manufacturing
- **8%** – Education
- **20%** – IT services
- **23%** – Other

GEOGRAPHIC REGION

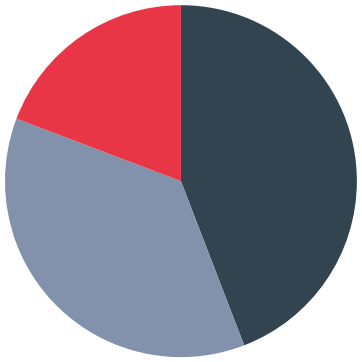


- **42%** – North America
- **3%** – South America
- **42%** – Europe
- **1%** – Africa
- **5%** – Middle-East and Asia
- **6%** – Australia and New Zealand

SIZE OF COMPANY:
(By Revenue)

- **31%** – Less than $10 million
- **30%** – $10M to $100M
- **23%** – $100M to $1B
- **16%** – More than $1B

SIZE OF COMPANY
(By Number of Employees)



- **44%** – Less than 500 employees
- **37%** – Between 500 and 5,000 employees
- **19%** – More than 5,000 employees

About DataCore

DataCore is a leading provider of software-defined storage and hyperconverged infrastructure solutions powered by Adaptive Parallel I/O technology, delivering higher performance, greater application workload productivity and cost savings. DataCore leverages the multi-core advances and cost efficiency of off-the-shelf x86 server platforms to overcome the IT industry's biggest problem, the I/O bottleneck. With DataCore, customers enjoy faster application response times and lower costs by making full use of their available computing resources to multiply productivity. The SANsymphony™ software-defined storage product pools diverse storage despite differences and incompatibilities among manufacturers, models, and generations of equipment. The software can span multiple locations and devices to bring them under the control of a common set of enterprise-wide data services for management automation and infrastructure simplification. DataCore Hyperconverged Virtual SAN software provides similar services using the internal or direct-attached storage spread across physical or virtual servers in a cluster.

The company has been privately held since its founding in 1998, and today has more than 10,000 customer sites across the globe. DataCore solutions are also available within turnkey appliances from hardware manufacturers including Lenovo.

Visit www.datacore.com or call (877) 780-5111 for more information.

Reference URLs

[1] Gartner, **2017 Strategic Roadmap for Compute Infrastructure**, Mike Cisek, George J. Weiss, Thomas J. Bittman, Andrew Butler, Arun Chandrasekaran, Chirag Dekate, February 9, 2017

[2] Gartner, **Top Five Use Cases and Benefits of Software-Defined Storage**, Julia Palmer, Dave Russell, April 1, 2016

[3] Neuralytix, **A Buyer's Guide to Software-Defined Storage (SDS): Establishing a Sustainable Data Strategy**, Ben Woo, May 22, 2016

[4] Gartner, **Top Five Use Cases and Benefits of Software-Defined Storage**, Julia Palmer, Dave Russel, April 1, 2016

[5] Wikibon, **Server SAN Readies for Enterprise and Cloud Domination**, David Floyer, August 28th, 2016

[6] Wikibon, **Server SAN Readies for Enterprise and Cloud Domination**, David Floyer, August 28th, 2016

[7] **Gartner Reveals Seven Myths for Hyperconverged Integrated Systems**, November 2, 2016

[8] **Worldwide Converged Systems Revenue Increases 11.0% Year over Year in the First Quarter of 2016, Surpassing $2.5 Billion in Value**, According to IDC, June 23, 2016

[9] **The Forrester Wave™: Hyperconverged Infrastructure (HCI), Q3 2016**, Richard Fichera and Naveen Chhabra with Glenn O'Donnell, Andrew Hewitt, Diane Lynch, August 16, 2016

[10] **Worldwide Converged Systems Revenue Increases 11.0% Year over Year in the First Quarter of 2016, Surpassing $2.5 Billion in Value**, According to IDC, June 23, 2016

[11] **Gartner Says Hyperconverged Integrated Systems Will Be Mainstream in Five Years**, May 5, 2016

[12] Gartner, **Survey Analysis: Software-Defined Data Center and Hyperconverged Adoption**, John P Morency, Scot MacLellan, October 4, 2016