An IDC Whitepaper

DECEMBER 2018

Digitalization of the Business with Object Storage

IDC ANALYZE THE FUTURE

Sponsored by HITACHI Inspire the Next

Digitalization and the Data Explosion Challenge

Digital transformation of enterprises has been the ongoing imperative for all enterprises in the last few years. The case for change is obvious. Customers are changing, technology is changing, competitors are changing, and business models are changing, and this is reflected by companies adopting digital technologies, with spending on digital transformation projects to soar past US$1 trillion in 2018.

Drivers of this change can be seen as external factors, such as millennials and their desire for immediate personalized fulfilment, government regulations and industry compliance on data management, such as Health Insurance Portability and Accountability Act (HIPAA), as well as rapid urbanization and environmental concerns. Furthermore, we have companies seeking to improve profitability, grow market share, and reinvent themselves. There is also disruption in a number of industries in which incumbents are being forced to change because of new entrants coming in with a different business and operating model.

Fundamental to driving the change is the reliance on data. Digital transformation is defined as "transforming decision making with technology." The focus on data is key, in which data is used to automate actions, predict outcomes, and create digital twins of products, processes, and people. This has resulted in a data explosion with forecast data to reach over 163ZB by 2025.

This explosion in the growth in data, whilst forecast by IDC over a number of years, has still taken many organizations by surprise. Clearly the ability to predict these growth rates has been severely under-estimated leaving many enterprises with severe data management challenges. At the same time, the nature of applications and their lifecycle is radically changing, and traditional methods to define storage requirements appear not to be up to the task any longer. Customer-facing applications can balloon out of expectations due to the global nature of business today, bringing with it significant compliance and resiliency challenges. So too, in the B2B environment, partner- and customer-facing applications are expected to have the same characteristics as pure consumer applications, requiring significant changes to the underlying network and storage infrastructure.

Object storage is one such technology that can help infrastructure managers take control of this escalation environment. Designed for the cloud-generation, and imbued with similar properties, object storage should be adopted by organizations to help enable the agility that businesses need in this digitally transforming environment.

Traditional methods to define storage requirements appear not to be up to the task any longer.

As far back as 2014 IDC identified that unstructured data growth would outstrip that of structured data, driven largely by different formats being used within the workspace such as images, videos, audio files and presentations. This growth is set to continue as new types of applications result in an increase in both data consumption and data production, resulting in ever more esoteric ways to present this information to decision makers. At the same time, the blending of business and social lives online creates a situation where many internal storage administrators find themselves managing data that is sourced from the Internet and their own employees. Whether for business purposes, or sometimes personal, the blurring of lines between personal and enterprise IT sees many smart phones being backed up to enterprise laptops, which in turn back up to the enterprise storage systems. Unless organizations create policies to handle this situation, the explosion on data could inundate less organized enterprises.

However, the biggest culprit of the data explosion of structured and unstructured data is *enterprises*, not consumers. As enterprises undergo their digital transformation, they digitize more and more of their business processes, products, and services. Transforming into a digital native enterprise (DNE) demands more reliance on the highest quality and most relevant data. With all their business processes, customer/supplier engagement, and products generating data, it is not surprising that a large proportion will be unstructured. This creates two major challenges — how to store the data in an efficient manner and how to consistently extract value from that data.

IDC defines a DNE as "an entity that is able to scale its business and innovate at a pace that is an order of magnitude greater than traditional businesses. It is driven by a customer-centric and empowered workforce that embraces risk taking as it seeks to continuously innovate. Technology and data are its lifeblood, fueling more efficient operations, new revenue streams, and customer loyalty."

Enterprise's Business and IT Initiatives for Data Storage

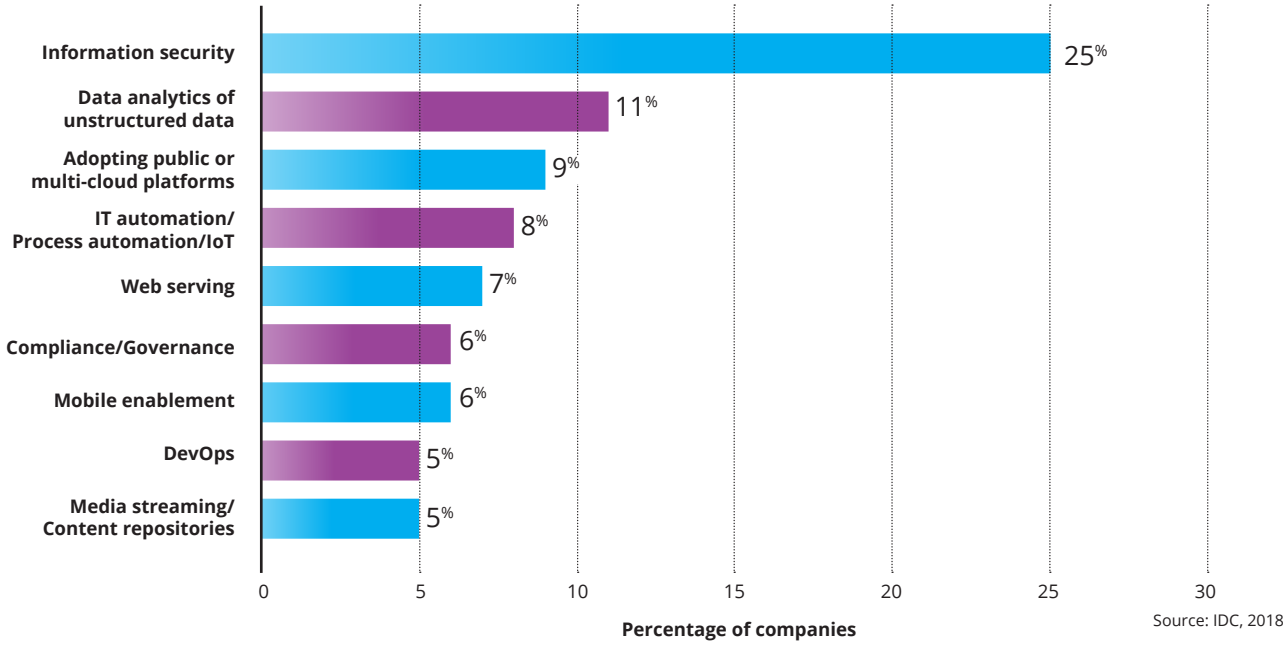
When we spoke with over 500 enterprises across Australia, Singapore, Hong Kong, China, and India, they told us that their top three IT priorities related to storage are: information security, analytics of their unstructured data, and adopting a public cloud/multi-cloud platform.

Securing data has become a top priority for many organizations, specifically this year, with the introduction of a range of local and trans-national legislation that requires organizations either to be able to report when a data breach has taken place and, in the case of a number of new regulations, face potential monetary fines for not securing this data sufficiently. However, this need to secure data is driven by a need to be compliant to legislation but does not drive much in the way of positive economic activity. At the same time the impact of ransomware on organizations that get hit with this malicious attack is immediately seen as a financial challenge. Either the (not advisable) cost of paying the ransom or the cost of not being able to access critical data has an immediate impact on business organizations.

FIGURE 1

Top Enterprise Storage Priorities

Question: What are your top 3 IT priorities related to storage?



Source: IDC, 2018

The real value of data management comes with the ability to be able to mine nuggets of information from the data, both held and managed locally, as well as a range of data sources that can be subscribed to. In the case of the ever-growing unstructured data, this poses a fresh set of challenges to being able to understand the data types, where this data is located, and what tools can be used to extract the information therein.

Consider the wealth of useful information that is exchanged across email systems. Both the content of the message and the attachments have the capability to yield significant insight into the business operations if the ability to store all this data in a single repository, and then analyze the various file attachment formats as well as the email content. To add to the complexity here, many organizations have moved to cloud-based email services, as well as CRM, ERP and a range of critical business systems that in the past had been on-premise.

The reasons for this move toward cloud is fundamentally based in economics, both in terms of being able to run leaner IT systems, but also the ability to re-deploy resources and respond to changing business needs much more rapidly than the traditional on-premise mode. Cloud offers a degree of automation that on-premise cannot (consider the need to scale storage and systems for peak workloads, this is now a "legacy" issue since cloud provisioning is a relatively easy operation) allowing IT resources to focus on the roles that deliver greater value. No longer should a storage administrator be so concerned about whether the resources are available and reliable, but can refocus on what data is where, and what value could that have to the organization.

Finally, we have observed that customers are not looking for a single environment strategy in their data retention today. Where data has identifiable value to the organization, or a certain level of sensitivity, customers will probably want that information in a private cloud they can manage. Whereas information that has been identified as having limited value, or free of information sensitivity, could be tiered off to ultra-long-term retention in the cloud. The critical part here is that the approach is not silos in themselves, rather they complement each other, and data can be moved back and forth across each as necessary.

In this region, unstructured data is generally retained for a period of at least five years, and even longer for organizations that must adhere to strict regulatory requirements.

Although about a third of unstructured data is stored long term today, this is expected to grow with the increasing digitization of the enterprise and changes to the regulations and the need to be compliant. Also, about 4 of 10 enterprises do not actively purge their storage of data that is of little value, leading to an accumulation of unstructured data and a growing level of data exposure risk.

The type of unstructured data stored will largely depend upon the industry that the organization is based as well as the focus senior management have on digital transformation. In some markets, healthcare data has to be kept for the entire term of an individual's life; for many financial markets, seven years of data has to be kept.

Organizations that understand the value of data will be storing certain types of relevant data for longer, perhaps not in its original format, but in a format that permits future analysis of this data. Consider the oil and gas industry where new computing power is allowing data that has been stored for decades to be reviewed, with some new findings.

FIGURE 2

Average Retention Period of Unstructured Data

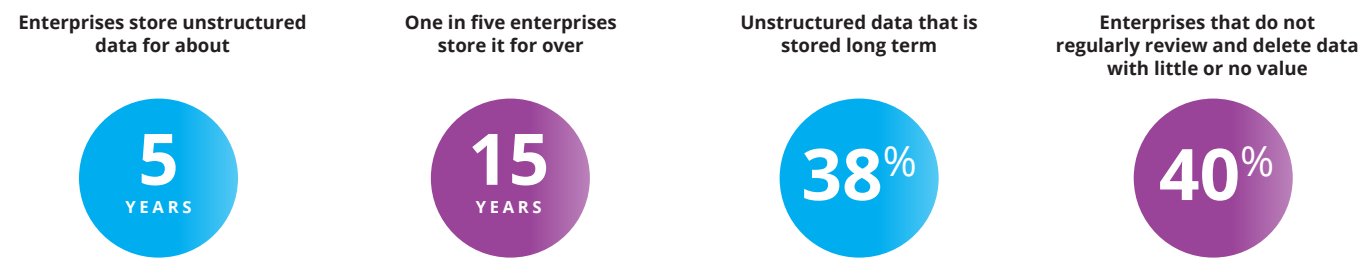
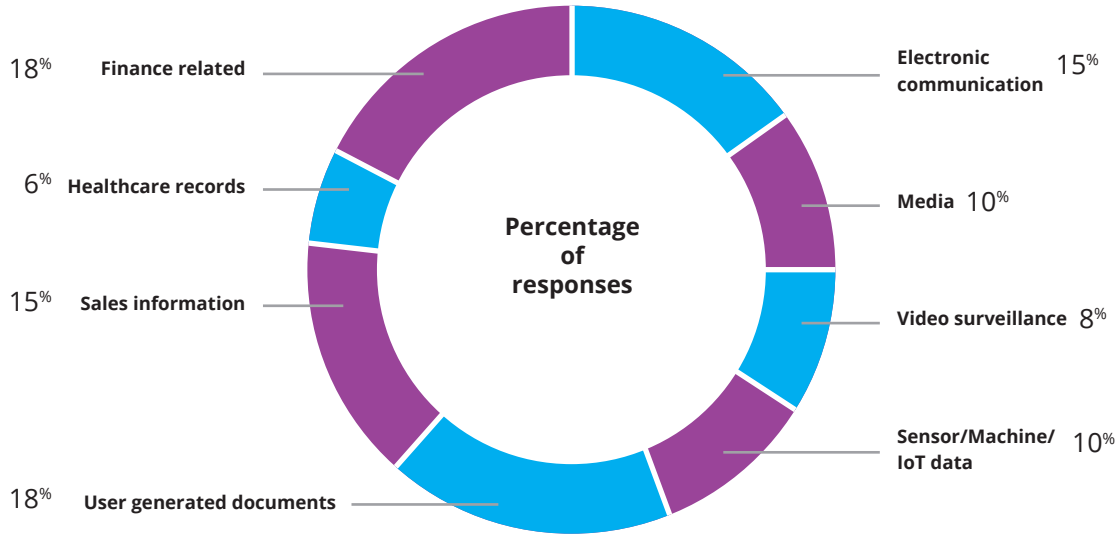


FIGURE 3

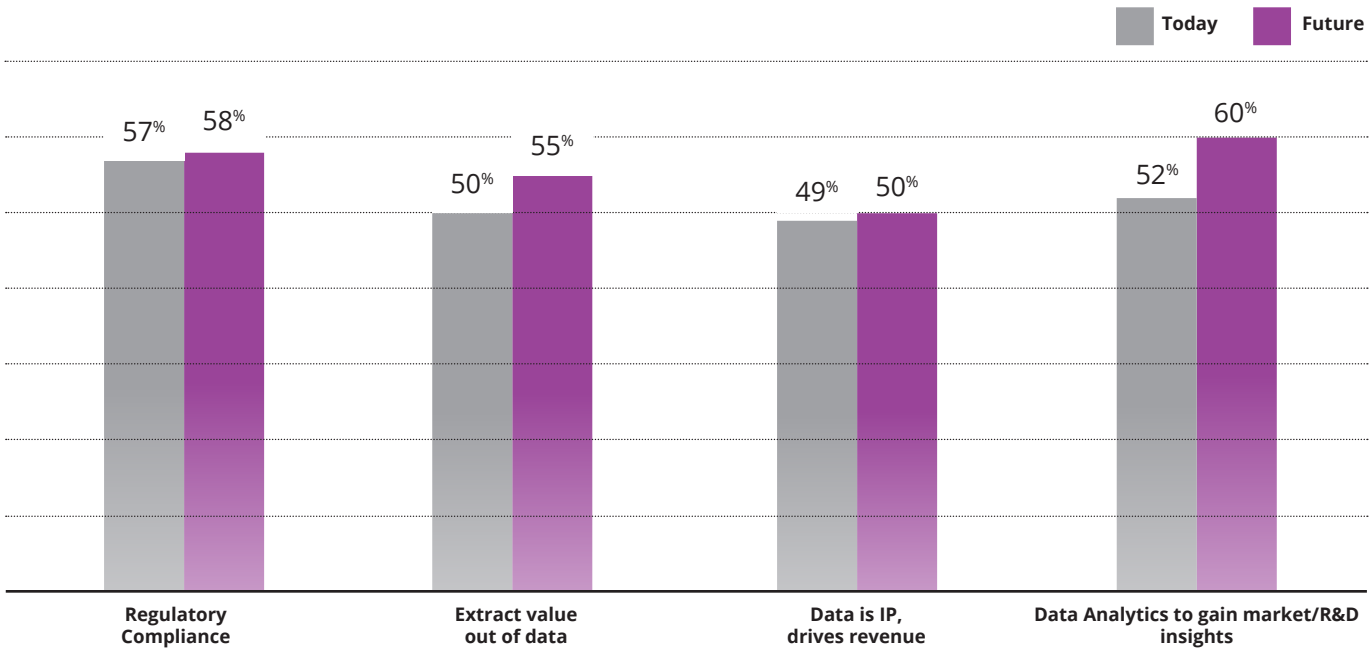
Types of Unstructured Data Stored Long Term



Source: IDC, 2018

FIGURE 4

Purposes for Storing Unstructured Data Long Term



Source: IDC, 2018

Some types of data do need to be purged occasionally. Some regulations demand that personally identifiable information that is no longer needed by the company should be purged and that this need is defined by any current engagement, not merely a desire to keep the data "just in case". The challenge here is generally, "where does this data exist within our environment?"

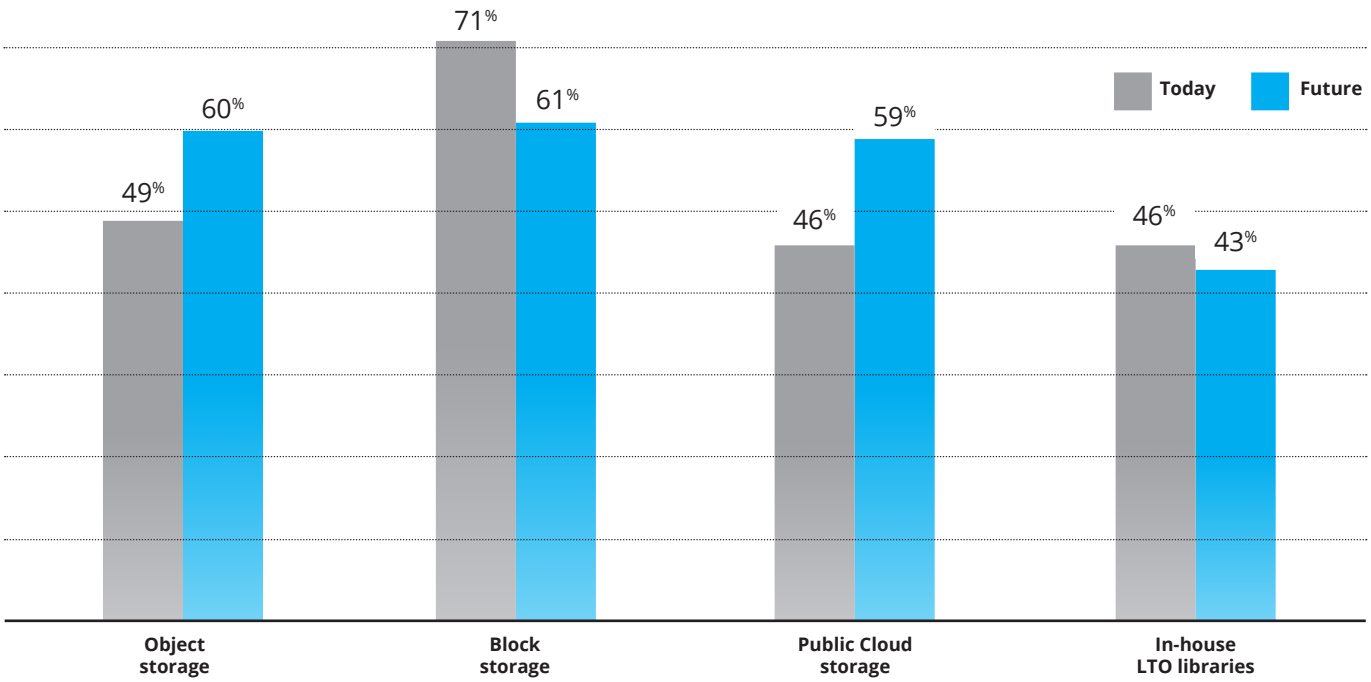
In the past, keeping unstructured data was mainly for regulatory compliance and record keeping, but more and more enterprises are keeping unstructured data for analysis to extract value, gain insights for market and product/services development, and sell for revenue. This requires a different approach to keeping unstructured data for long-term access and use, an approach that addresses both scalability and efficiency demands that align with data growth, and the performance and throughput requirements required for data analysis and processing. In order to achieve this goal of keeping both structured and unstructured data for longer periods, a different approach needs to be taken to the underlying storage technology that is used. One that addresses the needs of many data types, one that is architecturally aligned to the multi-cloud environment we now live in, and one that can scale on demand as this data continues to grow.

Enterprises today are using mainly traditional block storage to store their unstructured data. However, the percentage of enterprises relying on block storage for unstructured data is expected to decline in the near future with the adoption of object storage platforms. Given that object storage presents a more efficient and scalable platform for storing unstructured data, adds the ability to enrich the data with custom metadata, and is growing in popularity for analysis-based use cases, it stands to reason that using a more appropriate tier to balance the data center infrastructure is the more optimum approach for those who wish to create a data foundation capable of sustaining the daily business and supporting digital transformation efforts.

FIGURE 5

Platforms for Storing Unstructured Data

41% of enterprises surveyed are not aware of object storage. For those who are aware, about 49% are using object storage for unstructured data, with this growing to 60% in the future



Source: IDC, 2018

Object Storage Versus Block/File Storage

Object storage (also referred to as object-based storage) refers to a storage architecture in which data is organized into units, called objects, comprising the following three things:

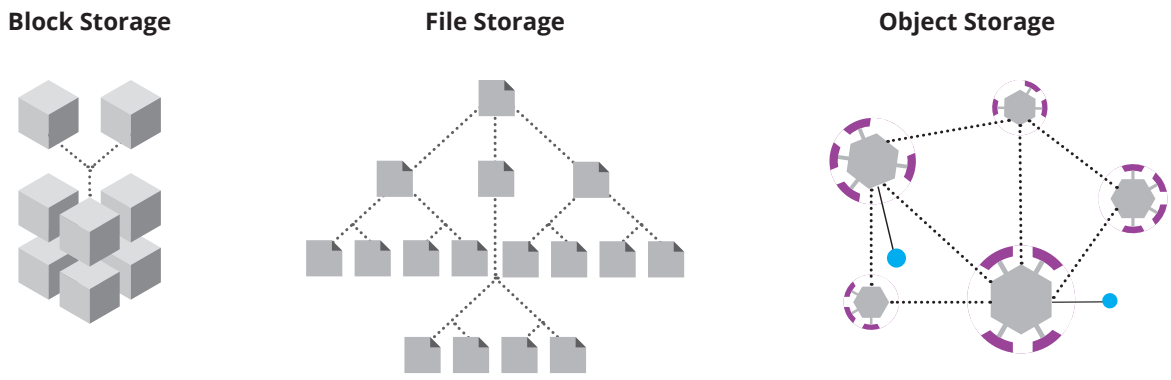
- **The data**. This is the digital content that is stored from electronic communications, user documents, photos, videos, audio, IoT data, and so forth.
- **The metadata**. The metadata is defined by the object storage creator and contains contextual information about what the data is, what it should be used for, its confidentiality, or anything else that is relevant to the way in which the data is used.
- **The unique identifier**. This is an address given to the object so that it can be found over a distributed system. This is to enable us to find the data without having to know the physical location of the data.

Block storage is a system whereby data is stored in volumes, also referred to as blocks. This is the traditional method for capturing and storing data from applications that typically require consistently high performance.

File storage is a relatively easier and less expensive technology to deploy than block, which also comes with metadata and directories to organize files. Over time, however, this simplistic approach can lead to challenges in performance.

FIGURE 6

Concept Diagram



Source: Hitachi Vantara, 2018

Why Choose Object Storage

Object storage doesn't split files into raw blocks of data but stores entire clumps of data in an *object* that contains the data, metadata, and the unique identifier. There is no limit on the type or amount of metadata, which makes object storage powerful and customizable. This approach can massively reduce the effort required to identify personally identifiable data that needs to, by law, be deleted.

Object storage is a more efficient technology for solving the increasing problems of data growth. As more and more data is generated, storage systems have to grow at the same pace. If we were to expand a block-/file-based storage system beyond 100TB or multiple petabytes, we would run into hard limitations with the storage infrastructure, durability issues, or escalating storage management overhead.

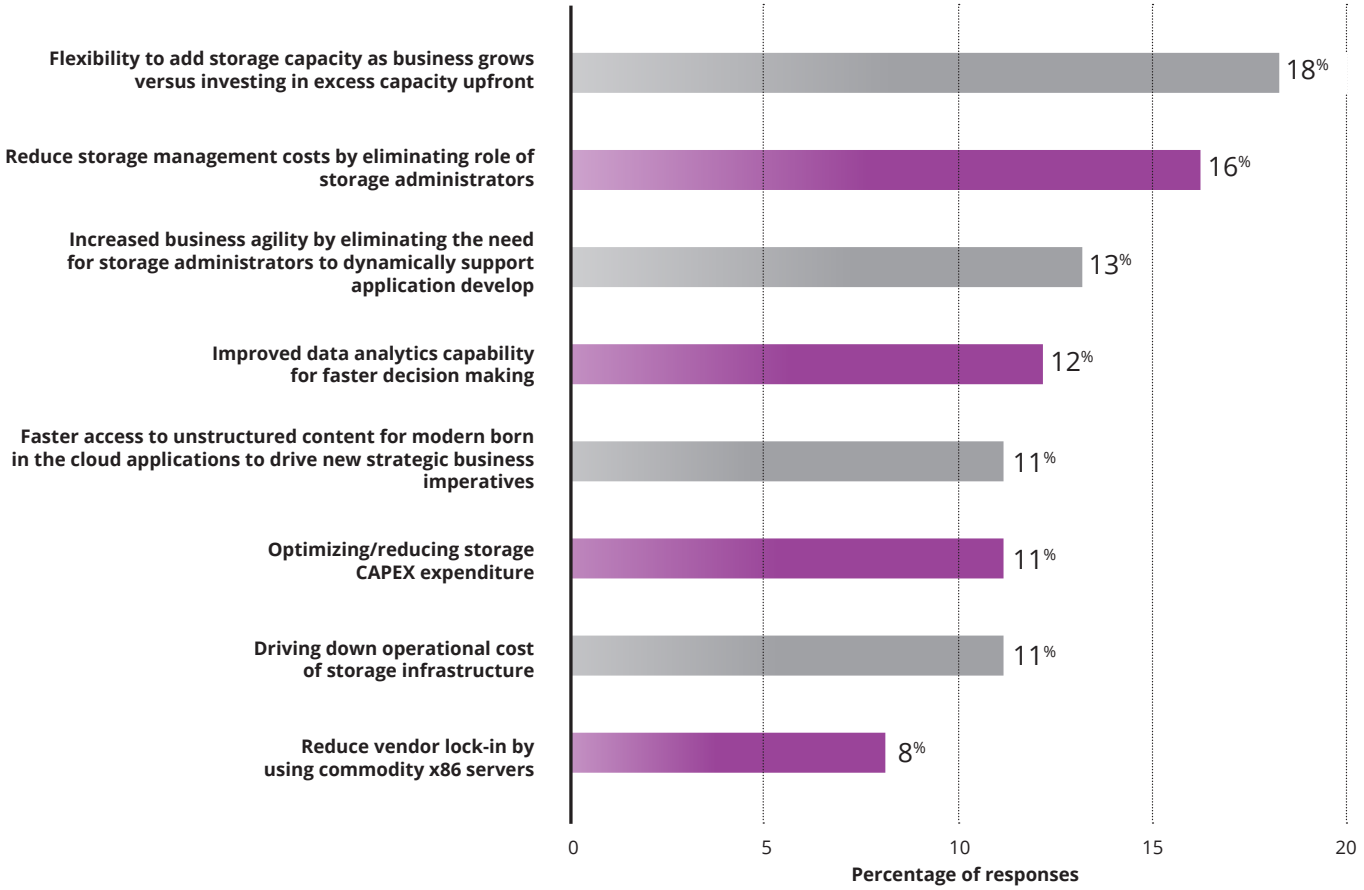
Object storage is able to address the provisioning management issues presented by the expansion of storage at petabyte scales and beyond. Object storage architectures can be scaled out and managed by adding additional nodes. The flat *name space* organization of the data, in combination with its expandable metadata functionality, facilitates this ease of use.

The top three business drivers behind the deployment of or the planned deployment of object storage for enterprises are in Figure 7. Object storage has historically demonstrated its cost-savings capability

FIGURE 7

Top Business Drivers for Deploying Object Storage

Question: What are the top three business drivers for deploying object storage in your organization?



Source: IDC, 2018

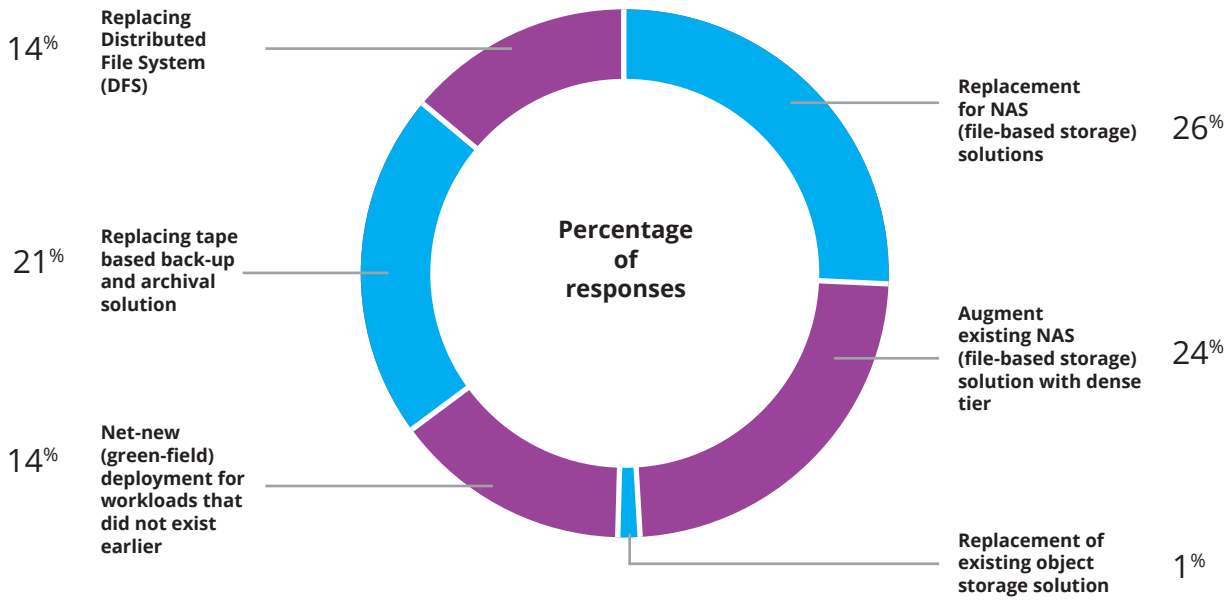
relative to core infrastructure costs by introducing a new storage tier based on cost-optimized hardware. Additionally, it can reduce management overhead with self-healing storage capabilities, in-place cluster upgrades and scale, and the flexibility it provides for any type of data storage requirement.

The decision to deploy object storage may differ between enterprises and between vertical markets. During a recent IDC object storage survey (see Figure 8), respondents stated that their main reason for deploying object storage was to replace or augment traditional Network Attached Storage (NAS) architectures. This use-case is becoming more and more common across industries. It is one thing to deploy a solution that is capable of growing in conjunction with the data stored on it, but it is now becoming more important to use a system that supports customized data descriptions (metadata) and rapid data identification. With object storage delivering on both of these capabilities, it is no wonder that it is removing technology silos to introduce a centralized data hub for all of the organization's most valuable assets.

Object storage is more than simply an alternative for NAS or tape backup solutions. The flexibility, reduced levels of support leading to an increased ability to support the dynamically changing business demands, back object storage a more compelling technology in the current business environment. The ease with which object storage can integrate with public cloud for elasticity, along with native data security through encryption using Restful APIs, has also been cited by the study respondents are the two leading technical reasons for its adoption.

FIGURE 8

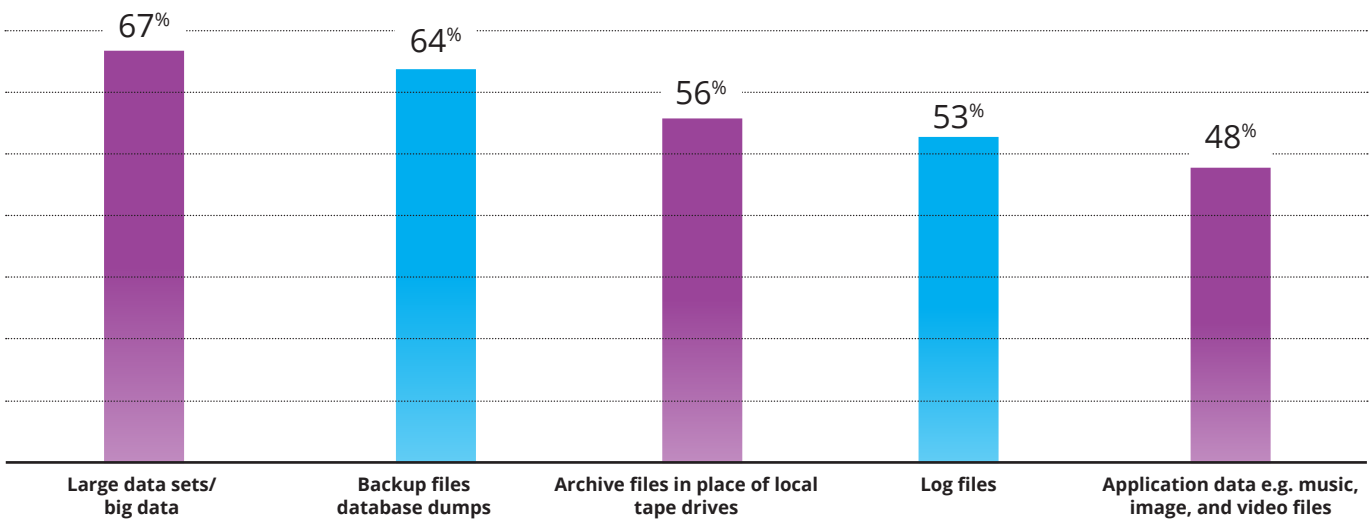
Object-Based Solutions Deployment Functions



Source: IDC, 2018

FIGURE 9

Types of Data Stored or Planned to Be Stored on Object Storage



Source: IDC, 2018

Case Studies of Three Object Storage Use Cases

CASE STUDY 1 Financial Services Organization in Asia/Pacific

THE CHALLENGE

A national bank in Asia/Pacific was working through its digital transformation journey into a digital-native enterprise. The bank's goal is to digitize its client processes to provide fast, accurate, and seamless transactional services to its corporate clients. A key effort in its digitization program is to move all its client's treasury functions and transactions into digital services away from the traditional manual or partial offline process. The organization currently has about 3PB of storage and growing at 30–40% annually. With the increase in digital transactions, the bank will need to build out its storage facility as banking regulations require to keep all information for 7 to 10 years. The bank's management felt that it is losing out on tremendous value as the data that is captured is not being mined and analyzed to extract value. With new services, such as chatbots, being deployed, the level of unstructured data is also expected to escalate.

THE SOLUTION

The bank decided to deploy object storage to cater to the escalating demand for storage as it moves all its processes to digital. The CIO felt that it would have been impossible to support the business transformation using normal block storage. The complexity of managing the high storage demand and cost will be exorbitant under block storage. Object storage was able to provide the bank with a storage solution that it could expand as digital transactions grow, especially when the bank needs to store additional data for seven or more years. This will compound the growth rate of its storage needs. The CIO felt that the metadata of the block storage will become more and more valuable in time as object storage makes it possible to search, retrieve, and analyze the information to extract value.

THE OUTCOME

- 30% savings from block storage in the form of lower storage management costs, resulting in 20% less manpower hired and 50% lower storage capex needed.
- Drastically improved their analytical capabilities of the underlying infrastructure, whilst also providing a more flexible storage system.

Savings from object storage

30%

The CIO felt that the metadata of the block storage will become more and more valuable in time as object storage makes it possible to search, retrieve, and analyze the information to extract value.

CASE STUDY 2 Imaging and Information Service Provider Organization in Asia/Pacific

THE CHALLENGE

A medium-sized surveying and geographic mapping service provider has embarked on an initiative to move all its services to a digital platform to make it more accessible as well as searchable. This is in line with its strategy to provide a differentiated service to its clients in the face of competition from other companies. It currently has about 400TB of data and is expecting this to grow into petabytes of data with its new service. The company also desires to extract more value from its data as its images capture a lot of geographic attributes, tags, and other data. As it moves all its image repository to digital storage, it will require the ability to track and retrieve the images based on the tags and attributes.

THE SOLUTION

The company is using object storage to store its data to reduce the complexity of managing the growing storage of unstructured image data. The head of the information center felt that the metadata of object storage will be invaluable to enable the company to provide a searchable database for its clients, which can be used for big data analysis.

THE OUTCOME

- 50% savings from block storage in the form of lower storage management costs resulting in 50% less manpower.
- Ability to tag and track data objects provide huge efficiency boost, reducing time to market for data requirements.

Savings from object storage

50%

CASE STUDY 3 Audio Visual Development and Processing Company

THE CHALLENGE

A local audiovisual (AV) production and processing company providing digital AV services is a leader in the entertainment sector, providing services to develop and process AV content for its clients in the music, television, and movie industries. It currently has a storage capacity of 20PB and is expecting this to grow as the entertainment sector grows. Because of the urgent nature of its business with short turnaround times, it needs storage to be able to quickly locate and retrieve the relevant AV files for processing and online access by its clients.

THE SOLUTION

The CIO felt that object storage, with its metadata function, helps the company classify the AV object based on many different attributes (e.g., genre, location, artist, and so forth) and provide fast retrieval. The time to regularly back up and retrieve large volumes of work is much shortened because of the way the media is stored as an object with metadata tagged. The object storage solution has a built-in index engine, which means there is no need to maintain another database server, resulting in a quicker backup of the application.

THE OUTCOME

- 70% savings from block storage in the form of lesser capex cost and 50% savings in manpower cost.
- Resources are more efficient due to hugely improved backup times, and new services offering previously archived media becomes a real possibility.

Savings from object storage

70%

Essential Guidance

Data, and how it is used, is at the epicenter of highly digitally transformed organizations and those that can best find value from this data are already emerging as the economic leaders in their market segments. In order to cater for the data demands that the cloud era and big data have brought, an evolving approach to data management strategies becomes a critical weapon in the arsenal of the CIO.

Organizations should adopt a mixed IT strategy, one that includes a Block and NAS solution to address smaller datasets and legacy applications, balanced with an object storage tier to handle data at (and beyond) the petabyte scale in a compliant fashion and support cloud-enabled applications.

In consideration for public and multi-cloud strategies, object storage provides a staging area where data can be analyzed, categorized, and its risk-value determined so that its managed movement to a cloud tier is within risk tolerance levels set by the business.

Object storage helps you manage unstructured data by letting you tag files with custom metadata that describe their contents. This allows you to track and index files without the need for external software or databases. All of your data is self-describing, which opens up an array of new possibilities for data analytics. You can index and search your files without any prior knowledge of a file's internal structure or what applications were used to create it. By giving your files context, you can perform analytics directly on your data.

Without the ability to scale rapidly in a multi-cloud environment, business organizations will be unable to compete in the longer term. Object storage systems have a role to play in this new world order and those that are already adopting them are quickly seeing the benefits of doing so.

To learn more about object storage and Hitachi Vantara's solutions, please visit www.hitachivantara.com

IDC ANALYZE THE FUTURE

Copyright IDC 2018. This IDC White Paper was produced by IDC Custom Solutions. 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. Email: AP_Permissions@idc.com

Sponsored by

HITACHI
Inspire the Next