



I D C M A R K E T S P O T L I G H T

Adoption of Object-Based Storage for Hyperscale Deployments Continues

November 2016

Adapted from *Worldwide File- and Object-Based Storage Forecast, 2016–2020* by Amita Potnis,
IDC #US41685816

Sponsored by IBM

The digital transformation of businesses continues to reshape the IT industry as businesses deal with a social and mobile workforce as well as big data and analytics that causes exponential data growth. Indeed, organizations are finding it difficult to store, manage, and provide value from ever-increasing data sets. As businesses grapple with their ever-increasing data sets, storage solutions such as object-based storage (OBS) lend themselves to solving the problems of scale, management, and costs for today and the future. This paper examines the considerations around choosing an OBS solution.

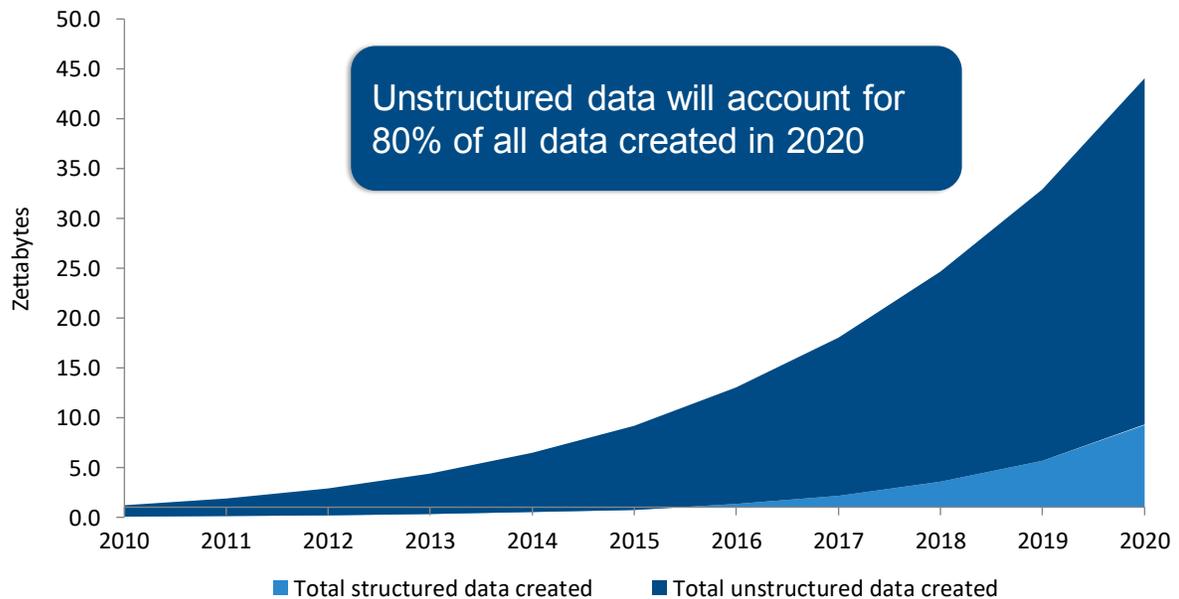
Introduction

Technology means change. Modern technology has penetrated mainstream business and has fundamentally changed its workings today. As businesses embrace digitization, they are now functioning within the frameworks of what IDC calls the 3rd Platform. Businesses are actively employing the four pillars of the 3rd Platform — mobile, cloud, social, and analytics — on a daily basis. IDC believes that this new phase is giving rise to millions of applications, billions of users, and trillions of things, and therefore the 3rd Platform is the new core of IT market growth. The ecosystem of any business successfully operating in the 3rd Platform includes a variety of essential elements, such as a mobile, geographically dispersed workforce; a change in use of storage strategies including cloud; proper management of growing data sets at the petabyte level and beyond; and an increase in hyperscale infrastructure environments. One of the most important drivers for this change is mobility of the workforce that enables efficiency and productivity in any business.

Cloud and big data serve as tools to store these petabyte-scale data sets and analyze them for trends. The use of social media and mobility typically not only helps businesses be more productive and better serve customers but also, in effect, creates the problem of increased data volume and management. This data can be structured or unstructured in nature. IDC expects that by 2020, the total amount of data created will reach 44 zettabytes, and 80% of it will be unstructured, as shown in Figure 1.

FIGURE 1

Capacity Growth by Data Type



Source: IDC, 2016

As unstructured data continues to rise and reach approximately 35 zettabytes over the next four years, businesses have to manage and strategize through three key trends:

- A broad range of mobile devices and new use cases, such as the Internet of Things (IoT), are increasing not just the scale at which data is generated but also the number of ways in which data is generated.
- Businesses participating in the software economy create and facilitate massive repositories of human- or machine-generated data for consumption. For example, companies such as LinkedIn, Uber, Airbnb, Yahoo!, Netflix, and Apple, as well as various government organizations, are supporting massive repositories of content that is consumed by users.
- Given the challenges associated with never-ending data growth, businesses are less likely to continue their investments in legacy storage solutions and instead will seek out highly optimized, scalable, low-cost, long-term storage alternatives.

In general, businesses of all sizes across many verticals have to deal with the aforementioned trends. Regardless of the exact nature of business, organizations are finding it difficult to store, manage, and provide value from ever-increasing data sets. A caveat to any potential solution for this dilemma is to also keep costs in check, which is a major point of concern for most, if not all, businesses.

In addition to the challenge of data growth, an increasing number of regulations and laws require longer retention of a broader range of digital data. For example, regulations now stipulate the retention of not only email and instant messages, as forms of electronic communications, but also public and enterprise social communications that occur over social networks. The increase in management, complexity, and IT spend required because of such new regulations impacts IT organizations and is spurring a fresh demand for technology that is agile, scalable, and cost effective.

In IDC's fall 2015 *Storage User Demand Study*, a majority of respondents (over 1,000 storage professionals) indicated that capacity scaling was the single-largest factor influencing a storage purchasing decision.

As businesses grapple with their ever-increasing data sets, storage solutions such as object-based storage lend themselves to solving the problems of scale, management, and costs for today and the future. OBS solutions are available in several delivery models, offering end users the flexibility to choose the optimal solution for their needs.

Object-Based Storage: A Definite Alternative

OBS employs a flat namespace spanning multiple physical hardware designed for multipetabyte environments. OBS solutions use flat structures to organize data. These are higher-level structures in which data is organized using an "account, container (aka buckets), and object" approach, wherein "objects" are analogous to "files" in file-based storage solutions. Accounts, containers, and objects are referenced by a metadata repository that manages attributes of data stored in that structure. OBS solutions generally operate on a per-object level (i.e., allow each object to be treated independently as far as policy management is concerned), whereas others operate at a container or an account level (i.e., policies are applied at the container or account levels). Several OBS solutions also leverage NoSQL databases as metadata repositories and persistent data stores.

An Edge with Object-Based Storage

The basic premise of OBS architecture is to deliver high scalability, reduced complexity in terms of management, and lower total cost of ownership (TCO). OBS solutions often beat traditional scale-out network-attached storage (NAS) systems in terms of performance while offering key benefits such as file ingestion and concurrent retrieval across multiple sites. In addition, many OBS solutions offer the flexibility to retrieve or push data to and from public cloud storage services (which are also based on OBS). Many of these benefits enable IT organizations to focus on finding value in the data as opposed to managing data repositories.

In general, the advantages of an OBS solution include:

- **Nondisruptive scale.** OBS solutions are generally scale out and therefore can grow as needed by simply adding capacity or performance. These solutions have a flat, global namespace and metadata capabilities designed for multipetabyte environments. The global namespace enables nondisruptive addition, removal, and replacement of infrastructure as needs warrant.
- **Cost.** Among the benefits of any OBS solution are the economic advantages it provides. Compared with traditional solutions, an OBS solution typically reduces capex and opex requirements and, in turn, TCO. Capex reduction comes as a result of lower cost per gigabyte, while opex savings are realized with dynamic scaling, upgrades, and migration. Traditional storage burdens end users with up-front initial investments and ongoing maintenance, management, and forklift upgrade expenses.
- **Accessibility.** OBS solutions lend themselves to long-term data retention with faster and simpler retrieval than legacy tape-based solutions. OBS solutions are routinely deployed as an extension to primary storage in the form of low-cost nearline archive storage. Thus, end users can protect investments in existing storage assets while making content more accessible. Many end users are using OBS to support numerous use cases, such as rich media (content delivery), big data, healthcare and life sciences, video surveillance, and backup/archive, as well as other new use cases.

A Choice of Delivery Models

Every business is unique, and so are its storage demands. OBS offers a variety of delivery models to choose from given the unique storage requirements of each business. Each delivery model has its own place in the market and brings its own advantages. OBS delivery models include:

- **Appliance (purpose-built integrated solution of hardware and software).** The decision to procure an appliance is driven sometimes by the need to have a limited vendor storage strategy that builds on existing relationships. When an end user chooses an appliance, it is often intended to also serve as a one-stop arrangement for maintenance and support contracts. OBS appliances account for the smallest share of revenue in the overall OBS market, but there is demand for this delivery model in the marketplace. OBS appliances can be deployed to house data on-premises, in a remote datacenter, or in a private cloud environment. Most, if not all, OBS appliances support movement of data to and from public cloud storage services.
- **Software only (hardware purchased and integrated separately).** One of the driving factors for deploying software-only OBS solutions is to lower costs while taking advantage of scale, agility, and accessibility. The key advantage in terms of cost for software-only OBS solutions is that the software is able to work with the underlying hardware to drive performance or other features. This reduces any added costs from the vendor side in terms of research and development, quality assurance, and testing that are passed on to end users. Another advantage of software-only OBS solutions is that end users can start small, making targeted investments in the choice of hardware at affordable price points, and grow their storage environment as needed. For this reason, many service providers that are looking to design and offer cloud-based storage services often choose this OBS delivery model. OBS software-only offerings also include open source software, which can be downloaded and used for no cost, as well as commercial or commercialized open source software that requires payment of a license fee. Like appliances, software-only OBS solutions can be deployed to house data on-premises, in a remote datacenter, or in a cloud environment. Many software-only OBS solutions also enable the movement of data to and from public cloud storage services of choice.
- **Public cloud storage services.** The key differentiation between this offering and the previously mentioned offerings is how the solutions are brought to market. Appliances and software-only OBS solutions are sold as products, but public cloud storage based on OBS technology can be used only as a service. Public cloud storage providers such as Amazon or Microsoft Azure offer OBS on proprietary implementations that they designed in-house that can be used only as a service. Other companies such as IBM offer end users not just public cloud storage services but also on-premises and hybrid cloud solutions.

OBS public cloud storage services solutions represent much of the OBS revenue; however, businesses are also deploying other delivery models of OBS such as appliances or software only. Businesses that own large sets of sensitive data may be wary of public cloud-based storage services not only for the sake of security but also for the hidden retrieval and long-term costs. Interest in software-only OBS platforms that can be deployed on commodity hardware is on the rise. At present, the OBS market has more software-only vendors than those offering purpose-built integrated solutions. According to IDC's *Worldwide File- and Object-Based Storage Forecast 2016–2020*, the overall OBS market is expected to grow at a CAGR of 8.7% from 2015 to 2020 and is expected to reach \$19.8 billion in revenue by 2020. While much of this revenue will be generated by public cloud storage service providers (such as Amazon, Microsoft Azure, and IBM), many enterprises and service providers are actively considering OBS solutions for their on-premises storage needs.

Leading Use Cases of OBS

Early adopters of OBS often used the technology for long-term data archiving and retention as well as backup. Today, OBS extends well beyond these limited use cases and is used as a NAS replacement or in net-new greenfield deployments for 3rd Platform workloads. As the market continues to gravitate toward digitization, a trend that is commonplace in any enterprise is that the line of business generally dictates IT initiatives as opposed to the IT department making storage decisions. The fundamental expectation from any IT organization is that the underlying infrastructure should work seamlessly and cater to the growing and stringent 3rd Platform storage demands. These demands include collecting, storing, managing, and analyzing massive data sets to make strategic business decisions without worrying about IT maintenance and operations.

Businesses of all sizes across many verticals (e.g., media and entertainment, healthcare, financial services, government and education, life sciences, manufacturing) are experiencing the same or similar phenomenon. For example, the healthcare industry increasingly relies on digitized MRIs and X-rays; media and entertainment is grappling with the issues of 4K video content; government agencies continue to utilize video surveillance or geospatial imagery to support national security; the life sciences market continues to focus more on genome sequencing and research. A relatively new line of business is that of collecting and assimilating digital information for and delivering digital information to corporations (business to business) or end users (business to consumer). IDC refers to this new economic segment as value-added content providers (VACPs) that create or add value to large information sources that end users use to gain additional insights in specific areas either to improve or to expand business. As businesses become more data oriented, OBS solutions are adopted by not just service providers or VACPs but also, increasingly, enterprise customers.

Conclusion

The penetration and the adoption rate of OBS have come a long way from the days when vendors had to explain what OBS was; today, enterprises are aware of this technology and actively considering it for their storage needs. OBS suppliers continue to innovate and build solutions that will lay the foundation of a new digital era for years to come. New OBS solutions also take advantage of solid state drives for performance instead of being limited to the use of high-capacity hard drives. Other OBS suppliers, such as IBM, are bringing to market a variety of OBS delivery models that can cater to a vast audience. As the number of OBS suppliers and the various flavors of delivery models and features widen, making the right choice can be a daunting task for any end user.

The key points for end users to consider before deciding on any OBS solutions are:

- Specifically, which workloads will potentially be moved to a new OBS solution? What is the projected data growth of workloads moving to an OBS solution compared with that of other workloads in the organization? Over time, are there other workloads that could be moved to an OBS solution?
- Which features and technology partnerships are a "must-have" in a given OBS solution for serious consideration? Is the OBS solution interoperable with existing storage infrastructure or a flexible deployment option to address the unpredictable nature of data or dynamic workload requirements? Is the OBS solution secure with proven availability? Which delivery model would best fit these needs?
- What is the track record of a potential OBS supplier in terms of scalability, data integrity, performance, and costs? Do these fit the requirements of the organization's workloads?
- Most importantly, considering the organization's needs and requirements, which OBS delivery model, among the many, is the optimal choice? Knowing the specific requirements for retrieval and retention policies will help in making this choice.

Data growth is real. The day-to-day management and maintenance problems stemming from this data growth are also real. An in-depth understanding of the current data sets, requirements, and necessary features, as well as future projections related to data growth and business strategy, will put any organization in a better position to make an astute decision on procuring the right OBS.

Finally, as end users employ the four pillars (mobile, cloud, social, and analytics) of the 3rd Platform in their day-to-day business, they are creating more content than ever before. Businesses therefore need a flexible, scalable, and simple storage approach that is also a cognitive model that creates value from the data generated and stored. An OBS solution that lends itself to these needs will help open up a world of possibilities and unleash creativity and innovation by way of defining new processes, applications, and experiences that elevate business.

A B O U T T H I S P U B L I C A T I O N

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

C O P Y R I G H T A N D R E S T R I C T I O N S

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

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

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