White Paper

Enterprises Rely on Public Cloud Object Storage to Manage Data Growth, Ensure Resilience, and Generate Value

Sponsored by: IBM

Andrew Smith
February 2021

EXECUTIVE SUMMARY

Public cloud object storage (COS) services have emerged as a leading solution for storing massive volumes of data (tens to hundreds of petabytes) and are widely known for their deployment flexibility (on/off premises or public cloud) as well as cost-effectiveness (with storage services typically measured in dollar per gigabyte per month). Many public cloud object storage providers offer services that enable enterprises to address their data growth, data resilience, and cloud migration strategies through a platform of tools and services, which go well beyond storage capacity. Leading cloud object storage providers do this by engineering an ecosystem of services, which deliver a combination of the following core capabilities:

- Storage tiering and cost optimization tools based on data life cycle or policy
- Data security and resilience via native capabilities for high availability, provisioning, replication, and management
- Policy-based user access and sharing
- Data and application migration/modernization capabilities
- IBM's data analytics services in the form of dashboards, reporting, or visualization tools

This IDC White Paper provides empirical data quantifying the public cloud object storage market and shows how these service offerings have grown over time. From a subjective standpoint, we analyze the current and potential value of cloud object storage solutions as an essential means to help enterprises manage the challenges associated with exponential data growth and data resilience. To do so, this White Paper specifically explores three use cases, and how they can be addressed using public cloud object storage services, including:

- Cloud data migration and infrastructure modernization
- Cloud data security, protection, and resilience
- Making unstructured data a usable asset for the enterprise

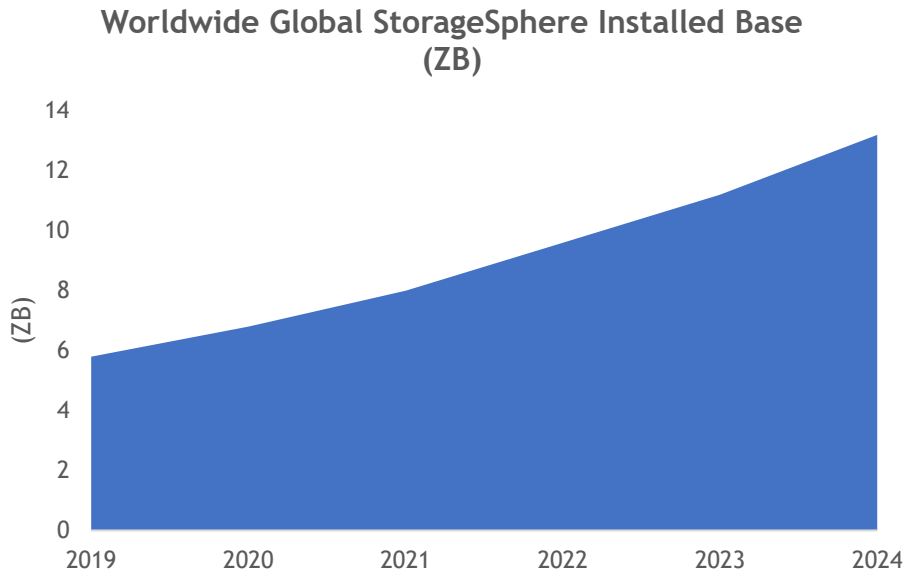
Finally, we provide a comprehensive overview of IBM's Cloud Object Storage (IBM COS) portfolio and how IBM solutions can accommodate each specific use case outlined previously.

SITUATION OVERVIEW

On average, enterprises indicate that they expect 30% data growth annually (primarily unstructured). Using this growth rate as a general guide, we can assume that an organization managing 10PB of data today will store upward of 13PB of data the following year. And this pace of growth is not expected to slow down. IDC's Global StorageSphere data indicates that, over the long term, the installed base of enterprise storage capacity is expected to grow at a 2019–2024 CAGR of 17.8%, resulting in an installed base of storage capacity of 13.2ZB in 2024, compared with 5.8ZB in 2019, more than doubling over the forecast period (see Figure 1).

FIGURE 1

Worldwide Installed Base, 2019–2024



Note: There are 1 million petabytes in a zettabyte.

Source: IDC, 2021

In many cases, this pace of storage growth may exceed existing infrastructure and quickly pressure the limits of both cloud and on-premises capacity plans. Unanticipated data growth variables (e.g., a global pandemic like COVID-19 driving a precipitous rise in remote work) put additional pressure on organizations to adapt to new operational realities. In many cases, organizations are slow to adapt to these changes due to infrastructure-related hindrances including:

- Storage silos caused by disparate solutions across infrastructure locations (e.g., datacenter and cloud) or business units
- Management overhead due to duplicative/redundant work across storage silos, resulting in inefficiency and more hours spent on storage infrastructure and management
- Inability to mine data for added value due to lack of tools that can report and visualize data across an enterprise's entire infrastructure footprint

- Reduced IT infrastructure/storage budget causes incremental refreshes, as opposed to comprehensive modernization efforts
- Inability to scale resources quickly and effectively or take advantage of new technologies due to limitations of current architecture

Many modern organizations are adopting cloud storage to address these challenges, and this growing adoption is reflected by the consistent expansion of the public cloud object storage market.

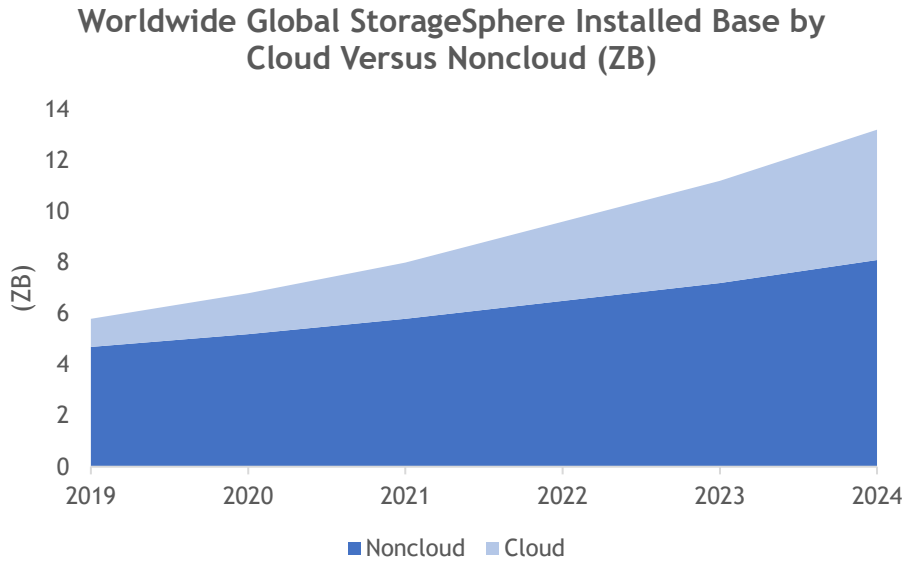
Public Cloud Object Storage Helps Enterprises Manage Exponential Data Growth

While the perceived value of enterprise data may be limitless, enterprise IT budgets are not. Many IT administrators find themselves stuck in this sort of data sustainability paradox, where their storage budget is stagnant, or shrinking as a proportion of total IT budget, but their organization's storage capacity is growing at a much higher rate. At the same time, business units and application owners are demanding faster access to data and the ability to integrate analytical tools and services to accommodate the organization's modern digital initiatives. How is it possible for an organization to do this? Enterprises must strike a balance between the promise of generating value from data and the reality of running their business applications cost effectively and securely.

As a result of this dynamic, we expect a growing proportion of the installed base of storage capacity to reside in the public cloud, where organizations can take advantage of cost profiles and levels of service automation and management that help them keep up with IT infrastructure demands. Putting this trend into perspective, IDC's Global StorageSphere data indicates that the installed base of storage in the cloud is projected to grow from approximately 19% of the total StorageSphere in 2019 to 39% in 2024 (see Figure 2).

FIGURE 2

Worldwide Installed Base Share by Cloud Versus Noncloud, 2019–2024



Note: There are 1 million petabytes in a zettabyte.

Source: IDC, 2021

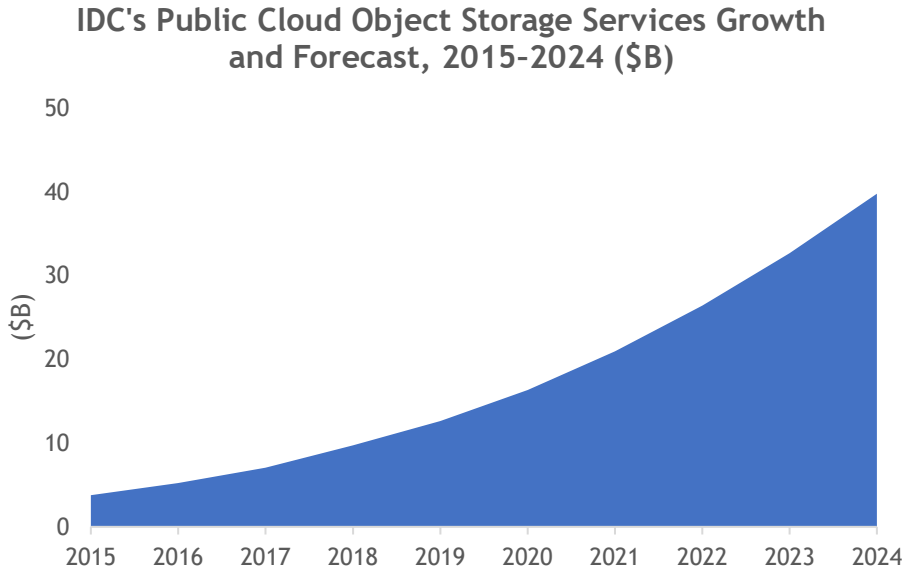
Object-based public cloud storage services are an ideal landing spot for much of this data. These cloud services can provide a data lake foundation, which combines massive scale with cost efficiency. Furthermore, these services allow for frequent (and increasingly performant) access options for some primary applications, and cost-optimized, long-term retention storage (archival and cold storage) of large repositories of unstructured data helps organizations ensure compliance and future proof their storage infrastructure.

The Continued Growth of the Public Cloud Object Storage Market

IDC tracker data indicates the public cloud object storage market has grown significantly over the past four years. We estimate that in 2020, the public cloud object storage market grew 29% to $16.3 billion. Looking forward, we expect the public cloud object storage market to achieve double-digit annual growth throughout the forecast period, reaching approximately $40 billion by 2024 (see Figure 3).

FIGURE 3

Cloud-Based Object Storage Market, 2015–2024



Source: IDC's Worldwide Semiannual Public Cloud Services Tracker, 1H20

Global Pandemic Impact on Public Cloud Storage

Readers may notice something odd about the forecast growth trajectory of the public cloud object storage market: it is not impacted negatively by COVID-19. This is because we see many organizations using the uncertainty and budgetary pressures generated by COVID-19 to promote cloud storage adoption. IDC's *COVID-19 Impact on IT Spending Survey* shows that expectations around public cloud IaaS (compute + storage) purchasing have steadily improved since May, and as of October, 46% of respondents indicated that they expect their public cloud IaaS spending to increase, while 23% of respondents expect spending to remain the same.

The unpredictable nature of the pandemic's impact on enterprises underscores the need to be able to adapt quickly and remain agile from both operational and technical perspectives. Budgetary constraints, coupled with new operational challenges (e.g., accelerated reliance on ecommerce and increasingly remote workforces), will only exacerbate challenges currently faced by customers relying on on-premises infrastructure (e.g., data silos, management, capex, and changing demands from net-new greenfield workloads). Cloud storage services will be a key contributor to both efficiency and cost savings enterprises need in the immediate future.

CONSIDERING CLOUD OBJECT STORAGE FROM IBM

IBM has built a robust portfolio of cloud object storage services over the past several years, all of which are grouped under the vendor's Cloud Object Storage umbrella. IBM offers several tiers of object storage to accommodate both primary and secondary storage use cases, with features including:

- The use of information dispersal algorithms by IBM COS across regional zones to achieve data resilience and high availability of object storage
- Access latency in the tens to hundreds of milliseconds depending on object storage tier, distribution, and data access requirements (e.g., "hot" versus "cold")
- Per-object/bucket retention policies
- Granular selection of "cold storage" tiers options (with various data access parameters and price points in terms of dollar per gigabyte per month)
- Direct integration with IBM's artificial intelligence (AI)/analytics services
- Physical bulk data transfer as well as high-speed network transfer

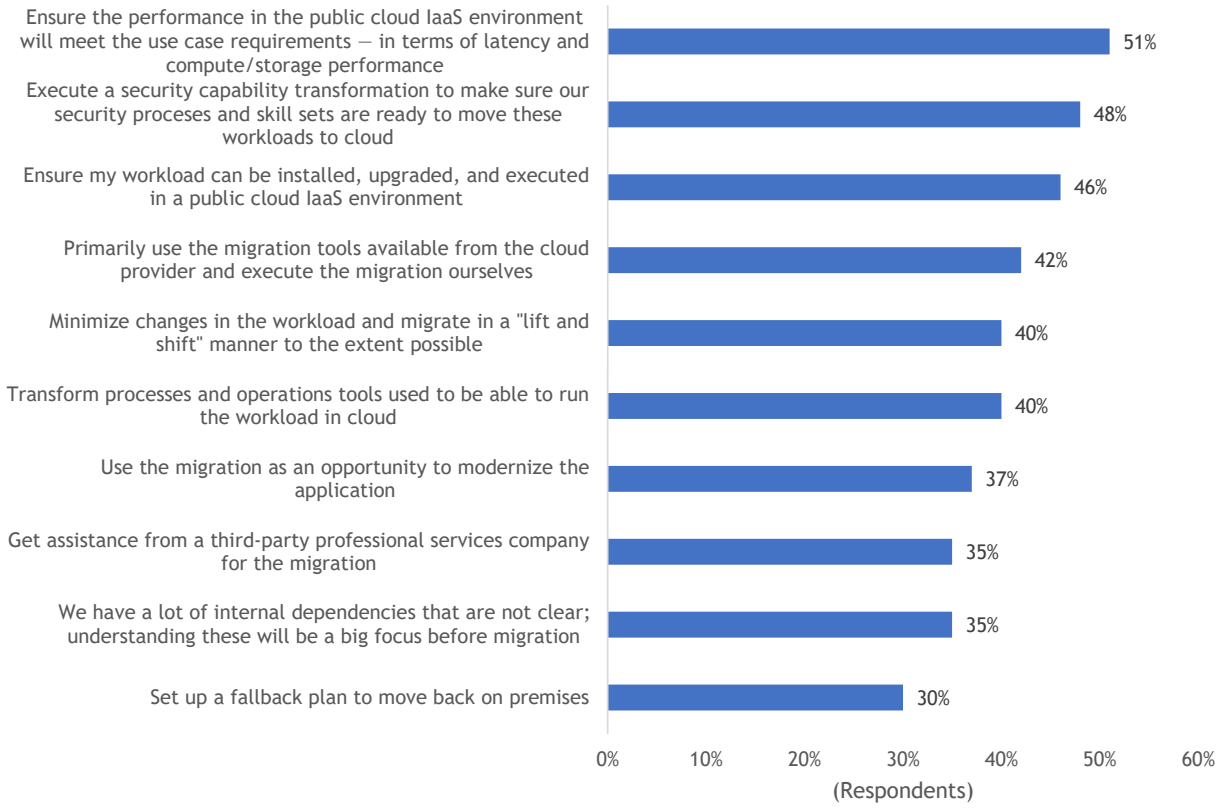
In addition to the aforementioned features, IBM COS benefits from a robust ecosystem of partner solutions that includes a range of ISVs and services providers designed to accommodate customer use cases for specialized industry cloud archiving, data migration to cloud, cloud file storage services, cloud data protection, backup, disaster recovery, and cloud-based analytics. The sections that follow explore three specific cloud object storage use cases and the challenges and opportunities associated with each use case. Finally, we illustrate how IBM COS can be leveraged to accommodate enterprise requirements across each scenario.

Use Case Example 1: Cloud Data Migration and Infrastructure Modernization

Enterprises face a host of decisions, opportunities, and challenges when it comes to infrastructure and data migration leveraging cloud infrastructure. IDC's 2020 *IaaS View Survey* indicates that workload performance, security, and execution are the top 3 priorities customers focus on to enable a smooth migration of workloads to public cloud (see Figure 4).

FIGURE 4

Top Priorities Enabling a Smooth Migration of Workloads into Public Cloud



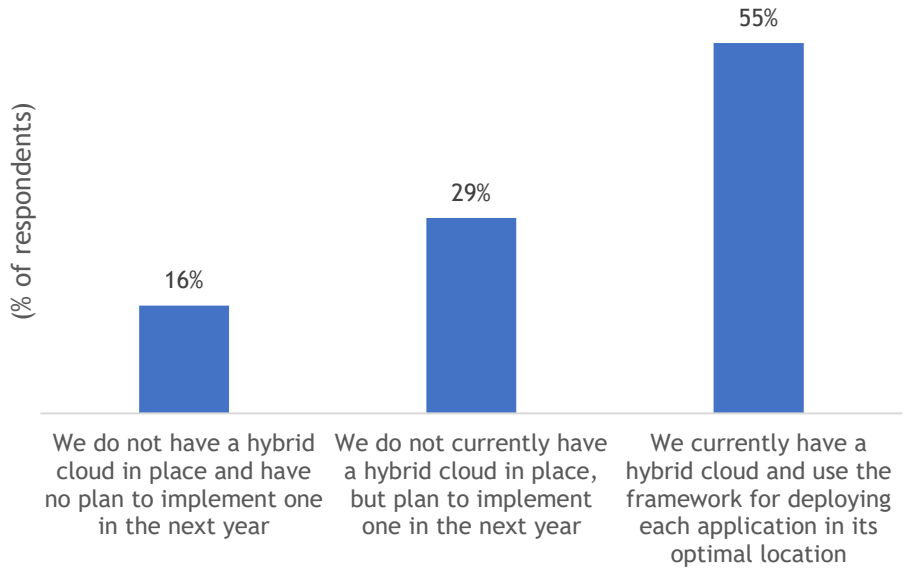
n = 1,500

Source: IDC's *IaaSView Survey,* 2020

Enterprises considering a workload migration are always faced with the concern that legacy applications simply may not work in a cloud environment. As enterprises move along their cloud journey from planned to active use of cloud infrastructure, regulatory restrictions and internal policy tend to become the next challenges to overcome. For these reasons, it is logical that so many organizations today are taking a hybrid approach to cloud (i.e., a mix of both on premises and public cloud) (see Figure 5). This allows organizations to optimize application and data placement within the right ecosystem based on the performance, security, or cost requirements. Hybrid cloud deployments also allow enterprises to account for the fact that many of these requirements will change over time, and workloads may need to be moved or distributed across environments as a result. Finally, from a cost perspective, hybrid cloud deployments give enterprises the flexibility to deploy infrastructure using both capex and opex models, as well as leverage or repurpose existing systems they may already own or operate.

FIGURE 5

Usage of Hybrid Cloud by Organizations



n = 1,500

Source: IDC's *IaaSView Survey,* 2020

How IBM COS Enables Hybrid Cloud Data Migration and Modernization Efforts

IBM COS offers a wide range of services, as well as integrations across the wider IBM portfolio, to help customers with hybrid cloud data migrations and modernization. Essential IBM products and services that apply to this use case and help customers manage the performance, execution, and data security of their hybrid cloud environment include:

- **IBM COS storage classes and tiers:** IBM COS offers a wide range of storage tiers that can be leveraged as a content store and a foundation for enterprise workload migration/modernization. These tiers range from ultralow-cost archive to more performant access tiers. IBM COS Vault, Smart Tier (Cool and Cold), Cold Vault, and Archive provide customers with several pricing (dollar per gigabyte per month) and data access options, allowing users to maximize efficiencies associated with their cold storage workloads and use cases. IBM COS Standard and Smart Tier can be used to give customers the flexibility and performance they need in terms of data accessibility. As mentioned previously, many cloud migrations begin with a storage-related use case focused on cost optimization. IBM's wide range of storage classes allows enterprises to engineer a highly tailored storage environment for their data migration needs.
- **Built-in high-speed transfer:** One of the major hurdles for cloud data migration can be the initial movement of petabytes of data and the logistical challenges associated with moving this data in terms of latency, bandwidth, and cost. Integration between IBM Aspera and IBM COS is designed to address this challenge, by enabling high-speed data transfer over established network infrastructure (e.g., 10Gb Ethernet). This provides enterprises with an integrated service to get their data from on premises to hybrid cloud and public cloud, or vice versa.

- **Hybrid cloud support:** IBM COS is positioned as the primary IBM solution enterprises can leverage to manage their distributed and hybrid cloud requirements. IBM offers IBM COS for on premises and as-a-cloud service with consistent functionality and APIs. IBM plans to integrate object storage with Cloud Satellite to offer deployment for edge, on-premises, and hybrid environments. Cloud Satellite allows users to eliminate infrastructure complexity and silos within their hybrid cloud ecosystems. This consistent view allows administrators to more effectively manage or provision resources, ensure deployment consistency (e.g., APIs), and implement and track compliance and security protocols.

Use Case Example 2: Cloud Data Security, Protection, and Resilience

Some organizations remain challenged to convey cloud IaaS as a secure and resilient solution when compared with owned on-premises infrastructure. Lack of visibility into the physical hardware and network configuration can be a jarring experience for traditional IT organizations. However, in many cases, cloud storage services are more resilient and redundant than an on-premises counterpart, and enterprises adopting public cloud IaaS indicate that better data protection, data security, and performance (compared with on premises) were each among the top considerations that drove their decision to use cloud IaaS (see Figure 6). Many enterprises now see data protection, security, and performance of cloud IaaS as a key differentiator, instead of a liability.

FIGURE 6

Factors That Drove/Are Driving Use of Cloud IaaS Instead of Alternatives

n = 1,500

Source: IDC's *IaaSView Survey,* 2020

How IBM COS Helps Establish a Foundation for Cloud Data Protection and Cyber-Resilience

IBM COS was built with enterprise data protection and resilience in mind. All storage tiers offer 99%+ data availability and 14-9s of durability. Additional data resiliency services and features include:

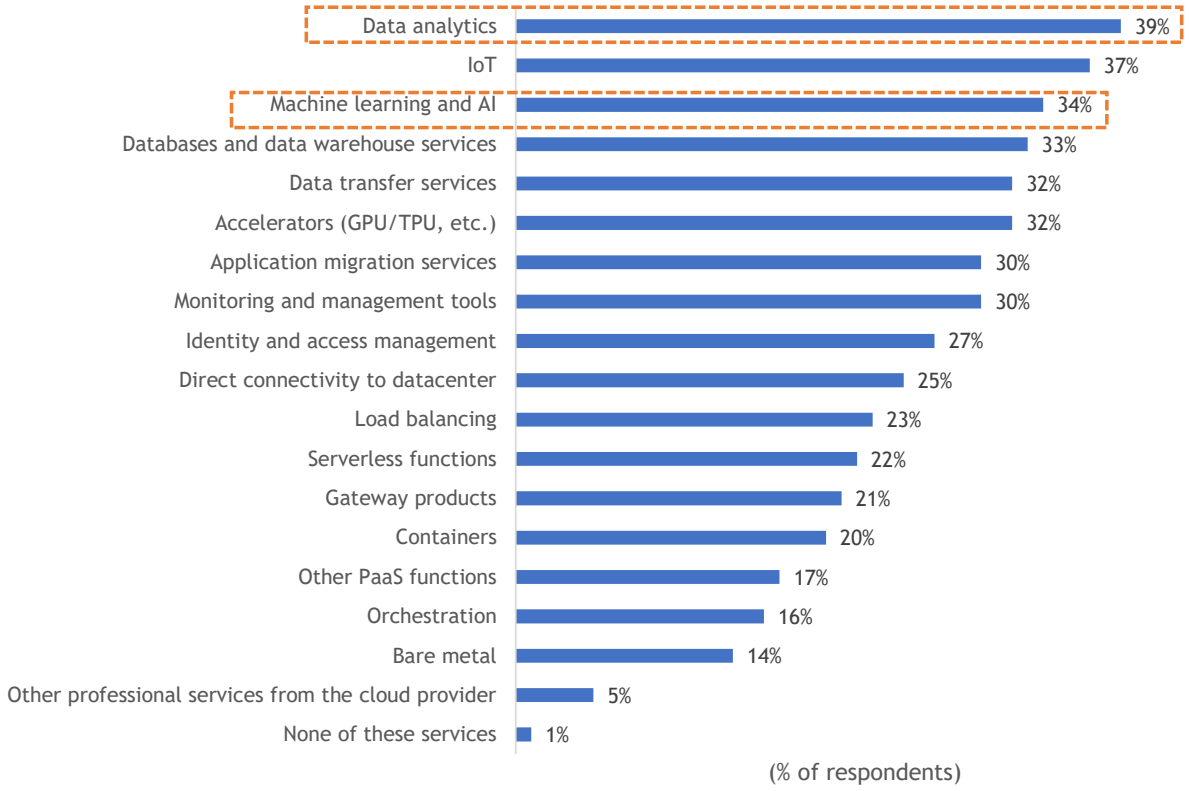
- **Native data access and resilience features:** Object and bucket immutability, replication, automated metadata creation and tagging, data auditing capabilities, access permissions, and API consistency across storage tiers are just some of the built-in features native to IBM COS, which help establish a foundation for data resilience.
- **Distributed failure zones:** IBM COS allows enterprises to establish multiregion data redundancy, a key component of modern data protection and disaster recovery, which can be quickly and efficiently established in the cloud without the need to develop new physical locations.
- **Data security and compliance features:** All objects stored in IBM COS are encrypted by default. Organizations can manage their own encryption keys or use IBM Key Protect (bring your own key) and/or IBM Hyper Protect Crypto Services (keep your own key). IBM Cloud Identity and Access Management allows users to assign policies to control who has access to the data in COS buckets as well as other IBM Cloud resources. IBM COS also boasts a list of compliance certifications from industry specific (e.g., HIPAA) to workload specific (e.g., WORM).
- **Data protection partnerships:** IBM COS offers a wide range of ISV partnerships to enable backup and disaster recovery leveraging IBM COS as a backup target or failover site. These partnerships enable the data staging and movement necessary to make sure that applications and workloads running on public cloud resources continue to operate and fail over as needed in the event of data loss or planned/unplanned disaster.

Use Case Example 3: Making Unstructured Data a Usable Asset for the Enterprise

Enterprises are forced to optimize decision making in a time-sensitive manner to remain competitive and meet customer expectations. Not only do their customers and partners demand this type of experience, but even internally, the expansion of business analytics and data science teams requires efficient and unfettered access to unstructured data. This can only happen if the existing data and the infrastructure that manages it are accessible, and enterprise data can be cost effectively retained and mined for insights in the form of correlations and patterns. To do this, enterprises are adopting tools that provide context — information about the data being analyzed and attributes (metadata) of the unstructured data. Increasingly, cloud storage platforms provide integrated data analytics and artificial intelligence services and software integrations to enable this work. In many cases, these services are embedded within the native environment and can simply be turned "on" or "off" when the customer needs them. This ease of use and integration has consistently driven adoption of adjacent services for cloud data analytics (see Figure 7).

FIGURE 7

Public Cloud Services Organizations Are Using or Planning to Use Within the Next Year



n = 1,500

Source: IDC's *IaaSView Survey,* 2020

How IBM COS Makes Unstructured Data Usable — Marrying Content with AI and Analytics

IBM has a history of productizing analytics solutions/services for enterprise use, and unstructured data stored in IBM COS benefits from integration with a range of solutions across the IBM Watson and IBM Cloud portfolios, including:

- **IBM Cloud Pak for Data as a Service:** IBM COS can be used as the storage foundation enabling a range of IBM's structured Cloud Pak for Data as a Service, which includes built-in capabilities for data ingest, collection, and extraction. IBM COS delivers a persistent storage layer that can be leveraged by the containerized IBM software solutions organized in a Cloud Pak for Data as a Service. For example, IBM COS can act as the cloud-based persistent storage layer for IBM Cloud Pak for Data by managing the integrations of multiple heterogeneous data sources, organizing that data, and presenting it in a hierarchical format to enable integration and querying by databases and data scientists.

- **IBM Watson services portfolio:** IBM COS can act as both the object storage repository and the write location for many Watson-based data analytics services. Watson Studio, Watson Speech to Text, and Watson Natural Language Understanding are all services that enterprises can integrate with IBM COS and use to develop AI and analytics applications for emerging use cases like call center automation or machine learning–based healthcare apps, which must access hundreds of thousands of images held in object storage in order to learn and develop pattern recognition.

Challenges to Consider

- IBM's COS is relatively new to the public cloud storage market (initially acquired by way of Cleversafe in 2015 and subsequently rebranded as IBM Cloud Object Storage). However, IBM has quickly built out its storage capabilities over the past five years and continues to invest in the segment. The vendor has maintained a steady pace of storage product/feature updates and has a detailed road map of planned features including access to new file types, new security and compliance capabilities, and new data management features. This road map is indicative of IBM's long-term focus and commitment to the company's cloud storage portfolio.
- With market growth also comes growing pains. In the case of public cloud object storage, this typically manifests in the form of skills shortages (in terms of either the platform or application to specific industries or use cases). This is a challenge that faces all cloud storage providers, not just IBM. But continued development of industry-specific applications and support and migration enablement tools and services are all critical to help overcome this challenge.
- IBM COS targets object storage use cases. Increasingly, cloud storage is branching out into high-performance and compute-led use cases, which require file- and block-based access. IBM doesn't currently have native services for these access types (e.g., NFS/SMB) but can be integrated via partners.
- As discussed previously, native cloud services for backup and disaster recovery are not available as part of IBM COS but can be added via partners. This may be viewed as a challenge by some enterprises looking for a single provider for their cloud-based secondary storage workloads.
- IBM's breadth of offerings are both a strength and a challenge. IBM's solutions in traditional (noncloud) storage systems segments (from converged systems to tape libraries) are unique and demonstrate expertise across a wide spectrum of storage deployments but may also create confusion or leave customers with a perception of legacy technology if they are not familiar with IBM's cloud portfolio.

Conclusion

As organizations grapple with the perennial challenges of infrastructure modernization, data growth, and resilience, many will look to public cloud storage services that can act as a performant, cost-effective alternative to traditional storage systems. In general, public cloud object storage services deliver the right balance of scale, complexity, and costs (high scale, low complexity, and low cost) to address a wide range of enterprise workloads. Public cloud object storage providers have capitalized on this broad applicability, which is reflected by the consistent, double-digit annual growth of the object storage market over the past five years. Today, all major public cloud object storage services deliver enterprise-grade levels of data access, resilience, and security. Organizations should feel confident in their selection of a public cloud object storage provider to address a range of modern IT storage challenges, several of which have been covered in this analysis.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com