IBM Cloud Object Storage System features and benefits

Object storage platform to help solve petabyte-and-beyond storage challenges

Highlights

- Scalability that offers a single storage system and namespace versus an ever-increasing number of limited-capacity storage silos
- Security features include a wide range of capabilities designed to help meet security requirements
- Reliability and availability characteristics of the system are configurable to best suit the customer's requirements
- Manageability that helps enable storage administrators to handle large storage capacity
- Flexibility of a defined storage solution that does not require specific or proprietary hardware

Introduction

The IBM® Cloud Object Storage (COS) System is a breakthrough cloud platform that helps solve petabyte and beyond storage challenges for companies worldwide. This Solution Brief summarizes its features that are designed to provide the scalability, availability, security, manageability, flexibility, and lower total cost of ownership (TCO) benefits.

The IBM COS System is deployed as a cluster combining three types of nodes as shown in Figure 1. Each node consists of IBM COS software running on an industry-standard server – IBM COS software is compatible with a wide range of servers from many sources, including a physical or virtual appliance, and IBM conducts certification of specific servers customers wish to use in their environment to help insure a quick initial installation, long-term reliability, and predictable performance.

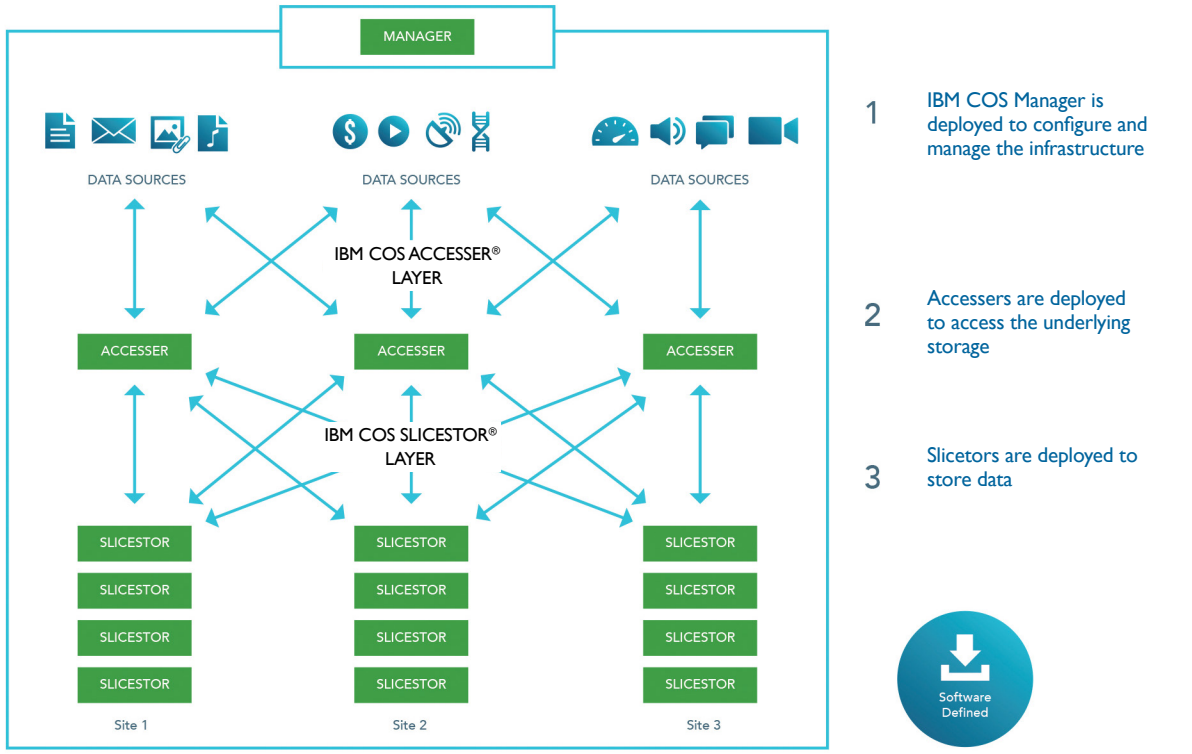

Figure 1: IBM COS System

The three types of nodes are the IBM® Cloud Object Storage Manager, IBM® Cloud Object Storage Accesser®, and IBM® Cloud Object Storage Slicestor®. Each IBM COS System has a single Manager node, which provides out-of-band configuration, administration and monitoring capabilities, one or more Accesser nodes, which provide the storage system endpoint for applications to store and retrieve data, and one or more Slicestor nodes, which provide the data storage capacity for the IBM COS System. The Accesser is a stateless node that presents the storage interface of the IBM COS System to client applications and transforms data using an Information Dispersal Algorithm (IDA). Slicestor nodes receive data to be stored from Accesser nodes on ingest and return data to Accesser nodes as required by reads.

The IDA transforms each object written to the system into a number of slices such that the object can be read bit-perfectly using a subset of those slices. The number of slices created is called the IDA Width (or Width) and the number required to read the data is called the IDA Read Threshold (or Read Threshold). The difference between the Width and the Read Threshold is the maximum number of slices that can be lost or temporarily unavailable while still maintaining the ability to read the object. For example, in a system with a width of 12 and threshold of 7 data can be read even if 5 of the 12 stored slices cannot be read.

Storage capacity is provided by a group of Slicestor nodes, which are referred to as a storage pool. In the diagram in Figure 1 above, 12 Slicestor nodes are grouped in a storage pool. A single IBM COS System may have one or multiple storage pools.

A Vault is not part of the physical architecture, but is an important concept in an IBM COS System. A Vault is a logical container or a virtual storage space, upon which reliability, data transformation options (e.g., IBM® COS SecureSlice™ and IDA algorithm), and access control policies may be defined. Multiple vaults can be provisioned on the same storage pool.

The Information Dispersal Algorithm combines encryption and erasure-coding techniques designed to transform the data in a way that enables highly reliable and available storage without making copies of the data as would be required by traditional storage architectures. By enabling reliability and availability without storing multiple copies of the data, the IBM COS System can offer significant TCO savings. For more detail on how Information Dispersal works, please see *The Definitive Guide to IBM Cloud Object Storage Dispersed Storage*.[1] For more information on the TCO benefits of the IBM COS System, please see *The Total Economic Impact™ of IBM Cloud Object Storage* from Forrester Research.[2]

Scalability

Storage systems should be able to handle virtually all current storage requirements and scale to help meet anticipated needs 5 to 10 years down the road in a single storage system and a single namespace, rather than via an ever-increasing number of limited-capacity storage silos. IBM COS software has been tried and tested at web-scale with production deployments exceeding up to 100 PBs of capacity at multiple customers, and the ability to scale to Exabytes (EB) while maintaining reliability, availability, manageability, and remaining more cost-effective. Whether your current needs are less than a Petabyte (PB), 10 PBs, or hundreds of PBs, the IBM COS solution can help meet that requirement due to the following capabilities.

Capabilities

Scalable, Internet-style, scale-out, architecture (distributed, shared-nothing, peer-to-peer).

Yottabyte-scale global namespace with 10^{38} object IDs available per vault.

Increase storage capacity and performance by adding Slicestor storage nodes.

Scale up to thousands of Slicestor storage nodes in a single system.

No practical limit on the number of Accessers per IBM COS System. Deploy as needed to help meet access performance requirements.

Network installation of IBM COS software across the IBM COS System nodes using PXE helps enable tens of PBs of capacity to be provisioned in near minutes following hardware installation.

Near linear increases in system throughput and HTTP operations per second as the system grows.

Security

From built-in encryption of data at rest and in motion to a range of authentication and access control options, the IBM COS solution includes a wide range of capabilities designed to help you meet your security requirements. These security capabilities have been implemented to help enable better security without compromising scalability, availability, ease of management, or economic efficiency.

Capabilities

Transmission and storage of data is inherently private and is designed with security. No copy of the data resides in any single disk, node, or location. Data is encrypted in motion using TLS and at rest using SecureSlice encryption.

SecureSlice encryption provides a high level of confidentiality of data at rest on Slicestor storage nodes as long as no more than N Slicestor nodes have their data exposed, where N = IDA Read Threshold-1. In typical IBM COS System deployments N ranges from 4 to 25 depending on scale and configuration.

SecureSlice is a standard product feature, with no additional license fee.

SecureSlice can be configured to use any of the following combinations of encryption and data integrity algorithms::

- RC4-128 encryption with MD5-128 Hash for data integrity
- AES-128 encryption with MD5-128 Hash for data integrity
- AES-256 encryption with SHA-256 Hash for data integrity

Total Layer Security (TLS) is supported on network connections within the IBM COS System for data-in-motion protection.

TLS is supported on Client to Accesser networks for data-in-motion protection.

Multiple authentication methods supported for data and management access:

- Username and password (internally managed)
- Active directory or OpenLDAP server
- S3 Secret access key
- OpenStack Keystone Identity Service
- Public Key Infrastructure (PKI) certificate and private key

One user may authenticate using multiple mechanisms (e.g., a username and password as well as a certificate and private key).

Certificate-based authentication of each node added to an IBM COS System (Manager, Accesser and Slicestor) by the IBM COS System administrator.

Critical configuration information is communicated in a security-enhanced manner or digitally signed to help prevent a potential outsider from assuming an administrator's role.

Vault-level access control: For virtually any vault, a user may be granted owner, read/write or read-only privileges.

Vaults may also be configured with classless inter-domain routing (CIDR)-schemed IP Access restrictions.

When vault security is not desired for a vault, it may be configured as anonymous read or anonymous read/write. This feature enables access to content in a vault without authentication.

Object-level Access Control List (ACL) support in the S3-compatible Cloud Storage Object API, enables the association of an ACL with each individual object.

Role-Based Access Control in the IBM COS Manager provides managed restriction of access to functionality by role for the following 6 roles:

- Super user
- System administrator – Security officer
- Operator
- Vault provisioner
- Vault user

Reliability and availability

In an IBM COS System, the reliability and availability characteristics of the system are configurable. For extremely demanding applications, reliability of 16 nines and availability of 8 nines can be provided. More commonly, customers can configure for more typical levels of reliability and availability and potentially achieve economic savings as a result. The configurability of the IBM COS System allows you to choose the combination of reliability, availability, and economic efficiency that best suits your requirements. Data durability is designed to be maintained over time by built-in integrity checking and self-repair capabilities.

Capabilities

The IBM COS Information Dispersal Algorithm (IDA) is designed to enable durable storage, helping provide reliability and availability without storing multiple copies of the data. Availability and reliability are maintained in spite of a potential failure of hard drives and other components, complete failure of IBM COS System nodes, and site outage or destruction.

For an explanation of how the Information Dispersal Algorithm works in an IBM COS System please see *The Definitive Guide to IBM COS Dispersed Storage*.[3]

Ability to configure the IDA to provide very high levels of reliability (exceeding up to 16 nines) or availability (8 nines), or to provide a lower level of reliability or availability with less physical storage capacity needed for the same usable capacity.

Flexibility to use different IDA configurations for different vaults.

Distributed Rebuilder uses all Slicestor nodes in the system to identify slices that are missing or corrupt and perform the necessary repair, such as:

- Restores slice data lost due to failure
- Restores slice data corrupted by a disk-level
- Unrecoverable Read Error

Disk lifecycle management – low-level monitoring of disk health in Slicestor nodes allows data to be moved from a failing drive to a healthy drive before the drive fails.

Multi-level data integrity – Incorporates checksums to handle physical media errors that often occur in large-scale storage systems. Integrity is checked at both the slice and object levels. Corrupted slices are not used and are repaired by the Distributed Rebuilder.

Manageability

The manageability of an IBM COS System helps enable storage administrators to handle up to 15X the storage capacity, freeing time for them to invest in other tasks. The IBM COS System is designed to provide always-on availability and can continue to serve storage applications while completing tasks that would require scheduled downtime in the majority of traditional storage systems. Software upgrades, hardware maintenance, storage capacity expansion, hardware refresh and physical relocation of the storage system are all supported with virtually zero downtime.

Capabilities

The IBM COS Manager is an out-of-band management console for the entire system. It provides robust configuration, administration, event monitoring, and reporting, as well as Role-Based Access Control support. The IBM COS Manager can be accessed via a web GUI, which provides a single pane of glass view into the IBM COS System, and via a set of Management APIs.

A robust set of management and monitoring APIs help enable integration with customer-provided management or monitoring tools:

- IBM COS Manager REST API
- SNMPv3
- RESTful device state and statistics interface – syslog
- Report export via HTTP command

Always-on design, virtually no downtime required to:

- Upgrade to a new version of IBM COS software
- Add Slicestors to increase storage capacity
- Add Accesser to increase access layer throughput
- Perform hardware maintenance
- Refresh hardware
- Move hardware to a new site/data center
- Change the number of sites/data centers across which the system is deployed (from 2 sites to 3 for example)

Top-level green/yellow/red indicator of the health of physical hardware underpinning IBM COS Manager, Accesser®, and Slicestor® nodes based on monitoring of drive, fan, and power supply status.

Drill-down capability to see more detailed information on any individual server in the IBM COS System.

The IBM COS Manager provides the following information for each Manager, Accesser, or Slicestor node:

- Node health
- IP Address
- Model
- SW version

Additional information for each individual Slicestor data drive:

- Drive health
- Drive capacity
- Drive model
- Drive serial number
- Drive firmware

The IBM COS Manager provides the following information for each vault:

- Name
- Description
- Creation date
- Vault health
- Capacity used (raw and usable)
- IDA width and Read Threshold
- Soft quota
- Hard quota
- SecureSlice enabled/disabled
- Object versioning enabled/disabled
- Delete restricted yes/no

The IBM COS Manager provides the following information for each storage pool:

- Name
- Capacity
- Slicestor nodes – Deployed vaults

Graphs provide visualization of key performance, system health, and usage indicators. The same data used to create these graphs is available via a REST interface for processing by other tools.

- Storage pool capacity and utilization
- Vault space used
- Client to Accesser throughput
- Accesser to Slicestor throughput
- Rebuild activity
- Node disk usage (MB/s)
- Node CPU usage
- Node network usage
- CPU temperature
- Fan speeds
- Hard drive temperature

Near real-time incident stream provides a to-do list of issues requiring operator attention to maintain the health of the IBM COS System at the current time.

Near real-time event stream provides a historical record system including conditions impacting nodes (IBM COS Manager, Accesser®, or Slicestor®), vaults, and storage pools. Event stream data can be filtered in a variety of ways when looking for specific events or patterns.

Generate alerts using email, SNMP traps, or syslog forwarding from incidents and events.

Audit stream provides a historical record of changes in the configuration of the system.

System provided reports provide information on IBM COS System health and configuration, which can be viewed via the IBM COS Manager or exported. Export can be done from the IBM COS Manager or via a RESTful interface. The following reports are included:

- Disk drive and devices
- IBM COS System compliance
- Storage pool usage
- Vault summary
- Device summary
- Failed FRU report
- Event report
- Firmware report

Extensive logging of activity for use in support situations with automated collection and options to forward logs to external syslog server or post using HTTP to the IBM COS support team.

Troubleshooting console provides an environment for low-level debugging of individual Accesser and Slicestor nodes within the Manager GUI.

Phone Home capability delivers relevant system information to IBM's support organization to enable proactive maintenance and help reduce issue resolution time.

Flexibility

IBM COS's software defined storage solution requires no IBM specific or proprietary hardware. Using IBM COS software with industry standard hardware systems helps enable an IT organization to potentially achieve a significantly lower cost per usable terabyte curve helping the organization to economically manage rapidly increasing storage capacity requirements.

Capabilities

SW-defined storage – runs on a wide range of industry standard hardware platforms.

IBM-certified hardware platforms from multiple suppliers. This certification includes integration with IBM's hardware component health monitoring and Disk Lifecycle Management functionality described earlier. The following certification types are included:

- IBM testing of the specific combination of HW and SW you deploy can result in significantly reduced time to production on initial deployment
- IBM guidance on optimal configuration of certified platforms
- Close monitoring of hardware health delivers more timely notification of hardware issues needing attention
- Disk Lifecycle Management allows data to be migrated off a failing Slicestor hard drive before that drive fails, helping to significantly reduce the amount of data that must be rebuilt due to disk failure
- IBM support staff are typically familiar with the specific combination of hardware and software you are using

An IBM COS System can be deployed in a single site with data protection via Information Dispersal without RAID and without replication.

An IBM COS System can be deployed across 2 sites using a vault- mirroring configuration. This configuration provides site level fault tolerance by maintaining a copy of the data in each of the two sites.

An IBM COS System can be deployed across three or more sites with Information Dispersal providing site-level fault tolerance without replication.

The number of sites utilized can be changed in the future in a way that requires no system downtime. For example, a system initially deployed across two sites can be later reconfigured and deployed across three sites.

The flexibility to specify different configurations across vaults sharing the same pool of Slicestor storage nodes helps enable the IBM COS System to support multiple applications and workloads with different requirements and characteristics. Vault level settings include the following:

- IDA width and Read Threshold
- Soft quota
- Hard quota
- SecureSlice enabled/disabled
- Object versioning enabled/disabled
- Delete restricted yes/no
- Accesser devices provisioned

The best IDA configuration varies based on factors like capacity (1, 10, 100, or 1000 PBs) and deployment 1, 2, 3 or more sites. Support for a wide range of different IDA widths and Read thresholds can provide valuable flexibility that is not present in replication-based storage systems or systems offering only a few erasure-coding settings.

Multiple storage interfaces are supported, including the following:

- Cloud Storage Object API
- an Amazon S3-compatible RESTful interface
- Open Stack Object API
- a Swift-compatible RESTful interface
- Simple Object over HTTP
- NFS via cloud storage gateway from one of IBM's technology partners
- SMB/CIFS/iSCSI from one of IBM's technology partners

Two options to consider, for increasing storage capacity over time:

- Add a new storage pool and create new vaults on it
- Increase the capacity of an existing storage pool making more capacity available for new and existing vaults

Support for multiple storage pools in a single IBM COS System offers flexibility to provision different combinations of performance/durability/cost for different applications or workloads within a single IBM COS System rather than deploying application-specific storage systems.

IBM® COS SmartRead™ feature optimizes read performance and allows consistent performance to be maintained even if some storage system components are impaired.

IBM® COS SmartWrite™ feature improves availability and allows consistent performance even when some individual components are slow to respond.

Adaptively increases the number of active WAN connections if latency becomes too high.

Packed Slice Storage optimizes performance and storage efficiency for small objects allowing the IBM COS System to efficiently handle a wide range of workloads including those with significant numbers of small objects.

Economics/TCO

IBM COS software uses distinctive Information Dispersal Algorithms (IDAs) to protect digital assets, rather than relying on storage of multiple copies, which can reduce the physical storage capacity required by more than 50 percent. The IDA encodes the data before storing it in such a way that data remains available in spite of numerous hard drive or server failures and even site outages, while only storing a single copy of that data. Also, the IBM COS System may be utilized up to 99 percent of its usable capacity without a noticeable drop-off in performance that is typical in most traditional storage systems as utilizations increases.

IBM COS systems can achieve TCO savings of up to 80 percent compared to most traditional storage systems as well as public cloud storage. TCO savings reflect not only reduced acquisition cost, but saving on power, cooling, and space due to the smaller hardware footprint, as well as lower manpower costs in storage management and operations.

About IBM Cloud Object Storage

IBM Cloud Object Storage provides organizations the flexibility, scale and simplicity required to store, manage and access today's rapidly growing unstructured data in a hybrid cloud environment. Relied upon by some of the world's largest repositories, our proven solutions turn storage challenges into business advantage by reducing storage costs while reliably supporting both traditional and emerging cloud-born workloads for enterprise mobile, social, analytics and cognitive computing. IBM Cloud Object Storage is built on technology from object storage leader Cleversafe, acquired by IBM in 2015.

For more information

Please call: 312-423-6640 or email: sales@cleversafe.com or visit the Cleversafe website: www.cleversafe.com.

To learn more about IBM Cloud computing, please visit http://www.ibm.com/cloud-computing/infrastructure/object-storage/

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
July 2016

1 The Definitive Guide to IBM COS Dispersed Storage, 2016, http://www.ibm.com/common/ssi/cgi-bin/ssialias?subtype=WH&infotype=SA&htmlfid=TSW03335USEN&attachment=TSW03335USEN.PDF

2 The Total Economic Impact™ of IBM Cloud Object Storage, a Forrester Total Economic Impact™ Study Commissioned by IBM, Forrester Research, Inc., March 2015, http://www.ibm.com/common/ssi/cgi-bin/ssialias?subtype=WH&infotype=SA&htmlfid=TSL03266USEN&attachment=TSL03266USEN.PDF

3 The Definitive Guide to IBM COS Dispersed Storage, 2016, op. cit.

Please Recycle