CLOUDIAN®

Modern Enterprise Data Protection with Cloudian

CLOUDIAN WHITEPAPER

Introduction

Not so long ago, if you asked a business executive, “what is the most irreplaceable part of your business?” they might have pointed at their building or their employees. The really forward-thinking execs may have chosen their customers. But the reality is that businesses have insurance to deal with the loss of a building; they can hire new employees; and they can acquire new customers. The one thing that is truly irreplaceable is their data. Lose that and you really are out of business.

The potential risks to your data can be as subtle as the malfunction of tiny areas on a disk drive or as enormous as the failure of an entire data center. That means it’s not enough just to protect your data — you must also select the best way to protect it. To deliver that protection in an economical and rapidly recoverable way requires a system that gives you complete control to determine levels of data protection based on the type of data.

In this paper, we’ll talk about the science of data protection and the advantages a Cloudian object storage system provides in keeping your data safe, easily recoverable and highly available.

Durability vs. Availability

To fully understand the strategies underpinning modern data protection, you need to know the terms used to describe the concepts that make up that strategy.

Data Durability: Durability refers to data-level redundancy and protection with an aim to ensure data is never lost or compromised.

Some of the key parameters used in calculating durability levels for a data protected system include usable storage capacity, number of unique locations, number of storage nodes and drives in each node, drive capacities, bit-error rate, annual failure rate of drives, average time to repair a failed drive and rebuild associated data, and the number of nodes and the number of sites that can be lost without impacting data availability.

Data Availability: Data availability describes the rate at which data continues to be available at a required level of performance to end users and applications in situations ranging from normal through “disastrous.” It defines the degree or extent to which data is readily usable along with the necessary IT and management procedures, tools and technologies required to enable, manage and continue to make data available. A problem may impact data availability without impacting data durability.

Object Storage and Data Protection

Unlike other storage options, object stores protect data as part of their design. The cluster replicates the data and can manage its recovery automatically. Objects are fully protected through the use of several possible methods in tandem:

1. Redundant Nodes

A “node” refers to a rack-mounted server filled with internal disks. Each disk is formatted and controlled by the object store as an individual place to store objects. In most systems, nodes are spread across the data center, or among multiple data centers as a way to provide protection in case of complete site failure. None of the disks need to use RAID protection; the RAID approach used to be the go-to technology for data protection, but the architecture of RAID arrays imposes a time penalty that today is considered unacceptable.

2. Full Object Copy

The basic method of protecting data is to make multiple copies of the data object and ensure the copies are located on different physical devices. An example of this is to use RAID mirroring across two different disk drive devices in order to tolerate a disk drive failure. An object storage system extends this concept to the cluster node level, whereby data copies are placed across different servers rather than drives, providing greater protection and availability against other server component failures as well as the disk drives. Typically, a system would be configured to make three copies of data across the cluster, although it is possible to have more copies for greater protection. A disk failure or node failure is not a cause for worry since there are always at least two other copies of the object available within the cluster.

Although mirroring is the simplest implementation of data protection and delivers better performance, the cost of storage is greater due to the high storage capacity overhead. Every additional recovery point costs 100% of the original data object.

3. Erasure Coding

Erasure coding (EC) is an alternative method of data protection in which data is broken into fragments, expanded and encoded with redundant data pieces, and stored across a set of different locations or storage media. If data becomes corrupted at some point in the disk storage process, it can be reconstructed using information about the data that’s stored elsewhere in the array. For example, in an EC 10/6 configuration, the 10 refers to the original amount of data and the 6 refers to amount of extra data added to protect from failure. Six drives, nodes or geographic locations could be lost or unavailable, and the original file would still be recoverable. Adding additional recovery points by increasing the number of parity fragments within a EC set, e.g. 10/7 configuration, improves data protection rates at a much lower cost than object copies.

Erasure coding is used instead of traditional RAID because of it provides ability to scale across cluster nodes rather than being restricted to only managing disk devices within a single server as well as providing the flexibility on the number of parity fragments.

4. Disaster Recovery with Offsite Replication

To protect against an entire site disaster, data replication is used to make copies of data distributed across multiple sites. This is done primarily to create up-to-date copies of data so that in the event of a disaster data may be restored.

Replication may be synchronous or asynchronous. Synchronous replication takes place in real time, and is preferred for applications with low or zero recovery point objectives (RPO) that mean that data must be able to be restored to the point in time of failure. This ensures that data is consistent — i.e., exactly the same — in multiple locations. This is crucial for certain regulated industries and for those organizations that have high-value data. This replication approach is typically more expensive to implement as the network infrastructure between sites becomes a limiting factor and money needs to be spent on minimizing latency to avoid application performance impact.

Asynchronous replication is typically time-delayed. It is designed to work over distances and requires less bandwidth as the application is not directly impacted by waiting for a remote write acknowledgement. This replication is intended for businesses that can withstand lengthier recovery point objectives. Because there is a delay in the copy time, the two data copies may not always be identical and therefore inconsistent. It is important to understand and manage your recovery points so that you can return your data sets to a known consistent state.

Cloudian's HyperStore allows an organization to select the appropriate data protection policy for each data set within a cluster according to the specific business and technical requirements for the protection of that data. HyperStore applies storage policies at the user bucket level. Storage policies may even support the combination of both replication strategies with the use of dynamic consistency levels, where the system tries to achieve synchronous replication to maintain data consistency, but will switch down to asynchronous replication if needed to maintain data availability.

How Many 9s Do You Need?

Ask any IT manager how much data they are prepared to lose, and the answer is always the same — zero.

The possibility of loss always exists, but companies can take steps to reduce that possibility, and do so in an economically responsible way.

Data durability is often described in terms of 9s — a shorthand method for describing the percentage of data durability. For example, AWS S3 storage durability is set at eleven 9s (99.999999999%), so for 10,000 objects, a single object may be lost once in 10 million years. Higher rates of durability are available, but, typically, the more 9s you add, the more you pay for your storage infrastructure.

Some of the key parameters used in calculating durability levels for a data-protected system include usable storage capacity, number of unique locations, number of storage nodes and drives in each node, drive capacities, bit-error rate, annual failure rate of drives, average time to repair a failed drive and rebuild associated data, and the number of nodes and the number of sites that can be lost without impacting data availability.

Varying the number of nodes, drives per node, and number of data centers hosting the object data (and extra data copies) directly impacts the level of data durability and ability to tolerate the risk of a failure and losing an object. As an example, data durability will be much lower when object data is hosted in a single node in a single data center as compared with hosting the same object data and copies on multiple nodes, each in multiple data centers, which enables much higher data durability and protection.

All data is valuable — so whether only for a small amount of data or with very large volumes of data, the concern is for failures or errors in the underlying storage media, particularly in the form of uncorrectable bit errors. After you have written an enormous amount of data, one of those bytes might not be correct. A high degree of data durability helps protect against this.

Many enterprises may be happy to have their storage design match Amazon Web Services' 11 9s, without the need for more. For some organizations, however, having more than 11 9s is important. Cloudian offers any level of 9s to suit customer needs. For example, some end users with significant regulatory requirements may opt to deploy more than 11 9s as a precaution against inadvertently losing critical data.

AWS S3 storage durability is set at eleven 9s (99.999999999%), so for 10,000 objects, a single object may be lost once in 10 million years.

Cloudian Solution Strengths

Cloudian's HyperStore object storage system brings all the data protection benefits inherent in object storage, plus additional benefits that make it a particularly strong candidate for backup applications:

Erasure Coding and Replication Capabilities

Cloudian HyperStore is flexible, permitting organizations to design their storage protection with erasure coding and /or replication. Unlike others, Cloudian does not require customers to commit to a single scheme, but can employ the right capability where needed. Replication may be a better choice for some clusters or locations, with EC or even different levels of EC for other parts of their storage. Another advantage with Cloudian is that when additional capacity is needed, customers can simply add additional nodes without having to employ the same EC configuration, as others can require. This added flexibility keeps the data protection cost-effective and affordable.

Bucket-Level Policy Setting

Rather than forcing you to set the same high-level backup policies across your entire storage infrastructure, HyperStore allows you to set up and manage policies flexibly and granularly at a bucket level. Cloudian's bucket-level policy setting provides highly precise granularity for SLA and management control through a number of parameters:

- Data Protection Control — allows you to select replication or erasure coding of data, plus single or multi-site data distribution
- Consistency Level — Provides control of replication techniques (synchronous vs. asynchronous)
- Access Permissions — Manages user and group control access to data
- Disaster Recovery — Enables data replication to public cloud
- Encryption — Offers data at rest protection for security compliance
- Compression — Reduces the effective raw storage used to store data objects
- Data Size Threshold — Provides variable storage location of data based upon the data object size
- Lifecycle Policies — Manages data management rules for tiering and data expiration

Architecture and Configuration

The HyperStore architecture allows managers to adjust the level of data consistency across the system. The system compares replicated data to ensure that copies are identical. If they are not, the system uses a time stamp to determine which copy is newer and bring other versions up to date. The system can also compare replicated data and uses its one internal logic to determine if discrepancies are caused by "disk rot," or malfunctions of the storage media; if this is detected HyperStore creates a new copy and then isolates the malfunctioning sector of media. All these activities are then reported via email to the storage manager. Managers can set policies to control the frequency of consistency checks to match the need for system performance.

In the event of a node failure or outage, HyperStore can stage data targeted at the failed node onto another node, then automatically move data back when the original target is back on-line. The HyperStore cluster automatically balances loads across the nodes with the express intention of not overtaxing individual nodes and disk drives, thus saving on hardware maintenance expenses. HyperStore can be architected to allow nodes of different sizes; for maximum security, many small nodes may be used to minimize data loss in the event of a drive failure. Larger nodes can be used if capacity is the major concern.

A key HyperStore benefit is that policies may be set across buckets or across the entire system, providing extensive management flexibility. For objectives where data durability is critical, managers can set durability at whatever level of protection is needed — 14 9s, 18 9s or even 26 9s or higher — in addition to thousands of different policies for protection of different data sets across a system if so desired.

The metadata engine is a critical part of the HyperStore architecture; it helps maintain the integrity of the global namespace while at the same time ensuring the searchability of the data being stored. The global namespace covers all data in all nodes, in all locations — including multiple regions globally as well as public cloud deployments.

HyperStore devices also include redundant power supplies and Ethernet ports to head off hardware failures that can impede performance and endanger data integrity.

HyperStore Algorithm

HyperStore employs a unique algorithm that detects data corruption, bit-rot and other issues and then automatically cures them. Automated version creation adds a level of protection to the data and the system overall.

The system constantly runs a background scan to verify the integrity of each object; if it detects any damage or missing data, the system will automatically build a new copy of the object.

The internal Cassandra database disk records recent activity and also serves as a record of all recent actions within the cluster. Because a failure of this node during a restore could cause significant problems, all activities written to the Cassandra disk are mirrored. Mirroring avoids the rare but extremely damaging event of a Cassandra disk failure during a restore.

Conclusion

Deciding on what level of data protection your data requires is a matter of understanding your company's needs, but also understanding the variables involved. Even having a level of data durability at 11 9s — which is 1 million times greater than the standard 20 years ago — doesn't mean that an even higher level of data protection isn't right for your company.

Cloudian allows you to manage the variables around data protection — and do so on a granular level, using different policies for different buckets. It gives managers any level of protection that is required – and enables them to assign different levels of data durability to different parts of their storage environments, and then change levels over time, as needed. Cloudian also offers architectural benefits that preserve data integrity and also limit wear on hardware to extend its lifecycle.

Cloudian HyperStore offers greater and more cost-efficient control over data protection and enables its users to do what's best for their organizations' data, today and for the future, with the same architecture.

CLOUDIAN®

CLOUDIAN, INC.

177 Bovet Road, Suite 450, San Mateo, CA 94402
Tel: 1.650.227.2380 | Email: info@cloudian.com | Web: cloudian.com