CASE STUDY

Object Storage

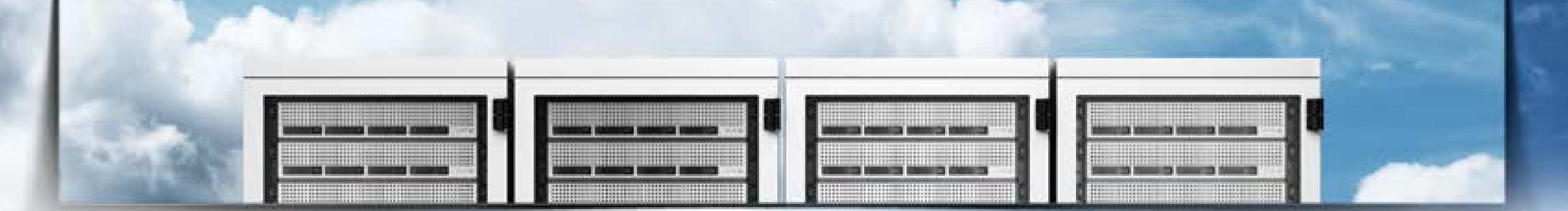
Why RAID is Wrong at Hyperscale

By Bob Murphy, DataDirect Networks

DataDirect™ NETWORKS

CLOUD ADVISORY COUNCIL

www.cloudadvisorycouncil.com



Why RAID is Wrong at Hyperscale

When making the leap to hyperscale storage systems, RAID should be left behind. It's time to consider object storage.

The following case study document is provided as courtesy of the Cloud Advisory Council.

For quite some time, it's been widely accepted that the best way to protect data on disks is with a redundant array of independent disks (RAID). RAID is a storage technology that distributes data across multiple disk drives. In the case of RAID,if a drive fails, it's possible to recover data from the data on the remaining good drives using additional parity data — an extra bit of information calculated from the stored data that helps with recovery when lost.

In fact, one of the dirtiest secrets in the storage industry today is its near-universal dependency on RAID-protected disks to prevent the loss of data. And it's true — at one time, RAID protection was a logical answer to managing data. However, that was when the industry spoke of drive sizes in gigabytes, and storage systems held terabytes of data on dozens or hundreds of drives. This is no longer tenable with today's multi-terabyte drives and multi-petabyte systems that contain thousands of drives.

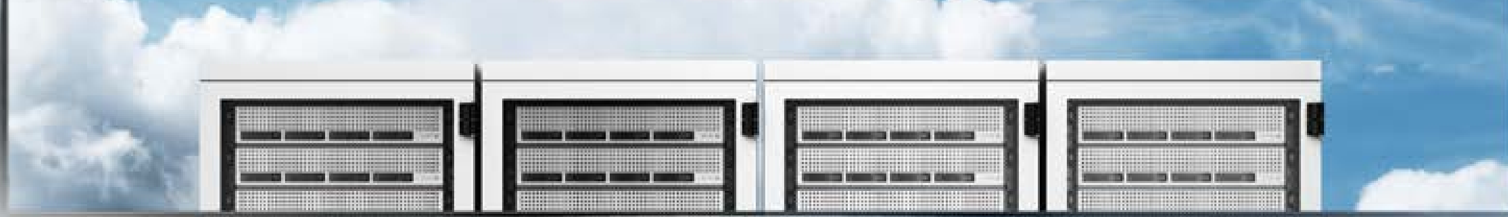
Today, data is growing exponentially, filling billions of files and taking up capacity — all of which must be quickly and efficiently shared between multiple data centers. As data grows, traditional storage solutions and approaches like RAID become complex to deploy, hard to manage and scale, and overly expensive. Now, a multi-terabyte drive failure takes days to rebuild, strangling the system's ability to provide data service. In addition, the sheer number of drives means that those failures are more and more likely to occur.

Among the most critical problems affecting RAID-protected data today is the length of time needed to recover data in the event of system failure, which — due to the fact that drive sizes have ballooned to three terabytes or more — can stretch untenable lengths. This lengthy recovery process is largely because in the event of system failure, the entire storage system becomes unresponsive and must be effectively taken offline for the complex parity calculations needed to rebuild the failed drive. While the system's processing resources are consumed with the rebuild, fewer resources are available for data services.

Most enterprise storage systems today rely on RAID 6 for protecting data in case of failure. Systems that use RAID 6 for protection can survive two drive failures and still recover all stored data. However, using RAID 6 in hyperscale environments pits usable storage against purchased raw capacity. With RAID 6, the usable storage is only 60 percent of raw capacity due to the extra stored parity data, creating data storage inefficiencies. And in today's big data environments, this means significant inefficiencies — which most organizations cannot afford to have.

This new reality has introduced a new storage approach for hyperscale environments: object storage. Object storage enables organizations to store unstructured data in the cloud, and it can also scale to billions of files and allow access from anywhere at any time. With the new reality of object storage and hyperscale storage systems that store multiple petabytes, many organizations have turned to techniques that are not RAID-based to protect data stored on disks.

CLOUD ADVISORY COUNCIL



This is especially the case in situations that entail storing immutable, unstructured data. This includes data driving the lion's share of data growth in hyperscale systems, such as photos, videos and music files, as well as PDF documents, medical records, scientific instrument data, and intelligence, surveillance and reconnaissance data.

Replication is central to these new techniques. Data stored on a disk involves simply replicating it to another disk or disks usually on a system in a remote location — a scenario also known as cloud storage. This provides both protection against a failed disk drive and built-in disaster recovery.
If one storage system fails, the data is instantly retrievable from the remote system. Hyperscale storage systems, such as Amazon's Simple Storage Service, OpenStack Swift and DDN's Web Object Scaler (WOS), are examples of hyperscale object storage implementations that make use of replication.

With a RAID 6-protected system, a failed drive is considered an emergency situation, requiring an urgent response to replace the drive since the system has only one remaining disk left before data is permanently lost. In contrast, a replicated object storage system can tolerate multiple drive failures and requires no urgent emergency action. As a drive locations to another drive to reestablish the replication policy, over and over again. This enables organizations to replace drives as they are able, such as during planned fails, data simply replicates from one of the surviving maintenance periods, and requires no emergency response. This becomes increasingly valuable in today's hyperscale world comprised of thousands of deployed drives.

The one downside of replication is that multiple copies of data obviously take up more disk space than a single RAID-protected copy. However, when compared to the storage inefficiencies introduced by RAID 6, this actually isn't that much disk space overall.

That's where erasure coding comes in as a viable storage option. From a storage perspective, erasure coding breaks objects into smaller pieces. It then spreads those pieces across multiple drives in a way so that the data is still recoverable, even if a number of those drives fail.

When applied correctly at the hyperscale level, object storage erasure coding entails erasing codes across multiple drives in a single local storage system, providing protection against a dual disk failure, as can happen in RAID 6. This application of local erasure coding is particularly valuable in high performance applications, where fast data storage and quick access to information is critical.

While similar to RAID 6, erasure coding used to protect against a double disk failure is not as onerous as RAID is in terms of time required for system rebuilding.

That's because with a rebuilt RAID-protected disk, parity calculations are required for every block on the disk, whether or not those blocks actually contained data before the problem occurred. With erasure coding, recovery calculations are only done on the areas where data was lost, significantly shortening recovery times.

Another appealing benefit is the disk reduction enabled by object storage. Many CIOs have realized that by reducing disks and system complexity, the total cost of storage ownership can drop by more than 60 percent due to the reduced power, cooling and staff time required for system management.

Not surprisingly, the federal government is showing great interest in object/cloud storage, and the uses within that i ndustry will likely set the stage for other industries to follow suit accordingly. Originally, the federal government found object storage appealing for benefits such as improved

(CONTINUED)
collaboration and data security. However, in the current budget environment, additional benefits have helped drive its appeal, including object storage's ability to help reduce disk cost by 60 percent in a multi-petabyte environment, which can have a significant impact on an agency's bottom line.

As has been shown, hyperscale storage systems can no longer rely on RAID data protection due to the multi-day rebuild periods, lost usable storage efficiency and compromised system performance, among other notable factors. Given the reality facing today's hyperscale immutable unstructured data storage systems, it's time for organizations to consider object storage and to increase their reliance on multi-site replication and erasure coding. Administrators and users alike will experience greater data protection and improved service as a result.

About the Author

Bob Murphy is the senior object storage product marketing manager for DataDirect Networks.
DataDirect™ NETWORKS

CLOUD ADVISORY COUNCIL

www.cloudadvisorycouncil.com