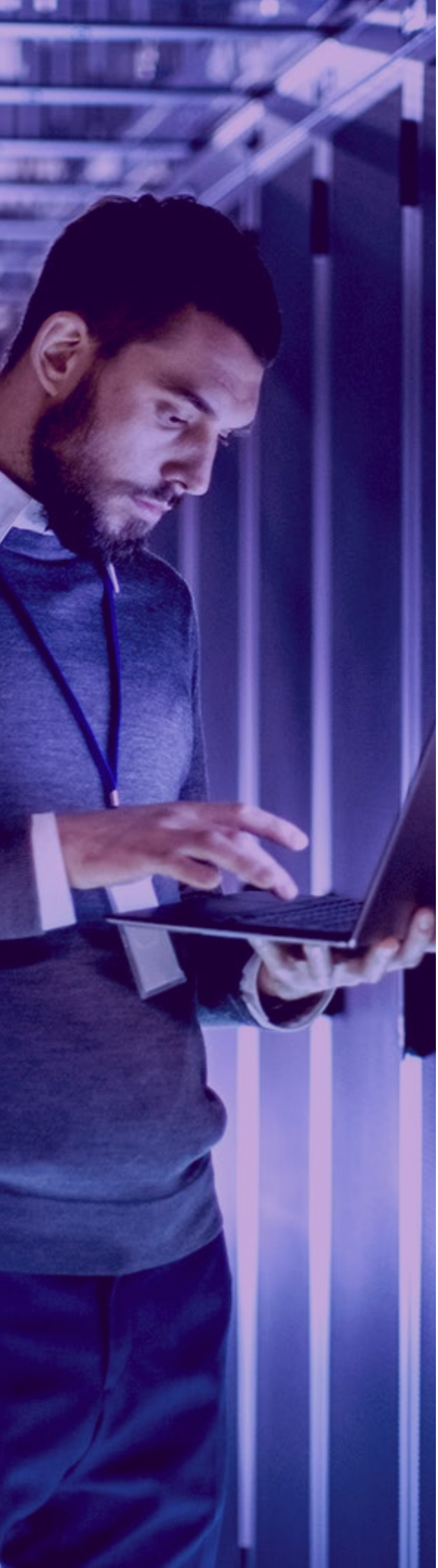
LOADBALANCER

WHITE PAPER

The essential ingredient in successful object storage solutions

Improving the availability and performance of large-scale data storage environments with load balancers



Introduction

The volume of data collected and stored by organizations globally is growing at an unprecedented pace. We're all generating more digital information, and so are the billions of internet-connected devices uploading data every second, on everything from vehicle movement to a home's temperature. At the same time, organizations are storing larger volumes of unstructured data, like social media posts, images and videos, and want to make all their diverse data assets available to analyze and use in the future.

Object storage is the next generation storage technology that's evolved to address these challenges. Infinitely scalable, object storage solutions won't only cope with the exceptional data growth anticipated; they're also designed to manage the inestimable data volumes resulting from machine learning and artificial intelligence in future. Object storage uses metadata to make all kinds of data easier to categorize, find and use for business intelligence and analysis. It also improves data protection, enabling organizations to identify if data has been accessed without authorization.

Leading object storage solutions vendors recommend that load balancers are installed with their products. Load balancers help ensure zero downtime, by automatically redirecting user traffic away from storage nodes with failed disks or seamlessly routing traffic to an alternative data center in the event of a local outage. Load balancers help optimize application performance by distributing traffic evenly across all storage clusters and preventing network bottlenecks. In industries like healthcare and media broadcast, it's absolutely crucial for object storage-based applications to be available to users on demand, 24/7 – load balancers deliver this high availability.

While some object storage vendors leave customers to source their own load balancers, the more established players recognize the advantages of partnering with load balancer vendors. This enables object storage vendors to test compatibility, plan optimum architectures, create step-by-step guides and offer joint support arrangements. Deployments become standardized making it easier to implement and support customer solutions.

If end users source their own load balancers for object storage, there is a lot of choice – many load balancers may work, but not all of them will work well. So organizations should look for load balancers that have been tested with their chosen object storage product and come with pre-configured functionality, deployment guides and a strong track record of success. Load balancers with artificial throughput limits will be unable to cope with the scalability demands of object storage solutions. And, not all load balancer vendors will have the knowledge and flexibility to provide the kind of holistic and specialized support that object storage environments demand so decisions about which product to use should not be made hastily.

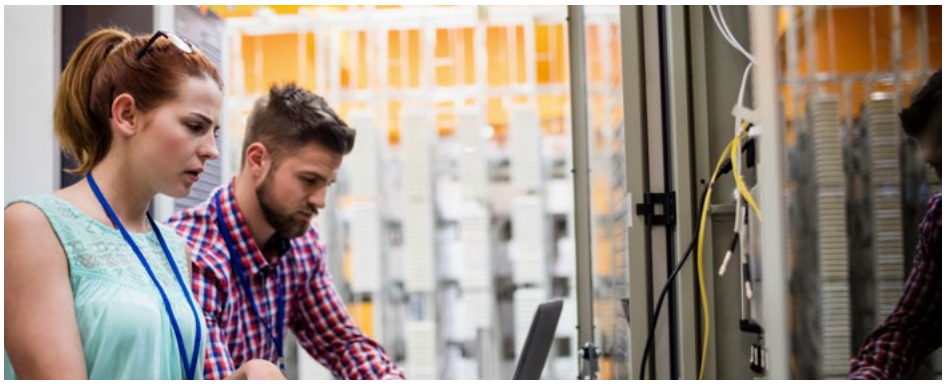
> *'Global Datasphere' will grow from 33 Zettabytes (ZB) in 2018 to 175 ZB by 2025*

Why is object storage important?

Object storage technology has emerged in direct response to the challenges created by data growth. Over the last decade, the volume of digital information generated worldwide has expanded at a tremendous rate, and IDC now predicts that the 'Global Datasphere' will grow from 33 Zettabytes (ZB) in 2018 to 175 ZB by 2025.[1]

This data growth is being fueled in part by the digitization movement; information assets previously stored using physical media, such as tapes, DVDs, CDs and filing cabinets, have been digitized on computers to save space, reduce storage costs and improve access. However, a more significant driver of data growth is the Internet of Things. There are now billions of devices that upload data to the Internet, from CCTV cameras to e-books and remote heating controls. These Internet-connected devices alone are expected to create over 90 ZB of data in 2025.[1]

The increasing demands of industry and data security regulations are also contributing to data growth. Faced with the need to comply with General Data Protection Regulation (GDPR) and similar legislation, organizations now have to record more data about 'who did what' in business processes, track changes to documents and record approval stages to satisfy the requirements of auditors. The end result is that data volumes are increasing not by the day, but by the second. According to one forecast, every individual in the world will generate, on average, 1.7 megabytes of data, every second, during the course of 2020.[2]



This data growth is placing a strain on traditional data storage systems, which were not designed to cope with such large data volumes. File-based and block storage methods are running out of space, costly to expand and difficult to manage and maintain. When ESG surveyed 373 IT decision makers in 2015, the rapid growth of data was already identified as a top challenge.[3] With the acceleration in data growth that has occurred in the five years since then, the concerns that these IT professionals expressed were well-founded. Data growth was then, is now, and will continue to be, a significant challenge for all organizations.

1 IDC White Paper, sponsored by Seagate, 'Data Age 2025: The Digitization of the World from Edge to Core', Document #US44413318, November 2018.
2 IDC iView, sponsored by EMC, 'Universe Study: Big Data, Bigger Digital Shadows and Biggest Growth in the Far East', December 2012
3 ESG White Paper, commissioned by Cleversafe, an IBM Company, 'The Rise of Object Storage', September 2015.

Yet, data growth isn't the only challenge that is necessitating new approaches to data storage. Organizations are now hoarding a wider variety of data, including social media posts, graphics, emails and videos. This data is 'unstructured', so it is not easy to understand what it is, how valuable it is to an organization, where to save it and how long to keep it. For example, a traditional data storage system would not be able to distinguish between a jpg image of a member of staff and a jpg image of a product blueprint.

> *As much as 80% of data is expected to be unstructured by 2025, so it has become an imperative to adopt more effective ways to not only store it, but also access and use it.*

As much as 80% of data is expected to be unstructured by 2025[4], so it has become an imperative to adopt more effective ways to not only store it, but also access and use it. Governments, public sector organizations, commercial businesses and not-for-profit groups are all looking for new ways to use their growing data assets to gain intelligence about their customers, citizens, supporters and markets. In the past, some academics have estimated that as much as 70% of data is not accessed again after being generated[5], but this picture is likely to change, as more and more organizations seek to use business intelligence tools and data analysis techniques to gain value from their 'Big Data'.

4 King T. '80 Percent of Your Data Will Be Unstructured in Five Years' [Internet]. Best Data Management Software, Vendors and Data Science Platforms. 2019 [cited 2020 Feb 13]. Available from: https://solutionsreview.com/datamanagement/80-percent-of-your-data-will-be-unstructured-in-five-years/

5 Leung AW, Pasupathy S, Goodson G, Miller EL. Measurement and analysis of large-scale network file system

What is object storage?

No longer 'new' technology, object storage (sometimes called object-based storage) provides a radically different approach to storing data that is more in tune with both the volume and types of data that organizations wish to store today. Object storage has become popular over recent years in industries such as medical imaging, CCTV security, media and broadcasting, where there is a requirement to store space-intensive, high resolution images and high density videos. Now, however, object storage is gaining increased attention across a far wider range of industries, and, in one recent global survey, 80% of organizations believed object storage could support their top IT initiatives.

> *"Object storage solutions are unlikely to replace traditional file and block storage methods completely – at least in the short term."*

The primary advantage of object storage is that it is scalable. Unlike traditional file and block storage methods, in which data is stored in files and folders in a hierarchy or 'tree' format, object storage puts data into a single, flat space or storehouse, often referred to as a storage pool. Each piece of data is assigned a unique identifier (a 128-bit random number) which enables it to be moved easily from one storage node to another, as the storage environment grows, without impacting user access or performance. The unique identifier also makes it easier for users to find and retrieve data in large data repositories.

Object storage is also the ideal solution to the challenges of unstructured data, as it makes it far easier for organizations to identify and use their data. In an object storage system, each data item is enriched with metadata, which allows the content of the data file to be better understood and analyzed. For example, an x-ray image might include metadata on the patient's age, location and medical condition. Organizations can use metadata to create data storage rules for how long to store certain types of data and group data for analysis in business intelligence systems.

Furthermore, object storage helps organizations to monitor data security and identify any breaches in policy. It does this by attaching a hash signature to each piece of data and automatically changing the hash signature each time the data is accessed or amended. As a result, organizations have a simple, automated way of detecting if any piece of data has been inappropriately accessed or tampered with, as well as proving compliance with data security standards.

Object storage solutions are unlikely to replace traditional file and block storage methods completely – at least in the short term. In environments where users need to create, access and amend files regularly, current storage approaches are likely to remain the best option. On the other hand, object storage solutions are an attractive option for companies that generate large volumes of data through artificial intelligence, machine learning and internet-connected devices, or need to archive large amounts of historical information. In reality, hybrid environments are likely to become the predominant model, with some data stored traditionally and some data stored in object storage.

Why are load balancers essential in object storage environments?

Load balancers fulfil a vital function that object storage solutions themselves cannot. In order for object storage to be scalable it needs have multiple servers, housing multiple disks. The object storage solution itself will have methods of protecting the integrity of the data stored, if one of those disks or servers should fail. However, when end users try to access data in one of the failed servers, they simply cannot. Therefore a load balancer is required to direct user traffic to an alternative server.

Leading object storage vendors, including Cloudian, OpenIO, Scality, Dell EMC and NetApp, all recommend the use of a load balancer in conjunction with their solutions. Correctly installed and sized for anticipated data volumes and usage, load balancers deliver two vitally important benefits: zero downtime and optimum application performance.

Zero downtime

Given the criticality of storage systems, zero downtime will be an aspiration – if not a critical business requirement – for all organizations. Within individual data centers, load balancers will perform constant health checks on storage clusters to make sure that there are no problems relating to overloading servers or disk failures, for example. If they detect a partial or complete failure anywhere in the object storage installation, they will seamlessly and instantly redirect user traffic to alternative servers, to prevent down time.

> *Load balancers play a pivotal role in ensuring that such disk failures have no impact on storage availability*

Organizations that adopt object storage frequently elect to install the solution across multiple data centers, in different geographic locations. In these instances, load balancers play an additional, vital role. Using Global Server DNS functionality, load balancers will make sure that if there is a failure with the storage provision in one data center, they can direct all user traffic immediately to the alternate data center.

It is a well-known fact that there is a high rate of disk failures in object storage systems, due to the large number of physical disks that these solutions depend upon. As many as 50% of support calls to object storage vendors from end users can relate to malfunctioning disks. Load balancers play a pivotal role in ensuring that such disk failures have no impact on storage availability, by health checking all disks and diverting traffic to functioning systems. When servers need to be taken out of service temporarily, to allow for disks to be replaced, load balancers redistribute the traffic, deliver an uninterrupted service for the user.

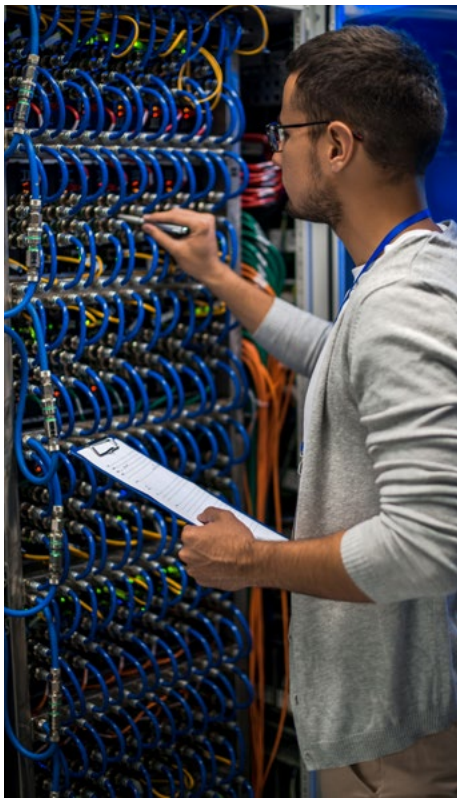
Optimum application performance

Object storage systems are increasingly being installed in environments where there are large numbers of users accessing large data files (such as video) or processing Big Data. In such situations, load balancers play a part in managing traffic to ensure that all users benefit from the same high quality experience. Object storage servers have a finite number of physical disks, which can become overloaded if too much user traffic is sent to the same place. Load balancers distribute traffic across all available object storage servers to deliver optimum storage performance for everyone.

> *Load balancers play a part in managing traffic to ensure that all users benefit from the same high quality experience*

Equally, load balancers improve the performance of object storage systems by managing the volume of traffic across the network. If, for example, a network has a 10 GB capacity, bottlenecks can occur if the traffic in one part of the network approaches this threshold. Load balancers make intelligent decisions about how to direct traffic through the network, without saturating the available network bandwidth.

The use of load balancers in object storage environments also has the knock-on effect of improving performance in those business applications that depend on the data storage infrastructure. If we take the example of healthcare, having a high performance storage solution will in-turn improve the performance of associated medical imaging solutions, making patient x-rays available more quickly.



How do object storage vendors know which load balancers to recommend?

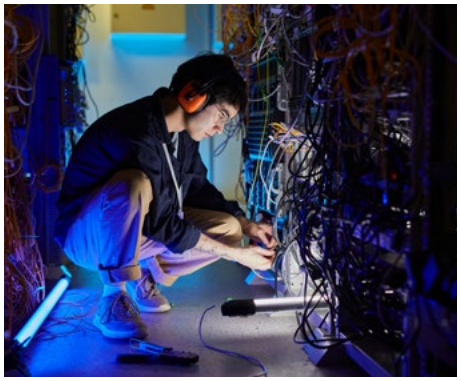
Some object storage vendors put the onus on their customers, the end users, to supply and install a load balancer alongside object storage solutions. While this approach gives customers the freedom to choose and the ultimate flexibility to design their own IT environments, it can create challenges.

For the object storage vendors, if every customer selects a different load balancer, every deployment will also be different, resulting in storage environments that are harder to support. With a large amount of variation in deployments, compatibility issues may also arise during or after installation that object storage vendors have never encountered before and therefore cannot resolve quickly.

Equally, from the customer's perspective, if the object storage vendor does not recommend a load balancer (or a shortlist of possible load balancers), then the IT department may have to spend a lot of time researching options and liaising with vendors to make sure the chosen load balancers are appropriate, which can delay the storage migration.

In light of these challenges, several leading object storage vendors are now not only recommending one or more specific load balancing products, but also forming partnerships with load balancing vendors to enable them to offer integrated or bundled solutions, including the load balancer.

When forming such partnerships, object storage vendors should consider:



Willingness to customize features

To create a truly integrated solution for their customers, object storage vendors should partner with load balancer vendors that are prepared to invest in developing features and functionality specific to the object storage solution. Loadbalancer.org has worked with several object storage providers to create platform-specific health check functionality that looks deep into the application layer to monitor performance and make sure the object storage solution is behaving optimally.

Consultancy and pre-sales support

The cornerstone of successful customer deployments is thorough planning and detailed architecture designs. For this reason, object storage vendors don't simply need to select a load balancer; they need to select a partner who can work directly with them, providing consultancy and pre-sales support and being as hands-on or hands-off as needed.

Knowledge of industry applications

It can be enormously helpful for object storage vendors to work with load balancer vendors who understand the industries in which their solutions are used and the challenges of these application environments. Some load balancer vendors, including Loadbalancer.org, have pre-existing relationships with vendors of solutions that are commonly used in combination with object storage, such as medical imaging, electronic health records and security.

Experience with file gateways

Many end user organizations that are currently planning migrations to object storage solutions are also evaluating file gateway solutions, also known as storage gateways or file management systems. These solutions make it easy for end users to name, save and retrieve files, by rendering the object storage interface more similar to traditional hierarchical storage solutions. Given this trend, it is important for object storage vendors to work with load balancer vendors who have established partnerships with file gateway vendors and experience of setting up load balancers in such complex environments.

Full commitment to support

Good quality support is essential, and object storage vendors need to be confident that their chosen load balancer partners will work with them to ensure that their customers have the best possible experience. Loadbalancer.org creates support frameworks and partakes in mutual training, to ensure that its own technical staff and technical staff at the object storage vendor are trained in all aspects of the whole solution. In this way, Loadbalancer.org becomes an extension of its partners' teams, providing a feedback loop to continually improve the delivery of support for customers.

> *It is important to look for solutions that have a proven track record of supporting the specific object storage platform selected*

How can organizations identify the best load balancers to use?

In theory, all load balancers are compatible with object storage environments. In practice, however, some load balancers will provide a more comfortable fit than others. If object storage vendors do not provide load balancers bundled with their solutions, then end user organizations will need to shortlist, evaluate and select load balancers to meet their own requirements. This can be a time-consuming process, but it is vitally important to undertake it thoroughly, as the wrong load balancer could destabilize the new object storage system.

Sometimes, organizations attempt to use load balancers that are already at work within their IT infrastructure to support the new object storage platform, in addition to the existing workload. This can be a costly mistake. It is easy to underestimate the throughput and scalability requirements of object storage systems, and inadvertently place a strain on all corporate applications.

Important criteria that organizations should consider when selecting a load balancer include:

Scalability without limits

Some load balancers have restrictive licensing agreements, which impose artificial limits on throughput, and these kinds of products can become costly or even obsolete in just a few years. Organizations should therefore seek scalable products and size them to take into account the likelihood of rapid, significant data growth.

Proven track record

Object storage solutions are, for many organizations, mission critical. Consequently, it is important to look for solutions that have a proven track record of supporting the specific object storage platform selected. The best option will always be load balancers that are well integrated with the object storage product and extensively tested in this environment, so that purchasing decisions can be made with complete peace of mind.

Product-specific deployment guides

Whether a new storage solution is being implemented by the chosen storage vendor, a third party IT services partner or the organization's own IT team, clear documentation will be invaluable. Load balancer vendors who are dominant in the object storage marketplace should be able to provide detailed deployment guides, for each object storage product, giving everyone clear and straightforward guidelines on how to set up the environment so that it works – first time.

Presales consultancy

Finally, it is important for organizations to start thinking about their load balancer requirements as soon as possible and seek pre-sales consultancy from the outset. It can be difficult to size the load balancer and ensure that it integrates optimally, if the architecture has already been planned. The earlier the requirements of a load balancer are considered, the better chance the project has of being a technical success.

Holistic support

Everyone knows just how frustrating it can be, to be passed from vendor to vendor when an issue arises, with no-one taking responsibility for finding a resolution. IT technicians at Loadbalancer.org understand object storage environments, and can look either side of the load balancer to see what is happening and find solutions to problems, without bouncing the support call on to other vendors.

About Loadbalancer.org

Loadbalancer.org's mission is to ensure that its clients' businesses are never interrupted. The load balancer experts ask the right questions to get to the heart of what matters, bringing a depth of understanding to each deployment. Experience enables Loadbalancer.org engineers to design less complex, unbreakable solutions - and to provide exceptional personalized support.

LOADBALANCER

Visit us: www.loadbalancer.org

Phone us: +44 (0)330 380 1064

Phone us: +1 833 274 2566

Email us: info@loadbalancer.org

Follow us: @loadbalancer.org