LOADBALANCER

WHITE PAPER

Load balancers: the lifeblood in resilient Object Storage

Performance is key when it comes to object storage. In this paper, we explore the pivotal role of a load balancer in a high-performing object storage system.

Data is big and getting bigger every day. According to an IDC (International Data Corporation) report, the Global Datasphere will grow from 33 ZB in 2018 to 175 ZB in 2025. Clearly, rapid advancement in digital technology is leading to unprecedented and exponential data growth. Like, what you order on Amazon, how many hours you spend at the gym, how many miles you drive every day, how often you tweet - all this data is being tracked, recorded and stored on virtual machines somewhere. While this may sound fascinating, the question is how do organizations deal with this growing amount of data? Managing their IT system has never been simple for most organizations, and today's massive data explosion is making it more complicated than ever. Adapting to a data-driven world of business and dealing with complex storage requirements is now a challenge for most organizations across industries.

> *Adapting to a data-driven world of business and dealing with complex storage requirements is now a challenge for most organizations across industries.*

Moving from traditional to object storage

Over the years, nearly all software applications were designed to read and write data based on File and Block Storage. While File Storage manages data organized into hierarchical file systems, Block Storage manages data as blocks within sectors and tracks. But these platforms fail to scale sufficiently to meet the growing storage needs, simply because conventional storage techniques just weren't designed to handle the data tsunami that the current business world is heading towards. They work well for smaller data sets but fall apart when attempted to scale. Besides, as more and more applications are now designed to talk directly to the storage nodes, without additional (file system) layers in between, the older concepts of data storage are slowly becoming obsolete.

> *Object Storage is a new storage paradigm with a very simple interface*

Therefore, as data explodes, organizations are forced to relook at their storage systems and invest in solutions that are sustainable and long-term. As a result, they are making a move from traditional storage options to cloud-based Object Storage, which is a reliable, efficient and affordable way of storing, archiving, backing up, and managing huge volumes of static or unstructured data. Designed to be massively scalable, Object Storage is a new storage paradigm with a very simple interface. It organizes information into containers of flexible sizes, referred to as objects. Each object includes the data itself and its associated metadata and has a globally unique identifier name. It is a "flat" structure and naming convention. The biggest benefit of such an architecture and interface is its ability to achieve enormous scale dynamically. Besides storing humongous amounts of data, Object Storage also allows access to large amounts of disparate data sources for analytics and advanced reporting which traditional storage fails to offer. No matter where a particular type of data is stored, Object Storage is intelligent enough to find that data whenever a related query is fed in. It also ensures improved efficiency while managing very large quantities of data, thus making it a high-performance, cost-effective solution ideal for long-term data retention.

But a poorly managed object-storage system can put an entire business at risk. Therefore, more and more organizations are looking to partner with reliable enterprise data storage vendors to take complete control of their storage systems and processes.

Role of load balancers in object storage

To ensure a robust Object Storage system for their customers, storage vendors must implement an effective load balancing solution into their architecture. Load balancing improves responsiveness and increases the availability of applications by distributing network or application traffic across a cluster of servers. Thus, it is the key to running a successful Object Storage system. While most object-based data storage vendors promise unlimited scalability as one of their biggest strengths, load balancing is the driving force behind it. It helps Object Storage scale better by adding health checks and failover. Without a load balancer, most data storage vendors would use a simple round-robin DNS solution which no doubt can provide scalability, but is not completely reliable when a storage node fails. Therefore, in order to scale their solutions adequately to meet customers' growing data demands, it is important for storage vendors to add load balancers to their storage cluster. While doing that, vendors can choose from the following options.

Layer 4 and/or Layer 7

The easiest choice is to use Layer 4 and/or Layer 7 load balancing which allows all the traffic to flow through the load balancer. It is automatic, instant and not even noticeable to the end-users. With it, most storage appliances by default undergo the same method. But vendors must ensure to use a large enough load balancing solution to scale up according to the throughput requirement.

Direct Server Return (DSR/DR) mode

Other options include using the Direct Server Return (DSR/DR) mode where network traffic bypasses the load balancer on its return path to the client-server and thus, allows endless scalability for retrieving data out of the client storage. This method suits write few and retrieve many use cases, where the write speed is less significant but faster read performance needs to be guaranteed.

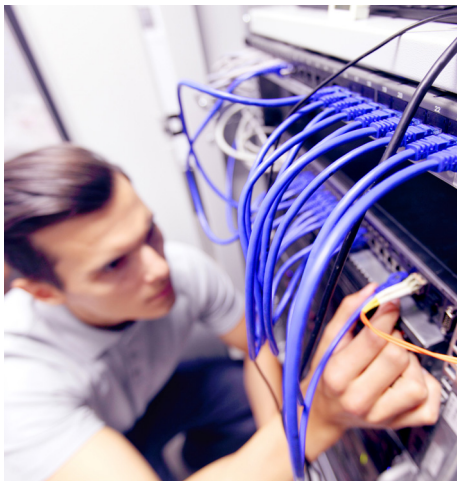
Global Server Load Balancing (GSLB)

Additionally, vendors can choose Global Server Load Balancing (GSLB), which is an upgraded version of the traditional round-robin DNS technique. GSLB adds health checks and location awareness and is the best choice for organizations looking at multi-site deployments. This solution offers the best of both worlds as vendors have quick guaranteed failover at the IP level provided by traditional Layer 4 or Layer 7 load balancing as well as location awareness and site failover provided by GSLB. Rather than using a simple round-robin DNS, storage vendors can use a GSLB to handle the entire load balancing requirement. This surely removes bottlenecks, but failover gets affected as it happens at the DNS level and is mostly reliant on the client software accessing the storage system.

> *Load balancing is the key to running a successful Object Storage system*

Load balancers facilitate zero downtime

The ability to copy data in multiple locations is another major benefit of using Object Storage. Data can easily be replicated within nodes and clusters among distributed data centers for additional back up, off-site and even across geographies. The storage system can be configured in a way that if a particular disk within a cluster fails, a duplicate disk is always available, ensuring the system continues to run without interruption or performance degradation. What makes this possible? Again, load balancing does the magic! Load balancers allow storage vendors to distribute (spread out) and store data in multiple locations to facilitate zero downtime in case of failover. For example, if data center A fails, a load balancer allows users of that locality to access the same data in data center B. Load balancers can use a myriad of Access Control Lists (ACLs), rules and topology information to direct users to the correct location to access storage. Typically, for multi-site deployments, storage vendors can use the GSLB with its topology feature which allows matching the source subnets to locations, helping users access their resources locally unless a failover occurs. When a site failover occurs, users are rerouted to a different location, thus ensuring uninterrupted data access.



Adjusting server weights and distributing traffic evenly is important for storage vendors to efficiently manage client object storage systems and prevent against complete data center failures. Load balancing facilitates this. By utilizing a “least connection” algorithm, load balancers can evenly distribute traffic among backends. They are equipped with intelligent decision-making capability which allows them to monitor the number of sessions sent to each node so that the new sessions can be scheduled to the least loaded node. Besides, they also work with Feedback Agents on real servers to additionally monitor CPU and memory loads and calculate the availability of resources left on the server. In case data center failures occur, load balancers are capable of allowing users with uninterrupted data access. How? Load balancers allow failovers to occur at the IP level and do not rely on client software to work. They fully load balance traffic without even a client typically needing to know that anything has changed. A load balancer quickly pulls back the failed backend node from the available servers and redistributes traffic amongst the remaining healthy nodes, thus resulting in a super fast failover. However, for vendors using GSLB alone, failover occurs at the DNS level where just like the IP failover, the failed location is removed from being served up to new users. The benefit of using GSLB is that it allows faster and maximum throughput although failover is slower as it is not handled at the IP level.

> *Load balancers allow storage vendors to distribute (spread out) and store data in multiple locations to facilitate zero downtime in case of failover.*

Load balancers – key architectural component in storage infrastructure

Overall, a load balancer helps Object Storage architecture maintain system firmness, improve system performance and protect against system failures. To leverage the benefits of load balancing, it is important that storage applications have load balancers installed as a key architectural component in their infrastructure. Orchestrating a load balancing solution into the storage infrastructure eliminates downtime, and ensures increased scalability, flexibility, redundancy and data protection. Data storage vendors should partner with service providers offering proven, easy-to-install result-driven load balancing solutions that are tightly integrated as part of their data storage infrastructure, running concurrently in the same environment as the application resources. Besides, installing standardized load balancing solutions across client sites helps vendors achieve consistent results and assured outcomes with every installation. Therefore, they should ideally look for service providers that understand their object storage requirements, and work together with them to design flexible load balancing solutions for specific storage needs.

Object storage – the perfect solution for data governance?

> *Load balancers eliminate downtime, and ensure increased scalability, flexibility, redundancy and data protection*

Clearly, data is every company's most strategic asset. As per Gartner, transforming into a digital business is the number one priority of most organizations now, and a Forbes article published last year suggested that data is the fuel for digital transformation. Therefore, how organizations manage, access, secure and utilize their data is critical to their business operation. It directly affects their business competitiveness and the ability to innovate. Hence, IT teams constantly look for flexible, scalable and easily manageable ways to preserve and protect the ever-growing data generated within an organization. And that is exactly where object storage shines.

Most business leaders today see Object Storage as the right technology solution for data governance. Companies that were born on the web already use it and others have started to join the bandwagon. Print, media, medical, law enforcement and many other industries are already heading towards object-based storage systems to manage their mind-boggling data banks. However, having a shiny new storage solution without accounting for proper HA (high availability) is practically of no use. While object storage typically scales endlessly on many systems, without a load balancer it is not reliable enough to be used in mission-critical systems. Object Storage needs load balancing to facilitate optimal performance. Storage vendors can no longer ignore the fact that deploying a powerful load balancing solution makes a significant difference in data storage systems. Load balancers ensure maximum application throughput, reliability, and guaranteed uptime – all things that make an Object Storage system truly unbreakable.

Source links

https://www.ibm.com/cloud/learn/object-storage

https://www.netapp.com/us/info/what-is-object-storage.aspx

https://www.gartner.com/smarterwithgartner/build-a-data-driven-organization/

https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

https://blog.seagate.com/business/enormous-growth-in-data-is-coming-how-to-prepare-for-it-and-prosper-from-it/

https://www.forbes.com/sites/cognitiveworld/2019/02/06/data-the-fuel-powering-ai-digital-transformation/#3dd60ade578b

https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/workforce/Seagate-WP-DataAge2025-March-2017.pdf

https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#4afcdce860ba

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