LOADBALANCER

WHITE PAPER

Load balancing deployment options for object storage

Optimizing high-performance object storage environments with the right application delivery controller (ADC) deployment method for your use case

Contents

1. Executive overview 2

2. Load balancing storage nodes (within a data center) 3-7

- Layer 4 Direct Routing
- Layer 7 Source Network Address Translation (SNAT)
- SDNS / GSLB Direct-to-Node

3. Load balancing storage clusters (across multiple data centers) 8-9

- Global Server Load Balancing (GSLB)
  - Weighted round robin
  - Topology weighting

4. Conclusion 10

Executive overview

Application delivery controllers (ADCs), commonly known as load balancers, are essential components of object storage solutions. Within data centers they direct traffic from clients (users' computers) to the storage node (physical or virtual server) that is best able to respond and meet the request for data.

Object storage clusters typically contain many nodes, and it is the load balancer that will spread the traffic equally across these resources to ensure that no single node is overloaded. In this way, the load balancer will improve performance and ensure faster responses. If a node fails, the load balancer will automatically and instantly direct traffic to an alternative node in the cluster, providing resilience and preventing unnecessary downtime.

Many object storage providers have two or more clusters of nodes, situated in separate data centers. These organizations will additionally need load balancers to share traffic evenly across the two clusters (if both are active), or to divert traffic instantly to a back-up cluster (if one is active and one is passive).

There are three primary methods of deploying load balancers within data centers. One of these is the SDNS / Global Server Load Balancing (GSLB) Direct-to-Node approach which is rising in popularity with object storage providers. It can work very effectively for organizations with consistent levels of requests and high throughput. However, it is not necessarily the best option in every situation.

Therefore Layer 4 Direct Routing (DR) can be a better approach for object storage providers with read-heavy apps, while Layer 7 SNAT offers flexibility for the vast majority of object storage providers, particularly those that want to be able to control traffic intelligently.

Organizations that need to balance traffic across multiple storage clusters, in multiple data centers, can use standard GSLB functionality. They can then fine tune how the load balancers operate using a variety of deployment approaches. Using weighted round robin or topology weighting, for example, organizations can control how traffic is shared throughout their object storage environments to deliver the best performance for users.

Given the range of options available, object storage providers need to take a considered approach and carefully select the load balancing deployment method that is best suited to their unique traffic volumes, packet types and customer requirements.

Load balancing storage nodes (within a single data center)

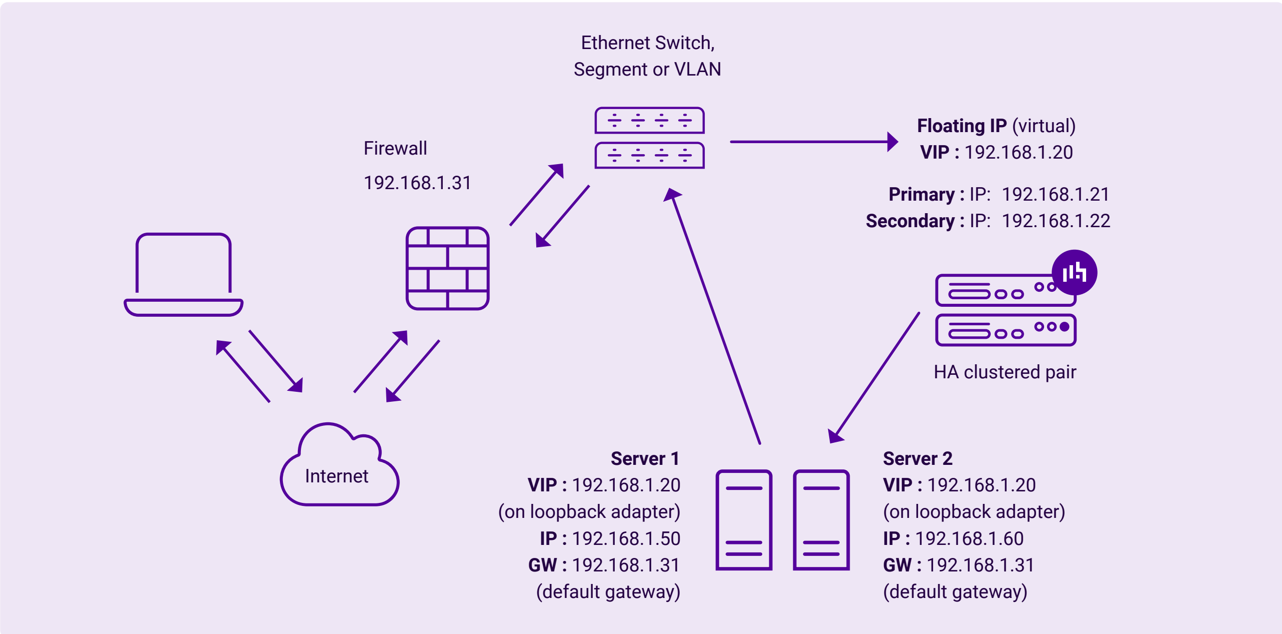
There are three main approaches to using load balancers within data centers to balance traffic across object storage nodes. No single deployment method is best for all storage environments, so object storage providers will need to carefully weigh up the pros and cons of each method.

1. Layer 4 Direct Routing (DR) mode

Layer 4 Direct Routing (DR) mode is a deployment approach that enables network traffic to pass through the load balancer in one direction only.

Requests are sent to the VIP that is on the load balancer. The load balancer then processes the request by selecting an appropriate Real Server and simply switching out the destination MAC address to that of the real server. The request is then forwarded on to the real server. Switching out the destination MAC is a simple process, is computationally inexpensive, and is therefore fast. When the server responds, however, it does so directly to the gateway, bypassing the load balancer completely.

Figure 1: Layer 4 DR mode



The main benefit of this deployment method in object storage environments is that it significantly reduces throughput through the load balancer. As traffic only traverses the load balancer in one direction, this approach also minimizes the likelihood of the load balancer becoming a bottleneck.

Layer 4 DR is, therefore, particularly beneficial for environments that only handle backups or stream large media files, as this deployment approach prevents large volumes of data from going back through the load balancer unnecessarily. If an organization is restoring a backup, the request for the data could be ten kilobytes, but the response

BENEFITS OF LAYER 4 DR MODE

- Traffic routed directly to the backend server
- Accelerates delivery of large backup and media files
- Reduces need for high-capacity load balancers

could be many terabytes. Equally, a request for a media file will be tiny, but a high-definition movie download could be four gigabytes.

Another key advantage is cost. Organizations that adopt Layer 4 DR can often specify and deploy smaller load balancers as they don't need to have the added capacity to be able to handle the return traffic, including backups and large media files.

Using Layer 4 DR to bypass the load balancer on return legs will lead to a small improvement in speed. The acceleration is likely to be measured in milliseconds, not seconds, though, so although this is an advantage, speed will generally not be the main reason for adopting Layer 4 DR.

A key disadvantage of this deployment method is that it comes with a heavy IT administration load. Nodes have to be manually configured and IP address changes have to be made on each individual node on the real servers. For this reason alone, Layer 4 DR is not used widely in object storage environments.

In the Layer 4 DR deployment approach, load balancers have the same virtual IP address as the real servers that they direct traffic to — and this can lead to a situation known as the address resolution protocol (ARP) problem. This difficulty arises because, at Layer 2, devices will send broadcasts asking, 'who has this IP address?'. If the load balancer and the real servers all have the same IP address (as they will in this deployment mode), they will all respond. Real servers therefore need to be configured to have the same virtual IP address as the load balancer, but not respond to ARP requests for it. Only the load balancer is allowed to respond.

LIMITATIONS OF LAYER 4 DR MODE

- Is more complex to manage and administer
- Causes address resolution protocol problems if not set up properly
- Cannot be used to optimize routing intelligently

If organizations buy pre-configured load balancers, the ARP problem will not be resolved as part of the standard set-up script, and they will need to implement a manual configuration to address it. This isn't difficult. Load balancer vendors provide scripts and tutorials showing IT teams what to do. However, most organizations either don't know about the ARP problem, or they don't realize that additional set-up steps are needed, and this can delay deployments.

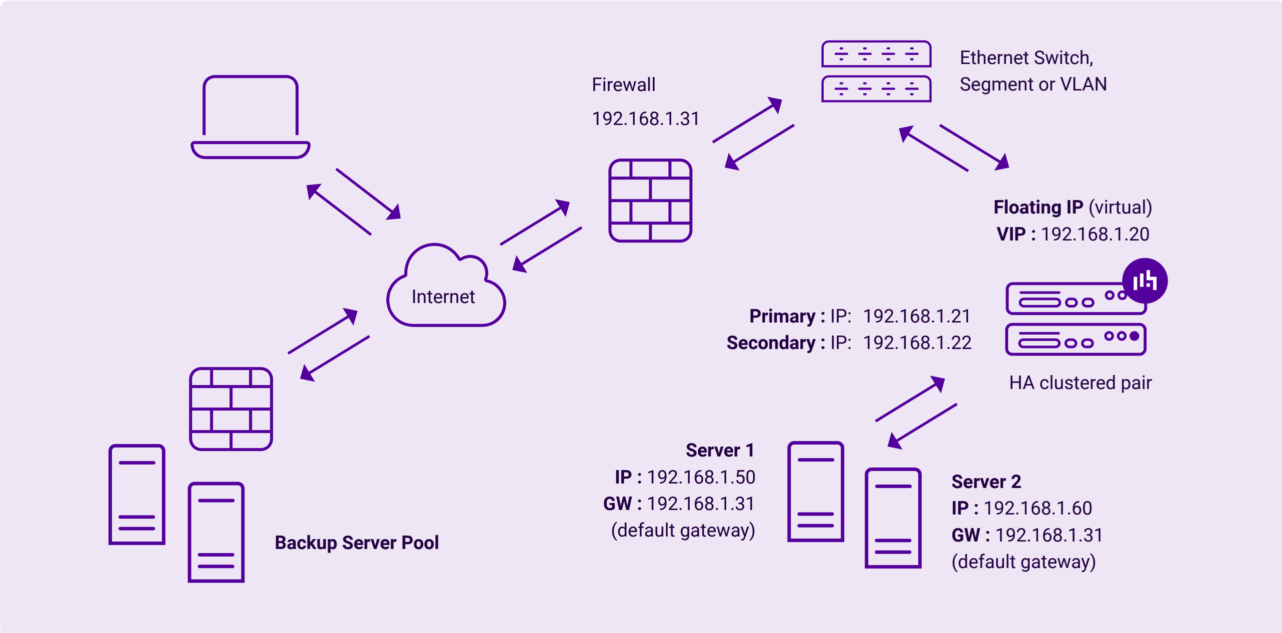
Another drawback of Layer 4 DR is that it operates, obviously, at Layer 4 and below, so only uses ports and IP addresses to make load balancing decisions. There is only one TCP connection through the load balancer, instead of two, which are necessary for configurations using a proxy server. Consequently, organizations that use Layer 4 DR cannot do any content switching or packet manipulation or investigate paths to route requests optimally.

2. Layer 7 Source Network Address Translation (SNAT) mode

In contrast to Layer 4 DR, Layer 7 Source Network Address Translation (SNAT) mode is a two-way street, with the entire packet traveling via the load balancer, in both directions.

Layer 7 SNAT mode uses a proxy (such as HAProxy) at the application layer. Inbound requests are terminated on the load balancer, and the proxy generates a new corresponding request that is forwarded on to the chosen real server. The response is then sent back along the same pathway, via the load balancer to the recipient.

Figure 2: Layer 7 SNAT mode



The significant advantage of Layer 7 SNAT is that it enables more intelligent decisions to be made about how packets should be routed within the network. It offers enhanced options such as HTTP cookie insertion, remote desktop protocol (RDP) cookies and connection broker integration and enables content switching and header manipulation rules to be implemented. Layer 7 SNAT can also be used to configure SSL termination and offloading to relieve the pressure on servers and improve application security.

Critically, Layer 7 SNAT enables organizations to persistently route traffic from one client to the same real server within specific timescales. This 'persistence' can help to improve reliability, speed up requests and enhance the customer experience.

BENEFITS OF LAYER 7 SNAT MODE

Improves the user experience with session persistence

- Provides greater flexibility to optimize package routing
- Works effectively in microservice architectures

Layer 7 SNAT is particularly effective in microservice architectures, as it allows bespoke content switching rules to be implemented for sending particular requests to specific end points. If all requests come into one place, they can be decoupled, sorted and routed to specific containers or smaller servers each handling a particular job.

In general, Layer 7 SNAT is easier to manage and administer, as IT teams do not have to configure and reconfigure real servers and manage IP addresses on end points. Only the load balancer needs to be configured, because the load balancer acts as a proxy for all traffic.

The trade-offs are speed and throughput. As the load balancer is handling traffic in both directions, throughput will be greater, and more network connectivity will be required to achieve the same read speeds. Since Layer 7 SNAT operates as a full proxy, it cannot perform as fast as Layer 4 DR mode. Computationally, however, the difference is negligible in most object storage environments.

LIMITATIONS OF LAYER 7 SNAT MODE

- Results in slightly slower performance
- Requires additional configuration to create transparency
- Demands more network connectivity

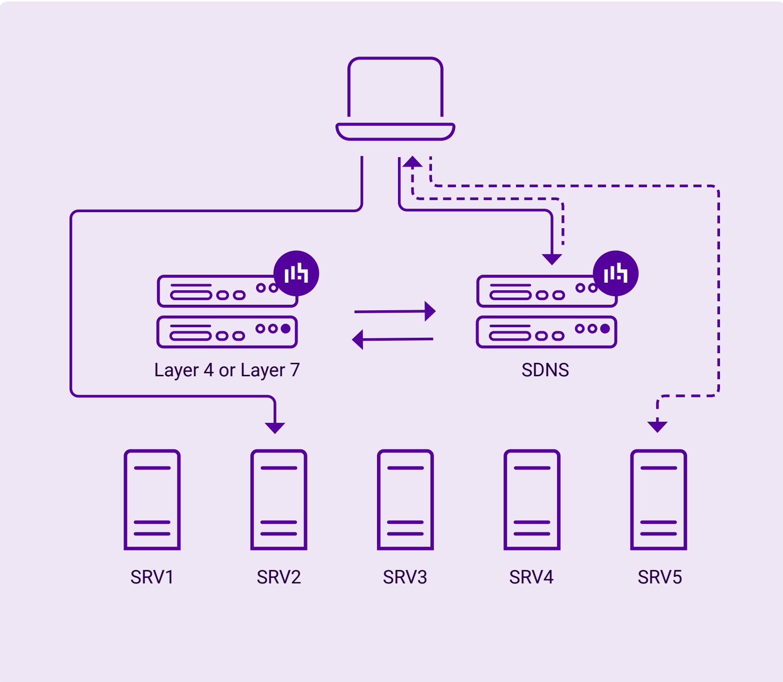
If transparency is a consideration, it is worth noting that in Layer 7 SNAT, real servers cannot see the actual source IP address of the client by default. This can be overcome in a few ways, if necessary. Such as inserting the X-Forwarded-For header, for example. In Layer 4 DR, in comparison, transparency is 'standard,' as real servers will automatically see the source IP address of the client.

3. SDNS / GSLB Direct-to-Node

Global Server Load Balancing (GSLB) is typically used to share traffic equally across two or more data centers that are geographically separate. However, it can be used in a non-standard way within a data center to enable traffic to travel to and from storage nodes, bypassing the load balancer completely. This new deployment methodology is known as GSLB Direct-to-Node, or SDNS for short.

The client will query its configured DNS servers for an IP address to send the request to. The DNS infrastructure will, in turn, delegate requests for the configured hostname to the load balancers. The load balancer decides which storage node (a virtual or physical server) is available and best able to process that request quickly. Then, the load balancer notifies the client of this decision, and the client responds by sending its packet directly to the allocated node. The server replies, delivering the requested data directly back from the node to the client, bypassing the load balancer for this return leg of the process too.

Figure 3: SDNS / GSLB Direct-to-Node



With SDNS, load balancers can be set up to perform dynamic weighting. This is the ability to change how requests to nodes in the cluster are distributed based on health checks or API calls to the cluster. If a node in a cluster is running out of storage, the load balancer can give it a lower weighting so fewer requests go to it, which improves the overall performance of the object storage solution.

SDNS is ideal for object storage providers with really large clusters of twenty or more nodes. As the load balancers only handle DNS traffic, and are not in path in either direction, they cannot create a bottleneck if traffic volumes are exceptionally high.

If an object storage provider has twenty storage nodes capable of connecting at ten gigabytes per second, for example, this is theoretically too much for a load balancer with a throughput capacity of 100 gigabits per second. In reality, it is almost unheard of for organizations to experience sustained traffic at this level, but with data levels increasing, these huge workloads are coming. SDNS provides a way of resolving this scalability challenge.

BENEFITS OF SDNS / GSLB DIRECT-TO-NODE

- Provides scalability in clusters with many nodes
- Improves performance through dynamic weighting
- Prevents load balancer bottlenecks

As with Layer 4 DR, hardware and throughput requirements will be lower, enabling object storage providers to reduce the specification and cost of load balancing. Throughput is not limited by the load balancer, but it will still be limited by the capacity of the network.

There are, however, some significant disadvantages to SDNS. If object storage providers previously used Layer 7 SNAT and then decided subsequently to move to SDNS, they would lose the ability for their load balancers to make intelligent routing decisions and use content switching.

Object storage providers would also lose the ability to achieve and manage persistence effectively. It is possible to emulate persistence by increasing DNS time-to-live (TTL) settings and extending the time that a cache is held. This approach will, however, direct traffic to the same node all the time, so is more likely to create hot nodes, with some nodes carrying a heavier load than others over time.

Another challenge results from client-side DNS caching. Clients can remember the DNS addresses previously provided by the load balancer and repeatedly direct requests to the same location, without following the load balancer's instruction. If organizations typically have high volumes of requests from a single client, in a short period of time—sometimes known as 'bursty' workloads—all these requests could go to just four nodes in a cluster of more than twenty. These nodes would become overwhelmed, affecting their performance, while other nodes would be underutilized. For this reason, SDNS is only really suitable for object storage providers with consistent workloads, where the number of requests is very steady.

LIMITATIONS OF SDNS / GSLB DIRECT-TO-NODE

- Removes ability to make intelligent routing decisions
- Hampers the use of persistence
- Causes hot nodes for workload flares

Load balancing storage clusters (across multiple data centers)

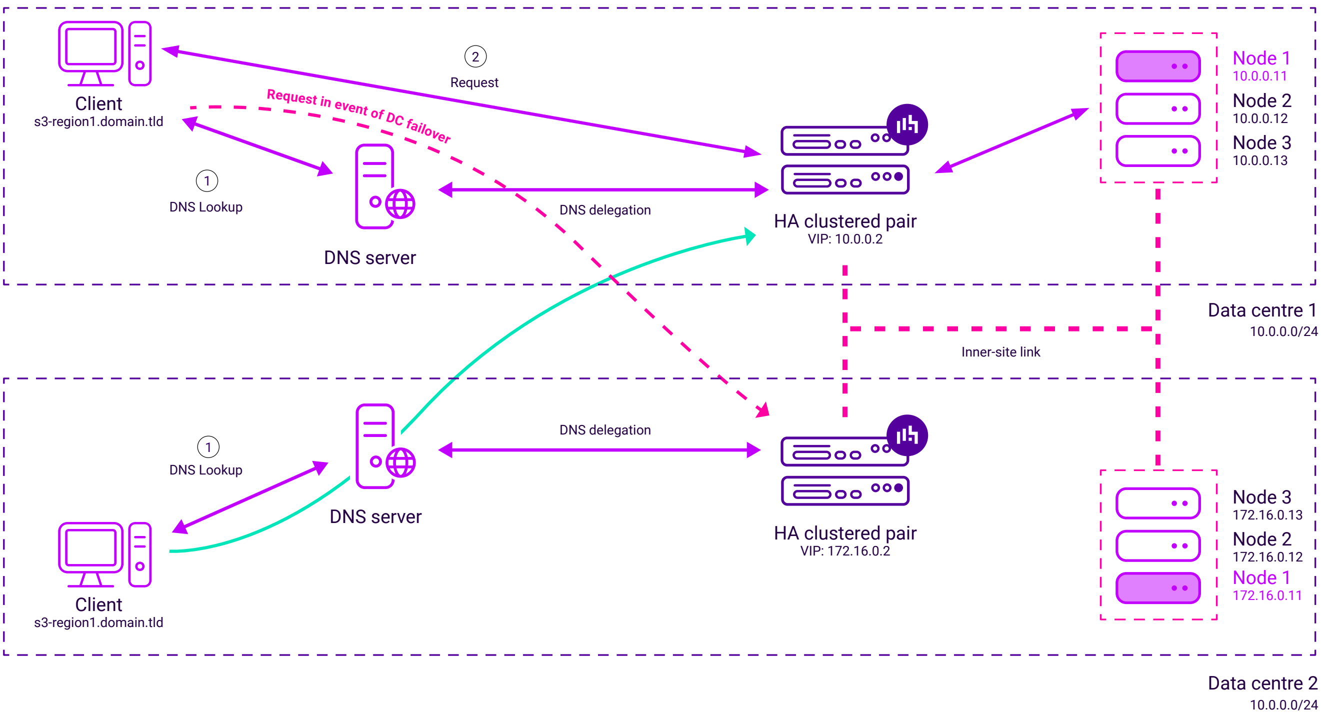
Many object storage providers have storage clusters located at multiple data centers. Consequently, they not only need load balancers to balance the traffic within their data centers; they also need load balancers to balance the traffic across data centers at different locations.

Global Server Load Balancing

Regardless of which load balancing approach object storage providers adopt within their data centers — whether they use Layer 4 DR, Layer 7 SNAT or SDNS — a standard GSLB configuration can be used to share traffic across geographically dispersed sites.

GSLB is commonly used by object storage providers that replicate their stored data across multiple sites and want to allow their users to be able to access data from any object storage cluster at any of their sites.

Figure 4. A regular Layer 7 with GSLB deployment (example FOgroup implementation shown)



GSLB receives all incoming traffic and then decides which data center to route the individual packets to. In the data centers themselves, the packets are received by the local load balancers, which then use Layer 4 DR, Layer 7 SNAT or SDNS to decide which storage nodes will serve up the data requested.

GSLB can also be configured in a variety of other ways including:

Weighted round robin

In a standard round robin approach, a load balancer configured with GSLB will allocate traffic in a circular, sequential way to all data centers, one after the other, so each site will receive the same number of requests. For example, a request from a client in London could be routed by a load balancer using GSLB to a Southampton, Newcastle or Reading data center.

Challenges can, however, arise if one data center has more or less capacity than another. In such situations, one data center can become overwhelmed, while another is underutilized.

The weighted round robin method can be used to assign weights to each data center based on the number of available physical servers or traffic-handling capacity. Traffic is then distributed to the data centers in a more intelligent way, based on each data center's ability to respond effectively.

Topology weighted

When typology weighting is used, the load balancer will consider the source of the request when making the decision about where to direct it. So, for example, a rule could be established to instruct the load balancer to send users from the London IP range to the Reading data center in the first instance, while users in the Portsmouth area could be directed first to the data center in Southampton.

If the Reading data center were to fail a health check, indicating that the Reading servers had either gone offline or were busy, then the load balancer would direct London IP traffic to Southampton or Newcastle instead. This is, in effect, a round robin method, but with preferential requests taken into account.

Using topology weighting to reduce the physical distance between the client and the object storage cluster in the data center will ultimately improve response times. This improvement in latency will be most noticeable when an object storage provider has data centers in different countries and topology weighting is used to ensure that a request from a user in the UK is not unnecessarily routed to a storage cluster in the USA, if another data center is geographically closer.

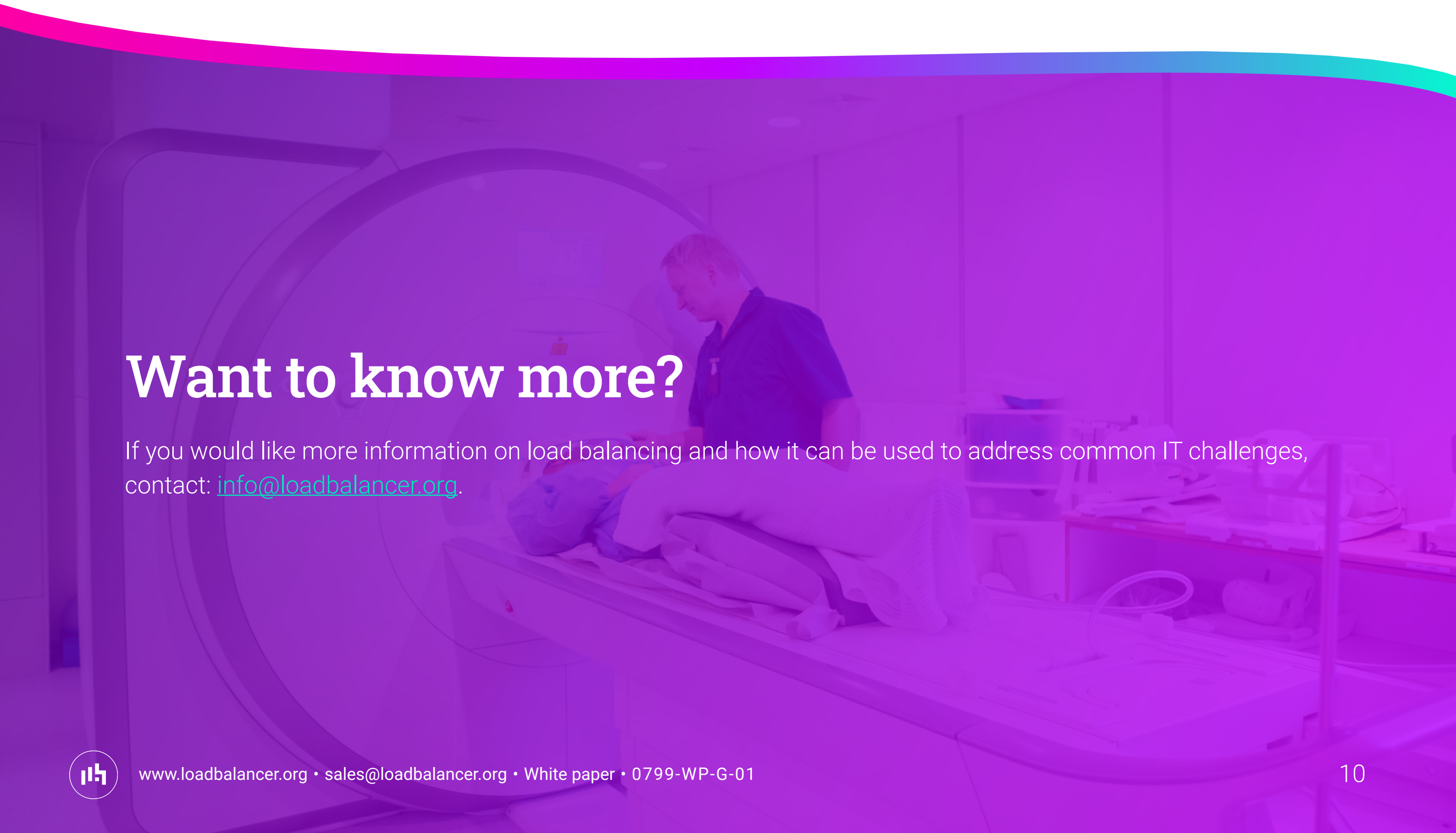
Topology weighting can also be used to overcome country-specific data governance restrictions. If, for example, a country does not permit certain types of data to be stored outside of the borders of that country, the load balancers can be configured to ensure that users are only served by a specific in-country data center.

Conclusion

Object storage providers will need to think carefully before deciding which load balancing deployment method to use. For load balancing nodes, in particular, there are three different deployment methods and they each have their pros and cons.

Most load balancing or ADC products on the market will be capable of supporting any of the deployment methodologies discussed in this whitepaper: Layer 4 DR, Layer 7 SNAT, SDNS and GSLB (with weighted round robin or topology weighting).

Getting it right can be a technical challenge for all but the most experienced engineers. Object storage providers should, therefore, leverage the expertise of a load balancing vendor that provides good support and advice, as well as detailed documentation to support seamless implementation, and step-by-step deployment instructions. The real question that arises is not which deployment method to use, but which load balancing vendor can help you optimize your object storage in this way.



Want to know more?

If you would like more information on load balancing and how it can be used to address common IT challenges, contact: info@loadbalancer.org.

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