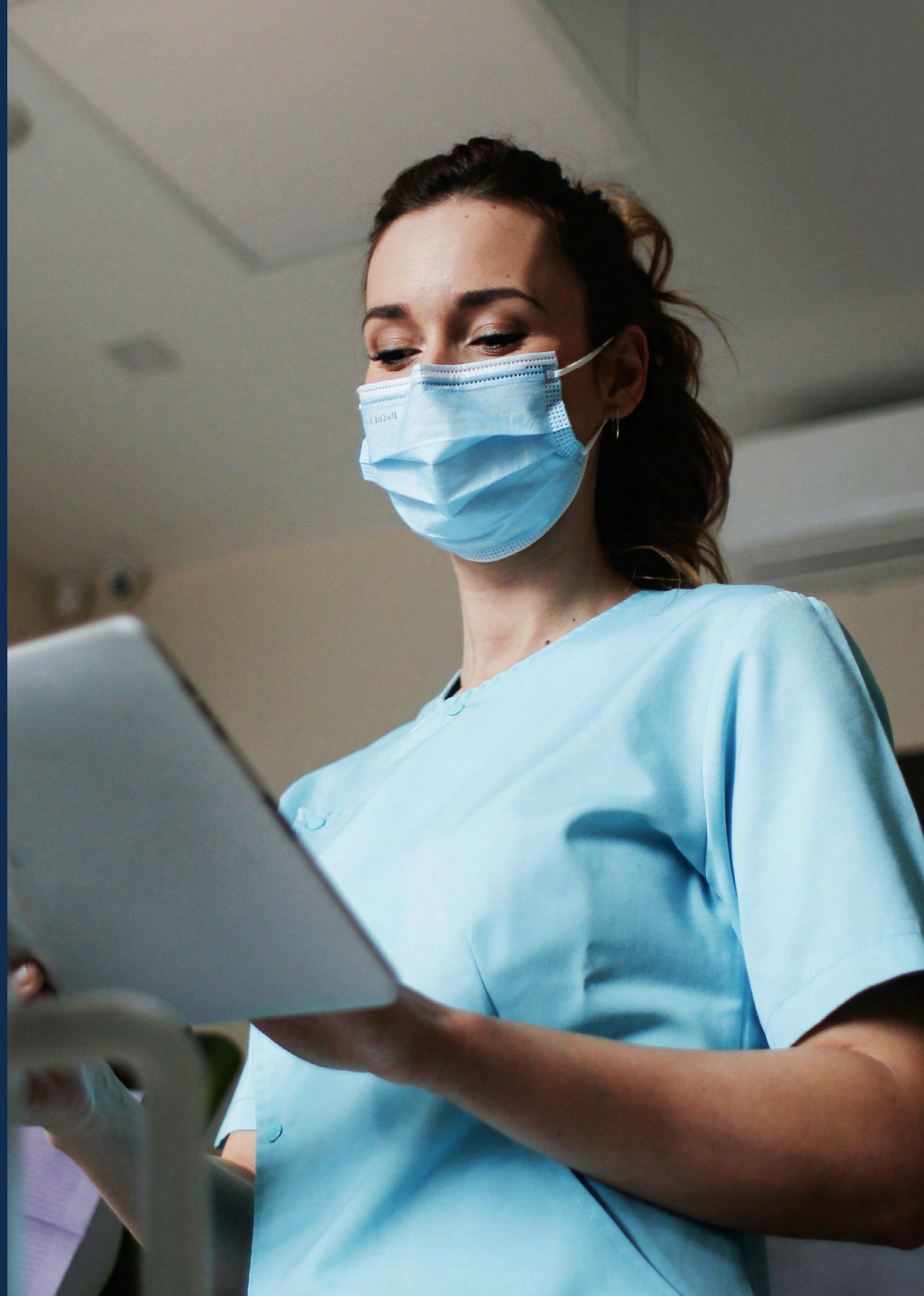
Evolving your approach to load balancing critical systems

loadbalancer.org



MULTI-APPLICATION LOAD BALANCING

Customizable, but complex

There are two primary approaches to using load balancers. The first is to use a single load balancer product or platform to balance the traffic for all applications within an organization. This method, called multi-application load balancing, gives IT teams just one load balancer technology to learn and support. The large load balancer solutions that are typically used for multi-application load balancing often come with an array of supplementary features, such as added security, and can be heavily customized to handle the traffic for different applications in different ways.

The more apps that organizations try to balance on a single load balancer platform, the more difficult the load balancer becomes to support and the more vulnerable it becomes to downtime.

In very large enterprises, these multi-featured load balancing platforms can be a good option. However, organizations need to make sure that they have the specialist IT skills in-house to manage and maintain them. As more and more applications are added to large load balancer platforms, and particularly as more customizations are made, the complexity of the application delivery environment can become overwhelming.

IT staff can struggle to complete upgrades and can inadvertently cause downtime in a mission-critical application, while performing routine updates.

The biggest challenge for organizations that adopt the multi-application load balancing approach is how to cope with the proliferation of applications. The number of apps used within medium to large enterprises has grown at an exceptional pace, as is likely to grow even more rapidly in the immediate future, as companies launch new apps to respond to COVID-19.

According to Okta's 2020 Businesses @ Work (From Home) study, the average number of apps in an organization has reached 88, up 21% from three years ago, and 10% of organizations now rely on more than 200 apps.[1] The more apps that organizations try to balance on a single load balancer platform, the more difficult the load balancer becomes to support and the more vulnerable it becomes to downtime.

1 https://www.okta.com/businesses-at-work/2020/

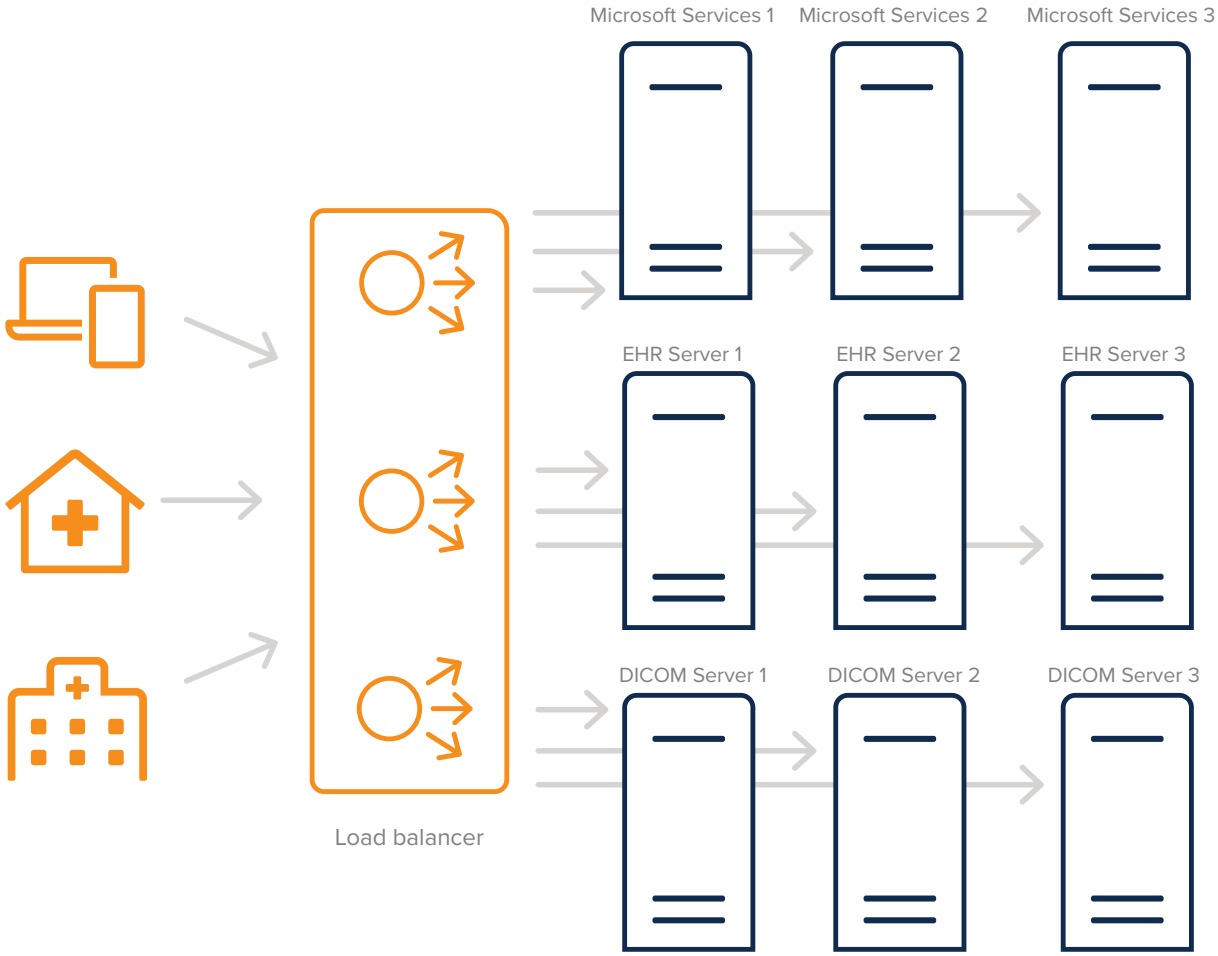
MULTI-APPLICATION LOAD BALANCING

Suitable for

- *Making complex application delivery customizations*
- *Accessing additional security features*
- *Being supported by load balancer specialists or external consultants*

Not recommended for

- *Balancing large numbers of applications*
- *Guaranteeing the uptime of mission-critical applications*
- *Being supported by IT teams without specialist skills*



Key considerations with multi-application load balancing:

WILL I USE ALL OF THE FEATURES I AM PAYING FOR?

The large enterprise load balancers, designed for multi-application load balancing, come with a host of additional features that are enticing. However, according to a 2018 NGINX and O'Reilly survey[1], 90% of companies only use basic load balancing, despite paying a premium cost for solutions with supplementary features.

CAN I MANAGE AND MAINTAIN A HIGH END LOAD BALANCER IN-HOUSE?

The more applications organizations put on a single load balancer, the more complicated they become to manage and maintain. Even simple upgrades that should take hours can take many months to plan and implement and will require expert skills. In an evaluation of the pros and cons of a single or dual-vendor approach for servers, Gartner Group points out that not relying on just a single vendor can mitigate risks[2] and the same can apply for load balancers.

ARE THERE ANY HIDDEN COSTS?

When considering the multi-application load balancing approach, it is important to take into account the additional and often hidden costs associated with managing and supporting a high end enterprise load balancer platform. It can cost twice as much as a new load balancer to bring an enterprise load balancer back into support, and organizations may have to hire consultants to help them implement or amend customizations, if they don't have the skills in-house. Centralized control can potentially increase security, but the result is slow deployments and difficulty making changes in the future.

DO I HAVE THE SKILLS TO MITIGATE SECURITY RISKS?

The customization features offered by large enterprise load balancers can be very useful, enabling organizations to specify exactly how traffic is directed to meet a business requirement. However, over reliance on customization can create security weaknesses, such as the one identified in F5's iRules in August 2019.[3] Similarly Citrix had to address a "critical flaw" in its application delivery controllers in January 2020.[4] As all multi-application load balancer implementations are unique, organizations will need to be able to detect and mitigate any security risks themselves.

HAVE I ACCOUNTED FOR PEAKS IN TRAFFIC?

It is important to understand the potential peaks in traffic for each individual application supported in a multi-application load balancing environment. If resources are not allocated correctly for each app, a peak in traffic in one app could adversely affect the performance of several others or lead to the failure of a critical business system.

WHAT ARE THE PROCESSING POWER REQUIREMENTS?

In a recent test performed by a storage vendor, a standard load balancer product only required 1 CPU and 2GB RAM, while the more complex enterprise load balancer required more than 4 CPU and 16GB of RAM to perform the same function and remain stable. The extra resource required for the large load balancer platform added significantly to the overall cost of the solution, whilst providing no extra benefit for the customer, as the extra features included would never be used.

WILL I BE ABLE TO RESIST THE URGE TO SWEAT THE ASSET?

When you have purchased your new enterprise load balancer, and sized it appropriately with capacity for unforeseeable, exceptional peaks in application usage, you need to have the discipline to keep this capacity ring fenced. Post procurement, organizations commonly make the mistake of looking at their usage data, thinking they have excess capacity and adding more applications to sweat the asset. This can overload the platform and lead to application downtime when peaks occur.

1 https://www.nginx.com/resources/library/survey-2018-new-infrastructure-for-a-multi-cloud-world/
2 https://www.gartner.com/en/documents/2503843/when-two-server-vendors-are-better-than-one
3 https://www.finextra.com/pressarticle/79455/serious-security-issue-in-f5s-big-ip-could-lead-to-cyber-breaches-en-masse
4 https://www.forbes.com/sites/kateoflahertyuk/2020/01/20/critical-citrix-fix-now-availableheres-how-to-apply-it/

PER-APPLICATION LOAD BALANCING

Flexible, but resilient

The alternative approach to multi-application load balancing is per-application load balancing. This involves using a dedicated load balancer (or more commonly a pair of dedicated load balancers) for each application. This approach is particularly recommended for mission-critical applications, as a pair of load balancers can be tuned specifically to ensure the high availability of the priority application, without having to take the needs of other applications into account. Far easier to set up, manage and maintain, they are also less likely to be compromised by human error.

The reduced cost of hardware, together with the availability of virtual and cloud-based load balancers, makes the per-application approach very viable.

In the past, it was cost prohibitive to have separate load balancers for each application. Now, the reduced cost of hardware, together with the availability of virtual and cloud-based load balancers, makes the per-application approach very viable, particularly when the reduced costs of support and management are taken into account. If it is company policy to standardize on single vendor solutions, to simplify the IT estate, then organizations can buy multiple load balancers from the same vendor and configure a pair to support each critical application. Furthermore, as it is much easier to configure load balancers when you only have to take into account the needs of a single application, this approach enables new apps to be launched much more quickly.

Load balancers installed for single applications don't generally come with as many supplementary features as the much more expensive load balancer platforms commonly used for multi-application load balancing. This, however, does not mean that their core load balancing capabilities are inferior. Easy-to-use load balancers still come with advanced, intelligent load balancing capabilities and are capable of routing and rerouting traffic in a fraction of a second to deliver exceptional user experiences.

Strictly speaking, the per-application approach involves allocating one pair of load balancers to one mission-critical application, but standard load balancers do have the sophisticated functionality needed to guarantee high availability for two or more applications simultaneously. Therefore, the definition of the per-application approach can be widened to encompass two mission-critical applications or a small cluster of similar applications.

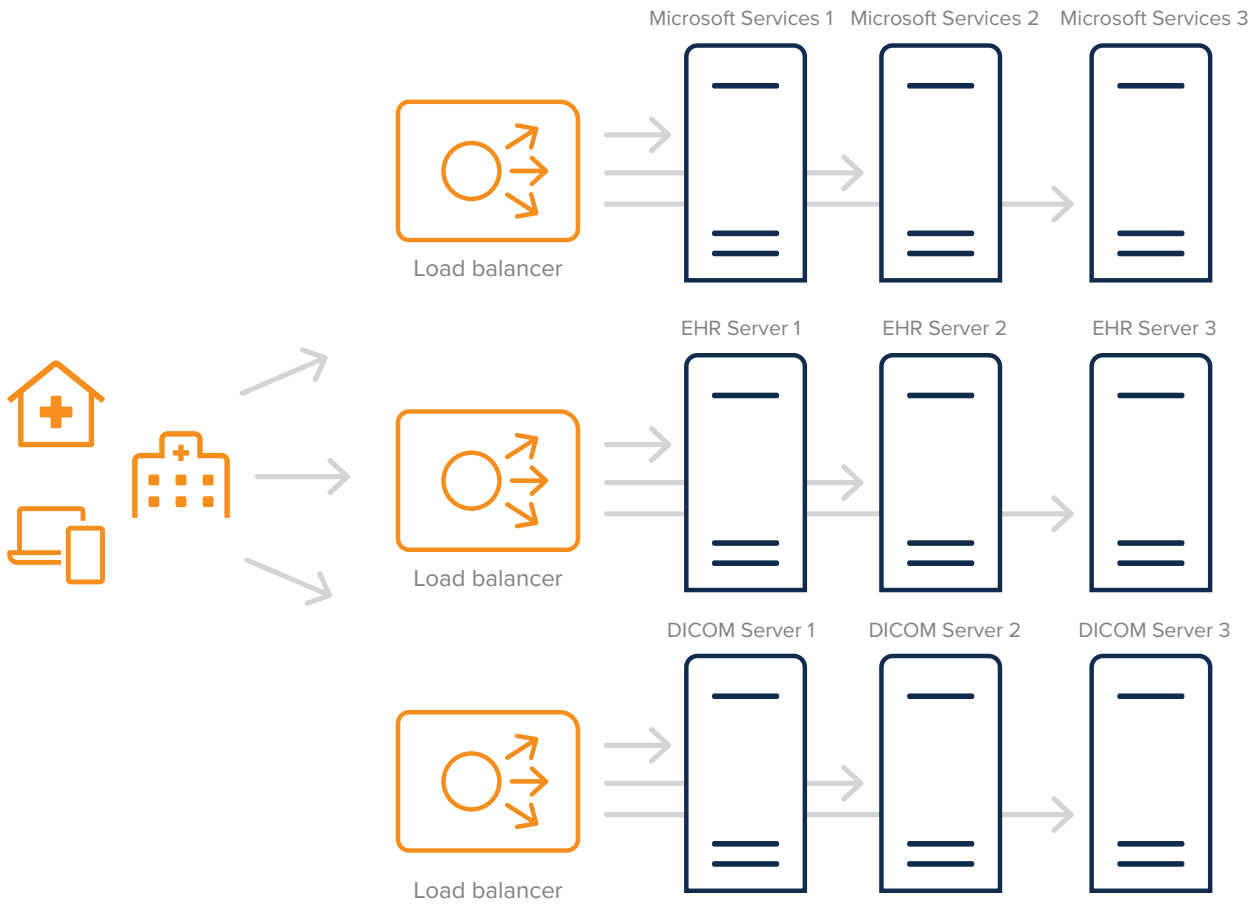
PER-APPLICATION LOAD BALANCING

Suitable for

- *Ensuring the availability of mission-critical applications*
- *Applying intelligent load balancing at lower cost*
- *Deploying new apps quickly*
- *Being easily supported by in-house IT teams*

Not recommended for

- *Accessing supplementary functionality beyond load balancing*
- *Having centralized control of all load balancing from a single dashboard*



Key considerations with per-application load balancing:

DO I NEED TO GUARANTEE UPTIME FOR MISSION CRITICAL APPLICATIONS?

The more critical an application is to the running of your business or the safeguarding of data, the more imperative it is for it to be protected by its own dedicated load balancer. It would be unthinkable for a vital medical diagnostic system to be taken down by mistake when a technician makes an upgrade to an email system.

WHAT DOES THE APPLICATION VENDOR RECOMMEND?

Application vendors that require the use of a load balancer with their solutions generally prefer their customers to have a dedicated load balancer product. Then, if their customers need help to address a problem or upgrade the application, they will be familiar with the set-up and can more easily make changes without putting other applications at risk.

WILL I HAVE THE AGILITY TO IMPLEMENT NEW APPS QUICKLY?

If speed of application delivery is important to you, then the per-application load balancing approach may be best. It can be significantly easier and therefore quicker to configure a load balancer specifically for a new app, than to try to adapt or customize an existing load balancer to accommodate the new solution, without jeopardizing the performance of other the other apps it already supports.

IS CENTRALIZED VISIBILITY AND CONTROL ESSENTIAL?

If it is, the multi-application approach might be best. The per-application approach is best suited to organizations that want to stay true to a decentralized approach.

WILL I GET THE THROUGHPUT I NEED?

It is essential to understand what throughput you will need, whether you opt for the multi-application or the per-application load balancing approach. To avoid future constraints and unanticipated increases in licensing costs, look for correctly sized solutions with high throughput options. Hardly any internet apps require more than 1G (unless they also stream video), while local network-area storage and object storage solutions could need 50G+.

HOW DO I MANAGE COSTS?

Standard load balancers, used for per-application load balancing, are a fraction of the cost of large enterprise load balancing platforms commonly used by large organizations to balance all the applications across the entire business. However, if you assign one pair of standard load balancers to multiple applications your costs could grow. Limit your expenditure by looking for site licenses that entitle you to deploy as many pairs of load balancers as you need for one fixed annual price.

WHAT SUPPORT WILL MY IT TEAM NEED?

Managing and maintaining per-application load balancers is much simpler than managing and maintaining a large enterprise load balancer that supports dozens or even hundreds of applications. Even so, you should check to see what level of support is available from your shortlisted load balancer vendors and make sure you can have direct access to engineers who can support you if needed.

DO I NEED A LOAD BALANCER AT ALL?

Ultimately, if a non-essential business application is available as Software-as-a-Service and your business is open to using the cloud, then this could well be a better option than using a load balancer. However, mission-critical solutions, like medical imaging, will need to stay on premise, because, despite the advent of 5G, the cloud is unreliable and is likely to remain so for the next five to ten years.

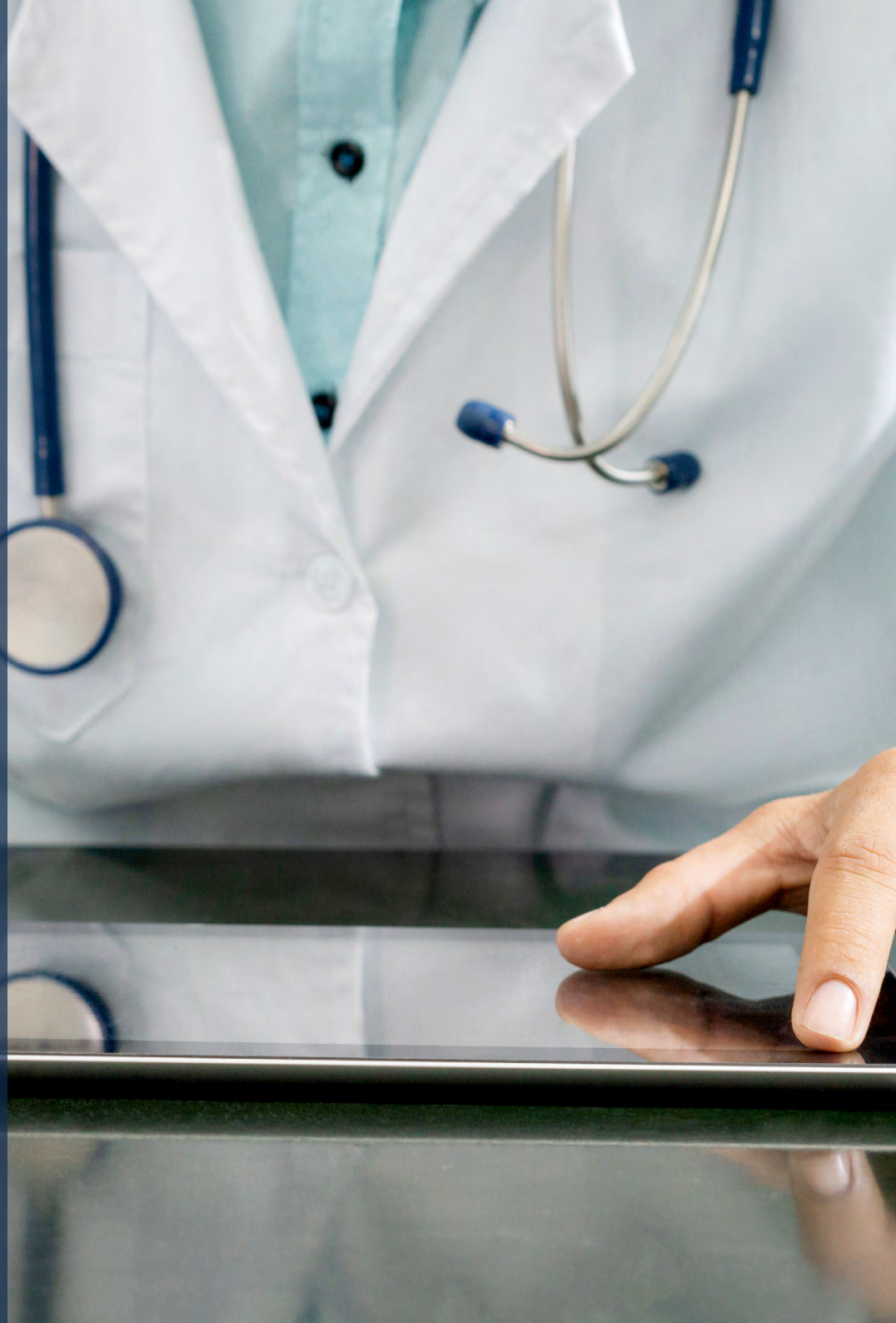
CASE STUDY
HEALTHCARE SYSTEMS

A healthcare provider specializing in delivering outpatient radiology and medical imaging systems undertakes over 72,000 scans and examinations for 55,000 patients every year in the USA. The organization relies heavily on two key systems to run its business: a Radiology Information System (RIS) and a Picture Archive and Communication System (PACS). It now uses a single, high availability pair of load balancers to manage traffic to both these mission-critical systems, at Layer 4 and Layer 7. The load balancers balance PACS application traffic across two application servers to maximize performance for users. In tandem, the same products direct RIS traffic to a production server and redirect it to a back-up server if a fault occurs or maintenance is required.

This organization is effectively using the per-application approach, but balancing two applications. The simple, straight-forward load balancer implementation is very easy for the IT team to manage, upgrade and maintain. At the same time, it improves business resilience and fulfils a critical role in optimizing the performance of the RIS and PACS applications used by 100 employees and over 800 referring doctors, as well as more than 5,000 patients per year.

Per-application benefits:

- **Delivers critical 24/7 access**
- **Reduced complexity**
- **Improved patient outcomes**



CASE STUDY
MEDICAL IMAGING

For many years, a leading supplier of medical imaging systems had advocated the per-application load balancing approach and had included load balancers as part of its diagnostic imaging solutions to ensure that medical practitioners could have uninterrupted access to patient data. However, the supplier had used a range of load balancers, from multiple vendors. This meant that different customers had slightly different installations, making it harder for the supplier to support its customers.

The medical imaging systems supplier subsequently standardized on a single load balancer vendor and worked with the load balancer vendor to create a bespoke installation guide. The per-application load balancing approach, together with the standardization on a single load balancer, enabled the supplier to accelerate the deployments of its medical imaging solutions using the same proven implementation approach for each customer. The supplier could also improve its customer service and support, as technicians have just one product and one standard network architecture to learn.

Per-application benefits:

- **Uninterrupted access to data**
- **Reduced complexity improves vendor support**
- **Standardized solution means quicker deployments**