



# International Journal of Multidisciplinary Research and Growth Evaluation



International Journal of Multidisciplinary Research and Growth Evaluation

ISSN: 2582-7138

Received: 26-02-2023; Accepted: 27-03-2023

www.allmultidisciplinaryjournal.com

Volume 4; Issue 2; March-April 2023; Page No. 605-607

## Architecting high availability solutions with google cloud load balancing

Upesh Kumar Rapolu

Houston, USA

Corresponding Author: Upesh Kumar Rapolu

DOI: <https://doi.org/10.54660/IJMRGE.2023.4.2.605-607>

### Abstract

The following research paper has provided a vivid description of the term High Availability. It has proved to be effective for cloud-based applications for determining fluent service delivery. At the same time, it has explored the integration of Google Cloud Load Balancing that plays a vital character in transforming high availability architecture. Furthermore, this has considered the essential components

that are needed for the overall enhancement of high availability such as improved performance, simplified management, enhanced security and scalability. Analysing the approaches has been rendered to control the risks leading towards a progressive road while using Google Cloud Load Balancing.

**Keywords:** Google cloud load balancing, cloud architecture, high availability, load distribution

### 1. Introduction

This research paper will provide an in-depth understanding of architecting high availability solutions with the help of Google Cloud Load Balancing. Google Cloud Load Balancing is found to be an essential tool that will be responsible for rendering to scale up the applications. This will be attained by the equal distribution of incoming traffic across multiple virtual machine instances that will help to support the heavy traffic. At the same time, it will be used to minimise the chances of performance biases by spreading the load across multiple channels. Moreover, the research paper will focus on defining the benefits of using Google Cloud Load Balancing and analysing the high availability solutions. Furthermore, this will stay abide to determine the approach that will be implemented for the overall development of architecting high availability solutions. Therefore, this will be analysed to manage the applications progressively and thus ensure a seamless user experience.

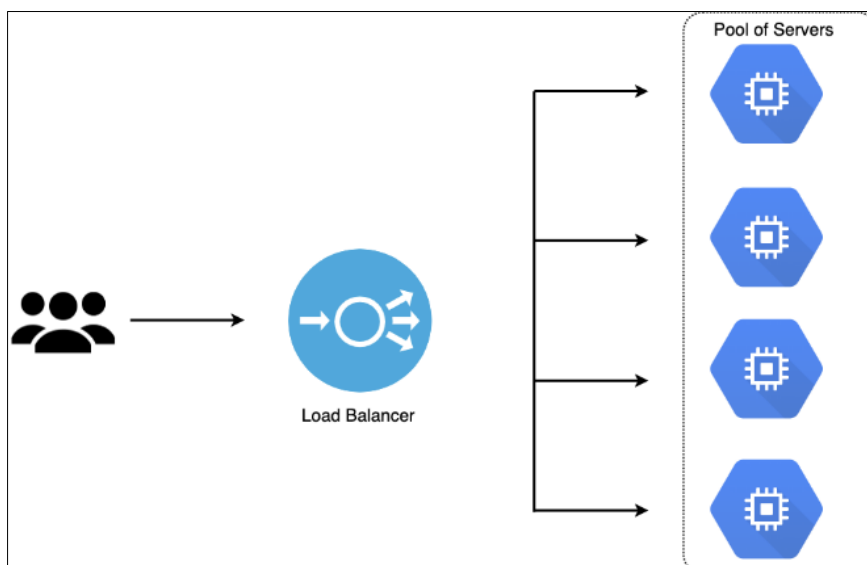


Fig 1: Google Cloud Load Balancing

## 2. Describing the overview of google cloud load balancing

The following section provides an overview of Google Cloud Load Balancing. It is observed as a necessary service which is used to manage and distribute the user traffic through several servers or via virtual machines also abbreviated as "VM". It also tends to minimise the chances of risks associated with the overloading of a host. At the same time, it is found that Google Cloud Load balancing functions by receiving the user traffic and thus distributing it to multiple servers<sup>1</sup>. However, this can have the possibility to divert or unexpected spikes in the traffic segment to increase application availability. It is considered to be effective as it mitigates latency and elevates security policies by ignoring congestion by creating robust architecture. There are different types of load balancing which are namely; Global HTTPS Load Balancing Regional Load Balancing and Network Load Balancing. Global HTTPS Load Balancing has the power to support eighth high-quality content that is distributed through various backend services<sup>2</sup>. The use of Regional Load Balancing is that it exchanges the traffic with a specific topographical locality. Moreover, Network Load Balancing tends to perform its function at the transport level by the utilisation of TCP traffic sharing.

## 3. Illustrating the benefits of using google cloud load balancing for high availability

This section illustrates the benefits of surging Google Cloud Load Balancing for High Availability in a progressive manner.

- 1. Improved Performance:** Catering with a consistency in the performance level that is featured like SSL offloading followed by coaching and intelligent traffic distributions. This has the probability of improving the overall performance and user experience. As a result, this allows Google Cloud Load Balancing to be scaled automatically<sup>3</sup>.
- 2. Simplified Management:** Fostering with simplified management allows Google Cloud Load Balancing to lower the operational burden on teams. This can be used to achieve High Availability by the involvement of systems that are made to remain functional even during unexpected outages or spikes in demand.
- 3. Enhanced Security:** It is observed that Google Cloud Load Balancing is termed to be vital as it enhances security through the integration of DDoS protection and SSL certificate management. This is used to secure the data received from the users in order to maintain optimum service continuity against combating potential threats<sup>4</sup>.
- 4. Scalability:** Google Cloud Load Balancing in terms of scalability is observed to be much more advantageous. This is due to the fact that it demands the user to fluctuate automatically by adjusting the number of backend instances that are meant to ensure constant performance during peak and off-peak time frames.

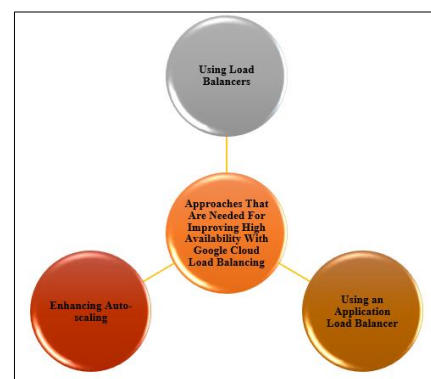
## 4. Analysing high availability solutions by using google cloud load balancing

The following section analyses high availability solutions by the utilisation of Google Cloud Load Balancing. This is achieved by allowing continuous monitoring and management followed by designing the architecture and understanding the financial conditions. The first factor is continuous monitoring and management which is required to achieve high availability. However, this is possible by the augmentation of Google Cloud Load Balancing. This helps to identify the bottlenecks and thus probably take justified actions. Secondly, in terms of designing the architecture, the Google Cloud Load Balancing needs essential elements such as backend services along with health checks and configuring the front end<sup>5</sup>. Moreover, backend services are used to construct multiple instances in diversified locations or zones making sure that latency is attained. The utilisation of health checks is used to keep a track record of backend service and redirect the traffic in the automatic mode meticulously<sup>6</sup>. Similarly, in the frontend configurations, the IP addresses are defined for the handling of incoming traffic in a mediating sense. Considering the financial status virtually provides sustainable benefits for non-critical workloads. As a result, this caters to effectively limit the possibilities of overall costs which in turn help to maintain High Availability.

## 5. Elucidating the approaches that are needed for improving high availability with google cloud load balancing

The following section elucidates the approaches which are used for the improvement of high availability with the application of Google Cloud Load Balancing. These approaches are described below.

- **Using Load Balancers:** Load balancers are used to distribute traffic across multiple servers in order to minimise the load on any one of the servers.
- **Using an Application Load Balancer:** An Application Load Balancer makes the entire application available by the usage of a single global IP address<sup>7</sup>.
- **Enhancing Auto-scaling:** The enhancement of auto-scaling is considered to be effective for the adjustments of the resources that are fully based on triggers.



**Fig 2:** Approaches for Improving High Availability by Using Google Cloud Load Balancing

## 6. Conclusion

The research paper has concluded that architecting high availability solutions with Google Cloud Load Balancing has proceeded to be effective in maintaining optimal application performance and thus satisfying the users. At the same time, the intricate use of load balancing strategies has navigated enduring effective traffic distribution followed by developing security and providing easy management through several virtual machines. Furthermore, the consistent reinforcement of high availability has rendered successful outcomes in fostering continuous monitoring with strategic architecture design and considering the financial status. This has resulted in limiting the risks and progressing to a sustainable path within the cloud-centric segment.

## Abbreviations and Acronyms

- HA- High Availability
- GCLB- Google Cloud Load Balancing
- VM- Virtual Machines
- DDoS- Distributed Denial of Service
- IP- Internet Protocol
- TCP- Transmission Control Protocol
- SSL- Secure Sockets Layer

## Units

- Computing information is measured in bytes.
- Data transfer rates and storage capacity are measured in bytes.
- Latency or response time is calculated in milliseconds

## Equations

- Load Balancing Efficiency:  $E = [L_{\text{target}} / L_{\text{actual}}]$ , where  $L_{\text{target}}$  is the ideally balanced load and  $L_{\text{actual}}$  is the actual load observed.
- Throughput =  $[ \text{Total Processed Requests} / \text{Total Time} ]$ , where this equation is used to calculate the throughput in requests per second.

## Acknowledgement

I would like to express my sincere gratitude to my professors for rendering their precious time, guidance and support throughout my research paper. Their extensive knowledge and experience proved to be instrumental for me to complete my research paper successfully within the given timeline.

## References

1. Javadpour A, *et al.* Improving load balancing for data-duplication in big data cloud computing networks. 2022.
2. Kruekaew B, Kimpan W. Multi-Objective Task Scheduling Optimization for Load Balancing in Cloud Computing Environment Using Hybrid Artificial Bee Colony Algorithm With Reinforcement Learning. IEEE. 2022. Available from: <https://ieeexplore.ieee.org>
3. Smith C, Jindal A, Chadha M, Gerndt M, Benedict S. FaDO: FaaS Functions and Data Orchestrator for Multiple Serverless Edge-Cloud Clusters. 2022.
4. Ashawa M, Douglas O, Osamor J, Jackie R. Improving cloud efficiency through optimized resource allocation technique for load balancing using LSTM machine learning algorithm. Journal of Cloud Computing. 2022; 11(1):1-14.
5. Kaur R, Luthra P. Load Balancing in Cloud Computing. 2022.
6. Bauskar S, *et al.* Data Migration in the Cloud Database:

A Review of Vendor Solutions and Challenges. SSRN Electronic Journal. 2022 Jan. Available from: <https://ssrn.com/abstract=3574927>

7. Boddapati VN, *et al.* Data Migration in the Cloud Database: A Review of Vendor Solutions and Challenges. SSRN Electronic Journal. 2022. Available from: <https://ssrn.com/abstract=3574927>