ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 6 | December 2024

www.ijirid.in

A Study on Fog Computing for Load Balancing Mechanism

Madhuri S. Aghav[1], Prof. S. A. Vyawhare[2]

[1]Research Scholar, CS&IT, Sanmati Engineering College Washim, (M.S), India

[2]Assistant Professor, CS&IT, Sanmati Engineering College Washim, (M.S), India

Abstract: *Recently, fog computing has emerged as a modern distributed paradigm and a complement to cloud computing for delivering services. Fog systems extend computing and storage capabilities to the network edge, significantly addressing the challenges of delay-sensitive applications while enabling location awareness and mobility support. Load balancing plays a crucial role in fog networks, preventing situations where some fog nodes are either overburdened or underutilized. By implementing effective load-balancing strategies, Quality of Service (QoS) parameters such as resource utilization, throughput, cost, response time, performance, and energy consumption can be enhanced. Although various studies have explored load-balancing techniques in fog networks in recent years, no systematic review has been conducted to consolidate these findings. This article systematically reviews load-balancing mechanisms in fog computing, categorizing them into approximate, exact, fundamental, and hybrid methods (covering studies published between 2013 and August 2020). Additionally, it evaluates load-balancing metrics, highlighting their advantages and disadvantages in relation to the selected mechanisms. The evaluation techniques and tools used in the reviewed studies are also examined. Furthermore, the article discusses key open challenges and future trends in load-balancing mechanisms for fog networks.*

Keywords: Fog Computing, Internet of Things, Load Balancing, Quality of Service, Systematic Review.

I. INTRODUCTION

Fog computing, an extension of cloud computing, is a geographically distributed paradigm that brings computing and networking capabilities closer to the network edge, end-users, and IoT devices through widely distributed fog nodes [1]. In traditional cloud-only architectures, most data requiring processing, analysis, and storage is transmitted to cloud servers, which can adversely affect latency, security, mobility, and reliability. For location-aware and delay-sensitive applications, the cloud alone often struggles to meet the extremely low-latency requirements. In contrast, the proximity of fog computing to IoT devices can significantly reduce latency and address these challenges effectively [2], [3].

Fog computing works in tandem with the cloud, supporting it while enabling a new generation of applications and services. In modern fog environments, users demand applications that provide quick responses and operate efficiently. To enhance QoS factors in fog networks, an effective load-balancing strategy is essential. Load balancing is particularly critical in fog environments due to the exponential increase in cloud computing loads [4].

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 6 | December 2024

Unlike cloud networks, fog networks are inherently heterogeneous and dynamic, requiring tailored load-balancing mechanisms [6]. The primary goal of load balancing in fog environments is to evenly distribute incoming loads across available fog nodes and cloud resources to prevent overload or underutilization. This approach maximizes throughput, performance, and resource utilization while minimizing response time, cost, and energy consumption [12].

II. LITERATURE SURVEY

Due to the unprecedented amount of data and the connection of over 50 billion devices to the Internet (based on Cisco estimation), handling that much of data with traditional computing models, like cloud computing, distributed computing, etc. is difficult [13]. Often privacy gaps, high communication delay, related network traffic loads that connect cloud computing to end-users for unpredictable reasons with the recent expansion of services related to IoT (like smart cities, eHealth, industrial scenarios, smart transportation systems, etc. [14]) are some challenges that affect cloud computing performance. To refer to some of cloud computing limitations and to bring cloud service traits so much closer to “Things”, as it is referred to, including cars, mobile phones, embedded systems, sensors, etc., the research community has suggested the fog computing concept [1].

Fog computing is regarded as a platform bringing cloud computing to end-users’ vicinity. “Fog”, as a term, has an analogy with real-life fog and was initially introduced by Cisco [1]. When the fog is nearer to the earth, clouds are up above in the sky and, interestingly, fog computing applies this concept, when the virtual fog platform is located closer to end users just between end-users’ devices and the cloud. In a similar definition, fog computing is suggested to make computing possible at the network edge, to send new services and applications specifically for the Internet future [15].

Bonomi, et al. [1], to give a more appropriate definition of fog computing for the first time, said that fog computing was not exclusively located at the network edge. However, it was a virtualized platform providing networking services, storage, and computations among the data centers and end devices of conventional cloud computing.

Fog computing is most often mistaken for edge computing, but we have major differences between the two. Fog computing applications are run in a multi-layer architecture that disconnects and meshes the software and hardware functions, permitting the dynamic reconfigurations for diverse applications while executing transmission services and intelligent computing. Edge computing, on the other hand, creates a direct transmission service and manages special applications in a fixed logic location. While Fog computing is hierarchical, edge computing is limited to a few peripheral devices. Besides networking and computation, fog computing deals with the control, storage, and acceleration of data-processing [16], [17]. An IoT client or smart end-device, to recognize fog computing from other computing standards, needs to utilize the following characteristics but not all of them while consuming a fog computing service [13], [18].

III. FOG ARCHITECTURE

The fog computing architecture reference model has become an essential research area. Recently, numerous architectures have been proposed for fog computing, most of which follow a three-layer

 DOI 10.5281/zenodo.14786908

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 6 | December 2024

structure [11], [16]. Fog networks extend cloud services to the network edge by introducing a fog layer between cloud and user devices. Figure 1 hierarchically illustrates the fog architecture [16], comprising three distinct layers:

3.1 Cloud Layer:

The cloud layer consists of high-performance servers and various storage devices that provide application services. It offers robust storage and computation capabilities to support the long-term storage of large datasets and extensive computational analysis. Unlike traditional cloud computing, however, not all computing and storage tasks are routed through the cloud.

3.2 Fog Layer:

Positioned at the network edge, the fog layer includes components like access points, routers, switches, and gateways that serve as fog nodes. These nodes bridge the gap between the cloud and end-user devices. End devices can easily connect to fog nodes to access services. Fog nodes are capable of computing, storing, and transmitting data received from IoT devices. The fog layer is ideal for latency-sensitive applications and real-time analysis. The connection between fog nodes and the cloud data center is facilitated by the IP core network. Fog nodes collaborate with the cloud to enhance storage and computing capabilities.

3.3 User Device Layer:

This layer is closest to the physical environment and end users, comprising IoT devices such as sensors, smartphones, smart vehicles, cards, and readers. Although some devices, like smartphones, have computing capabilities, they mainly function as smart sensing devices. These geographically distributed devices gather data about physical objects or events and transfer it to higher layers for processing and storage.

End devices connect to fog nodes using wired or wireless technologies such as 3G, 4G, Wi-Fi, ZigBee, and Bluetooth. Fog nodes themselves communicate using IP core networks for interconnection and collaboration with the cloud.

3.3.1 Load Balancing in Fog Systems

In fog computing, load balancing ensures an equitable distribution of workloads across resources, enabling continuous service availability even in the event of component failure. This is achieved by provisioning and de-provisioning application instances while optimizing resource utilization. Since data centers exhibit heterogeneity among hosts and distinct traffic patterns, a tailored load-balancing mechanism is crucial to improving network performance and application efficiency.

Load balancing prevents resource overload or underutilization by distributing workloads across available resources. It can be implemented using either physical equipment or software. The objectives of load balancing include maximizing throughput, optimizing traffic, minimizing response times, and enhancing resource utilization and scalability. Load balancing techniques in fog networks can be static, dynamic, or a combination of both. [10]

DOI 10.5281/zenodo.14786908

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 6 | December 2024

- **Static Methods:**
  These rely on pre-existing system information and pre-programmed rules. They are less efficient in dynamic environments due to the unpredictable nature of user behavior.

- **Dynamic Methods:**
  These outperform static approaches by adapting to current system states. Dynamic methods rely on specific policies:

  Transfer Policy: Determines conditions for transferring tasks between nodes, involving task migration and rescheduling.

  Selection Policy: Evaluates whether a task should be transferred, considering factors like migration overhead and execution time.
  Location Policy: Identifies underutilized nodes and transfers tasks to them, verifying resource availability for rescheduling or migration.

  Information Policy: Collects system node data to support decision-making in other policies.
  These policies work in conjunction to ensure tasks are efficiently processed, either locally or on remote nodes [9].

3.3.2 Challenges in Fog Computin

Fog computing, an evolution of cloud computing, addresses IoT-related challenges at the network edge. However, its distributed and heterogeneous nature presents several challenges:

- **Service-Oriented Challenges:** Many fog nodes have limited resources, making it difficult to deploy large-scale applications. Developing distributed applications requires suitable programming platforms and policies for task distribution among IoT devices and fog infrastructure. [8]

- **Structural Issues:** Fog infrastructure comprises diverse components, often not designed for general-purpose computing. Customizing computational units, selecting appropriate devices, deployment locations, and resource configurations are critical challenges. Efficient inter-nodal cooperation and resource provisioning strategies are essential due to the distributed nature of fog nodes.

- **Security Challenges:** Fog computing relies on conventional networking components, making it vulnerable to security attacks. Ensuring privacy, authenticated access, and data center integrity in a widely distributed model is challenging. Implementing robust security measures while maintaining QoS remains a significant hurdle [5].

DOI 10.5281/zenodo.14786908

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 6 | December 2024

IV. SYSTEM ARCHITECTURE

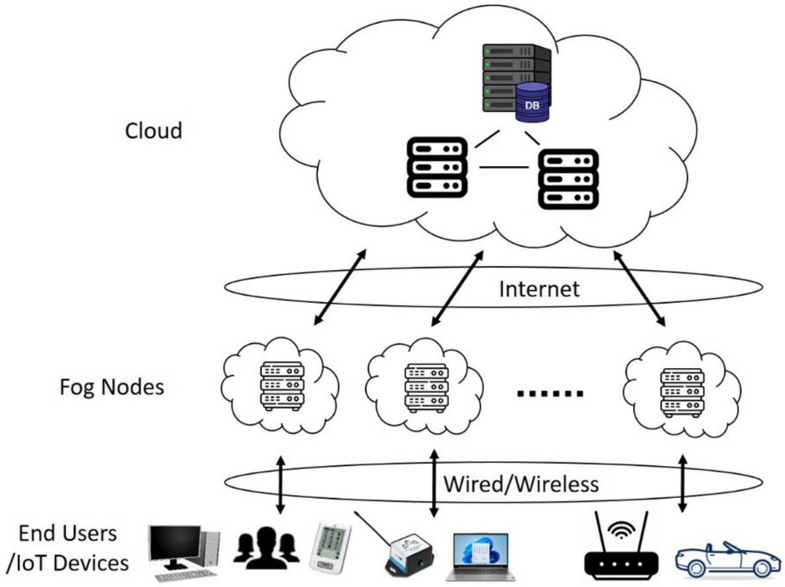

Figure 1: The fog Computing System Architecture [7]

V. PROPOSED SYSTEM

Cloud load balancing refers to the process of distributing workloads and computational resources across a cloud environment. It allows organizations to manage application or workload demands by allocating resources across multiple servers, networks, or computers. This approach involves regulating the flow of workload traffic and demands over the Internet. With Internet traffic increasing exponentially and accounting for nearly all current traffic growth annually, server workloads are also rising rapidly, often leading to server overloads, especially for highly popular web servers.
To address server overload, two primary solutions are commonly used:

- **Single-Server Solution:**
  This involves upgrading an existing server to one with higher performance capabilities. However, this approach has limitations, as the upgraded server may quickly become overloaded again, necessitating another upgrade. Additionally, the upgrading process can be complex and expensive.

- **Multiple-Server Solution:**
  A more scalable and cost-effective option is to create a cluster of servers that function as a unified, scalable service system. Server cluster systems are particularly suitable for managing network services. Cloud-based servers can enhance scalability and availability through server farm load balancing. This method ensures effective distribution of traffic and workload and is advantageous for various services, including HTTP, SMTP, DNS, FTP, and POP/IMAP.

 DOI 10.5281/zenodo.14786908

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

www.ijirid.in

Volume 3 | Issue 6 | December 2024

VI. CONCLUSION

In the 21st century, the fog computing (FC) paradigm is expected to remain a focal point for researchers in both industry and academia due to its immense potential. It enables the efficient distribution of services and computational resources through fog nodes (FNs) positioned at the edge of cloud networks. This paper specifically addressed the task scheduling challenge within fog computing environments, aiming to optimize task execution based on available processing capacity and remaining energy. With thousands of users potentially accessing a website simultaneously, managing the influx of requests can become overwhelming for applications. Such scenarios may result in system failures, underscoring the importance of efficient load balancing mechanisms.

REFERENCES

[1] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in Proceedings of the first edition of the MCC workshop on Mobile cloud computing, 2012: ACM, pp. 13-16.

[2] A. Yousefpour et al., "All one needs to know about fog computing and related edge computing paradigms: A complete survey," Journal of Systems Architecture, vol. 98, pp. 289-330, 2019/09/01/ 2019.

[3] N. Auluck, A. Azim, and K. Fizza, "Improving the Schedulability of Real-Time Tasks using Fog Computing," IEEE Transactions on Services Computing, pp. 1-1, 2019.

[4] S. Aslam and M. A. Shah, "Load balancing algorithms in cloud computing: A survey of modern techniques," in 2015 National Software Engineering Conference (NSEC), 2015: IEEE, pp. 30-35.

[5] Resul Das and Muhammad Muhammad Inuwa, "A review on fog computing: Issues characteristics, challenges and potential applications," Elsevier Journal telematics and Informatics Reports, vol. 10, June 2023, 100049.

[6] P. Jamshidi, A. Ahmad, and C. Pahl, "Cloud Migration Research: A Systematic Review," IEEE Transactions on Cloud Computing, vol. 1, no. 2, pp. 142-157, 2013.

[7] Tang, S., "Performance Modeling and Optimization for a Fog-Based IoT Platform", IoT 2023, 4, 183–201.

[8] C. Jatoth, G. R. Gangadharan, and R. Buyya, "Computational Intelligence Based QoS-Aware Web Service Composition: A Systematic Literature Review," IEEE Transactions on Services Computing, vol. 10, no. 3, pp. 475-492, 2017.

[9] S. Junare and M. Karande, " A systematic review on Load Balancing Mechanism in Fog Computing", Journal of Progressive Research,, vol. 01, Issue 01, April 2024.

[10] M. Haghi Kashani, A. M. Rahmani, and N. Jafari Navimipour, "Quality of service-aware approaches in fog computing," International Journal of Communication Systems, 2020/01/10 2020.

[11] M. Rahimi, M. Songhorabadi, and M. H. Kashani, "Fog-based smart homes: A systematic review," Journal of Network and Computer Applications, vol. 153, p. 102531, 2020/03/01/ 2020.

[12] S. Bazzaz Abkenar, M. Haghi Kashani, E. Mahdipour, and S. M. Jameii, "Big data analytics meets social media: A systematic review of techniques, open issues, and future directions," Telematics and Informatics, p. 101517, 2020/10/14/ 2020.

[13] P. Hu, S. Dhelim, H. Ning, and T. Qiu, "Survey on fog computing: architecture, key technologies, applications and open issues," Journal of network and computer applications, vol. 98, pp. 27-42, 2017.

[14] P. Asghari, A. M. Rahmani, and H. H. S. Javadi, "Internet of Things applications: A systematic review," Computer Networks vol. 148, pp. 241-261, 2019/01/15/ 2019.

[15] E. Marín-Tordera, X. Masip-Bruin, J. García-Almiñana, A. Jukan, G.-J. Ren, and J. Zhu, "Do we all really know what a fog node is? Current trends towards an open definition," Computer Communications, vol. 109, pp. 117-130, 2017.

[16] O. C. A. W. Group, "OpenFog reference architecture for fog computing," OPFRA001, vol. 20817, p. 162, 2017.

[17] C. Mouradian, D. Naboulsi, S. Yangui, R. H. Glitho, M. J. Morrow, and P. A. Polakos, "A comprehensive survey on fog computing: State-of-the-art and research challenges," IEEE Communications Surveys & Tutorials, vol. 20, no. 1, 2017.

[18] M. Iorga, L. Feldman, R. Barton, M. Martin, N. Goren, and C. Mahmoudi, "Fog computing conceptual model, recommendations of the National Institute of Standards and Technology," NIST Special Publication, pp. 500-325, 2018.

