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 2 | April 2024

A Systematic Review on Load Balancing Mechanism in Fog Computing

Prof. Manjiri Karande[1], Ms. Sakshi Dinkar Junare[2]

[1]*Assistant Professor, Padmashri Dr. V. B. Kolte College of Engineering, Malkapur, India*

[2]*Post-Graduate Student, Padmashri Dr. V. B. Kolte College of Engineering, Malkapur, India*

Abstract: *Recently, fog computing has been introduced as a modern distributed paradigm and complements cloud computing to provide services. Fog system extends storing and computing to the edge of the network, which can solve the problem of service computing of the delay-sensitive applications remarkably besides enabling location awareness and mobility support. Load balancing is an important aspect of fog networks that avoids a situation with some under-loaded or overloaded fog nodes. Quality of Service (QoS) parameters such as resource utilization, throughput, cost, response time, performance, and energy consumption can be improved with load balancing. In recent years, some research in load balancing techniques in fog networks has been carried out, but there is no systematic review to consolidate these studies. This article reviews the load balancing mechanisms systematically in fog computing in four classifications, including approximate, exact, fundamental, and hybrid methods (published between 2013 and August 2020). Also, this article investigates load-balancing metrics with all advantages and disadvantages related to chosen load-balancing mechanisms in fog networks. The evaluation techniques and tools applied for each reviewed study are explored as well. Additionally, the essential open challenges and future trends of these mechanisms are discussed.*

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

I. INTRODUCTION

Fog computing, which extends from the cloud and is a geographically distributed paradigm, brings networking power and computing into the network edge, closer to end-users and IoT devices both because of being supported by wide-spread fog nodes. Most of the data, in cloud-only architectures, requiring processing, analysis, and storage, are transmitted to the cloud servers, which may have an influence on latency, security, mobility, and reliability adversely. With the existence of location-aware and delay-sensitive applications, the cloud on its own comes across some problems in meeting the extremely low latency requirements of these applications; the proximity of the fog layer to the Internet of Things (IoT) devices may remarkably decrease latency and meet the needs of extremely-low. Fog computing always interacts with and supports the cloud, creating a novel generation of applications and services.

Nowadays, in fog computing environments, users need applications that give quick responses whenever they want to access anything and work fast. To improve QoS factors in a fog network significantly, we can use an efficient load-balancing strategy because load balancing is regarded as an

 DOI: 10.5281/zenodo.11114949

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 2 | April 2024

important issue. Many studies have been done to balance the cloud computing load because the load on the cloud increases enormously. Being heterogeneous and dynamic, the fog networks cannot directly apply most of the load-balancing mechanisms of cloud computing; The goal of load balancing in a fog environment is to distribute the coming load between available fog nodes or cloud, based on one mechanism, to avoid overload or under-load of fog nodes. This mechanism can maximize throughput, performance, and resource utilization while minimizing response time, cost, and energy consumption.

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 data with traditional computing models, like cloud computing, distributed computing, etc. is difficult. Often privacy gaps, high communication delays, and 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.) 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.

Fog computing is regarded as a platform bringing cloud computing to the end-user's vicinity. “Fog”, as a term, has an analogy with real-life fog and was initially introduced by Cisco. 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 the end user's 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.

Bonomi, et al., 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 centres 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 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 processing. 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.

DOI: 10.5281/zenodo.11114949

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 2 | April 2024

www.ijirid.in

III. FOG ARCHITECTURE

Fog computing architecture reference model is an important study topic. Recently, a wide range of architectures has been suggested for fog computing, mostly obtained from a structure with three layers. A fog network expands cloud services to the network edge by suggesting a fog layer between the cloud and user devices. As it can be observed, Figure 1 illustrates the fog architecture hierarchically, having three layers as follows:

- **Cloud Layer:**
  The layer of cloud computing has been composed of different storage devices and high-performing servers and creates several services of applications. It bears robust storage and computing abilities to back the permanent storage of a large amount of information and extended computation analysis. However, it should be noted that all computing and storage tasks do not pass the cloud which is not the same as traditional cloud computing architecture.

- **Fog Layer:**
  The fog layer is located at the network edge, which consists of a couple of fog nodes like access points, routers, switches, gateways, etc. They are spread between cloud and end devices. In order to get services, end devices might easily connect with fog nodes. They are capable of computing, storing, and transmitting the received sensed data. The latency-sensitive applications and real-time analysis can be performed in a fog layer. In addition, we can refer to the connection between the cloud data centre and the fog nodes by the IP core network. To get more robust storage and computing capabilities, fog nodes have the responsibility of interacting and cooperating with the cloud.

- **User Device:**
  The layer of a user device is so close to the physical environment and end-user. This layer is composed of different IoT devices, like, sensors, cellphones, smart automobiles, cards, and readers. Although cell phones and smart vehicles have computing capabilities, they are utilized as just smart sensing devices. Generally and geographically, they are widely distributed and responsible for sensing feature data related to events or physical objects and for transferring them to upper layers to be processed and stored. Here in the architecture, all end devices or smart objects are connected with fog nodes by technologies with wired or wireless connection access such as 3G, 4G, wireless LAN, ZigBee, Bluetooth, and Wi-Fi. Wireless or wired communication technologies to help the interconnection and intercommunications of fog nodes. IP core network helps fog nodes to be linked with the cloud.

IV. CONCLUSION

In the 21st century, the FC paradigm is expected to remain of interest to researchers in industry and academia because of its incredible potential where services and computing resources are distributed in effective FNs residing at the cloud computing network edge. In this paper, we concentrated on the task scheduling problem in the environment of FC to ensure effective task execution according to the

 DOI: 10.5281/zenodo.11114949

available processing capacity and remaining energy. Thousands of people have access to a website at a particular time. This makes it challenging for the application to manage the load coming from these requests at the same time. Sometimes this can lead to system failure.

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] 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, pp. 416-464, 2017.

 DOI: 10.5281/zenodo.11114949