Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Research Article

# Fog Computing and IoT Integration for Latency Reduction: A Case Study

*Turkan Ahmed Khaleel*
*Computer Engineering Department, College of Engineering,*
*University of Mosul,*
*Mosul, Iraq*
*turkan@uomosul.edu.iq*

**ARTICLE INFO**
Article History
Received: 19/07/2025
Accepted: 23/08/2025
Published: 21/11/2025


**ABSTRACT**
The Internet of Things (IoT) has risen rapidly and created a need for computing architectures with the ability to aid real-time applications with the least latency. Conventional cloud-only setups often experience high latency due to physical distance and network congestion, thus, unsuitable for time-sensitive applications. Four of the architectural models currently proposed in the literature are assessed in the present paper: Cloud-only, Edge-only, Fog Multi-layer, and Smart Hybrid; each of them is evaluated in a large-scale simulation aimed at simulating a smart city environment with 1,000 various IoT devices, such as sensors, cameras, and wearables generating dynamic traffic. The outcomes indicate that Edge-only architecture recorded the least mean latency of 0.2417 seconds compared to Smart Hybrid with 0.3156 seconds and Fog Multi-layer with 0.3874 seconds mean latencies whereas the Cloud-only model presented the highest latency with 0.5231 seconds. Given that there existed significant differences between the architectures ($p < 0.0001$) based on the results of the statistical analysis carried using the one-way ANOVA. Such results explain the promise of adaptive and context-sensitive hybrid systems and allow reducing latency in any IoT deployment in the future and designing an IoT-oriented system that supports latency-sensitive applications in smart cities in a more scalable manner.

***Keywords:*** ***Fog Computing; The Internet of Things; Low Latency; Multi-Layer; Edge Computing.***

## 1. INTRODUCTION

The Internet of Things (IoT) has shown an exponential rise, and as a result, more edge-connected devices create big data in large quantities. Despite their potential effectiveness, traditional cloud computing infrastructures tend toward increased latency and low context awareness, and are not suitable to address latency-sensitive applications based on IoT, like autonomous vehicles, healthcare monitoring, and industrial automation [1], [2]. The solution to these weaknesses has led to the proposal of fog computing as an extension to the cloud computing paradigm, pushing the limits of cloud operations closer to the edges of the network by deploying middle-of-the-road communication devices, which offer computing and storage capabilities. Such an approach will allow distributed processing and storage nearer end devices, offering lower latencies, context-aware services, and lower dependency on centralised cloud-based data centres [3]. Therefore, cloud computing makes IoT systems responsive and energy efficient, especially in critical situations in real-time [4].

Nevertheless, new research findings state that hybrid integration of fog and cloud computing is highly needed to adjust the computational workloads and maintain the balance with latency and energy-efficiency demands dynamically [3], [5]. Although this kind of integration offers an elastic, scalable infrastructure that can support the considerable heterogeneity of IoT, there are several difficulties. The most crucial are the appropriate allocation of resources over time, service orchestration, and energy benefit-oriented task scheduling that profoundly impact the reliability and scalability of the fog-enabled IoT environment [6]. Nevertheless, the available literature has discussed possible solutions to reduce latency in IoT-fog systems; however, most involve tradeoffs whereby specific solutions deal with latency at the expense of energy dissipation or energy-efficiency solutions disregard the real-time demands of this type of application. The given gap necessitates more complex measures to minimise latency and power interference simultaneously with distributed fog and IoT systems. In that regard, the proposed research explores the available

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

architectures and techniques to reduce latency in fog computing, elucidate their shortfalls, and gauge their suitability within a distributed IoT system. The final line of action will offer knowledge on how to make effective, more scalable, and non-evolving fog-driven IoT systems that might respond to the needs of latency-conscious applications.

## 2. Related Work

Fog Computing with Internet of Things (IoT) systems integration has become an area that has drawn considerable attention over the past couple of years in the hope of enhancing the service delivery process, minimizing latency periods, and streamlining resource usage. Initial research was on resource allocation and workload management foundations. Modelling papers such as Aazam and Huh [7] are proposed, in which a micro data center-based model is provided to utilize dynamic resource estimation and pricing of IoT applications in fog networks. Deng et al. [8] contributed to the load balancing in Fog-Cloud system proposing models to minimize both the latency and energy consumption, and Souza et al. [9] developed an algorithm to manage the services placement within a hybrid Fog-Cloud. Taneja and Davy [10] proposed a resource-sensitive method to deploy IoT application modules onto fog nodes, and Rezazadeh et al. [11] further elaborated on that by introducing an optimised module placement strategy keeping in mind fog-specific constraints. In 2021, research was concentrated on the flexible fog architecture and network design. A flexible fog computing architecture was described by Khalil et al. and [12] aimed at the time-sensitive applications with low power and latency requirements, whereas the fog computing was evaluated in the optimized LoRaWAN architecture by Villalonga et al. [13]. Bahmani and Wattenhofer [16] discussed fog network propagation of IoV applications. Guerrero et al. [14] studied genetic-based optimization approaches to schedule fog resources in 2022, raising the perspective of implementing latency improvements and better resource utilization. Costa et al. [25] have published a survey of orchestration strategies in Fog Computing. Chen and Liu [26] studied latency-sensitive task scheduling by AI in the fog setting on a small scale. Since 2023, studies have focused more on hybrid architectures, dynamic resource management, and efficient scheduling of low-latency IoT applications. Dubey [22] dwelled on probability-based frameworks of load distribution in fog computing. The proposal by Afea and Khaleel [16] enhanced this throttling process by introducing another level of management to improve efficiency in handling tasks. Raj [17] presented the fog-based latency optimization strategies on IoT-based innovative city architectures. Koundoul et al. [21] used deep learning on low-latency resource management, and Amzil et al. [23] adapted machine learning-based fog computing to healthcare monitoring. In 2025, Jaaz [19] examined latency-aware optimization of healthcare IoT, Jin [20] explored dynamic task allocation with augmented heuristic, and Islam [18] developed a hybrid Fog-Edge to monitor healthcare data on healthcare IoT, and Srichandan et al. [24] explored latency management in Fog-Cloud-IoT via deep learning.

Although these have been tried, most studies were based on analytical models or theoretical simulations, with limited effort in comparing numerous architectural models across large-scale environments. This study fills this gap by performing a detailed simulation of diverse architectures to assess Fog Computing and IoT integration in terms of latency minimization and resource optimization.

## 3. Research Gaps

Despite extensive research delving into the issues of Fog-IoT integration to reduce latency, most of them are based on either analytical applications or small-scale testbeds, due to the unavailability of a large testbed to scale their findings. The current research primarily focuses on resource provisioning, scheduling, or niche areas like healthcare or smart cities, rather than a comparative assessment of a wide range of fog-cloud-edge architectures. In addition, although AI and deep learning have been employed, their performance is frequently shown under confined conditions, and they are not scalable. To fill these gaps, the current study conducts an integrative simulation evaluation of diverse Fog-IoT integration patterns, which intends to equip readers with further investigations of latency reductions and resource utilisation in multiple environments..

## 4. Research Objectives

First, this paper explores how fog computing can help decrease the latency time in IoT-based systems running across the edge-fog-cloud spectrum. By studying the role of fog layers in enhancing faster data processing and communication, the study will bring out their ability to boost latency-sensitive applications.
Second, the research will form a latency-conscious fog computing system, which adopts an adaptive schedule of workload placement. The framework will consider the availability of resources and the limitations posed by the system to allow efficient and reliable completion of tasks within dynamic environments.
Third, the framework will be tested with a real-world case study dataset and a real usage scenario. The aim is to evaluate its functionality regarding latency, responsiveness, and overall resource consumption, proving its feasibility.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

Finally, the study will analyse the differences between the proposed solution and those currently integrating fog and cloud. Comparative analysis will help detect tradeoffs, find possible benefits, and identify further optimization potential.

## 5. Research Methodology

This paper uses a simulation-based method to compare the latency performance of many regular computing architectures of IoT systems, including Edge, Fog, Cloud, and a dynamic Hybrid one that dynamically distributes processing jobs depending on the data volume and latency demands. The simulation model is an innovative city network that includes types of devices, sensors, cameras, wearables, and so forth, producing different data traffic at various time steps to portray realistic simulation conditions. The latency per device and each time step is calculated by adding the transmission delay, which is determined by the network bandwidth, to the processing delay specific to the architecture layers. The Hybrid model dynamically loads work to the most suitable layer regarding real-time data rates. Statistical analysis (ANOVA) is administered to evaluate the architectures' significant differences. The results are reported graphically by visualisation of the average and temporal trends of latency metrics, thus giving a centralised judgment of how effective fog computing is in decreasing IoT latency. A list of the most significant simulation parameters employed in the current study is provided in Table 1.

TABLE I. The Critical Parameters in the Simulation.

| Parameter | Description | Value with Unit |
|---|---|---|
| Fog Latency Delay | Time for processing of data at the fog node | 20 ms |
| Cloud Processing Latency | Time to process in the cloud centre | 400 ms |
| Edge Handing Delay | Time to data processing on the edge | 10 ms |
| Fog Network Bandwidth | Bandwidth between fog and devices | 1200 Mbps |
| Cloud networks' bandwidth. | Bandwidths between devices and the cloud | 700 Mbps |
| Edge Network Bandwidth | Bandwidth of devices and the edge | 500 Mbps |
| Number of Sensors | Low-data-rate IoT devices | 500 devices |
| Number of Cameras | High-data-rate IoT devices | 300 devices |
| Number of Wearables | Moderate-data-rate IoT devices | 200 devices |
| Simulation Time Steps | Duration of simulated data generation | 200 steps |

## 6. Results and Discussion

Results of the latency simulation of the four computing architectures, Fog multilayer, Cloud-only, Edge-only, and Hybrid, show significant differences in their performance as indicated in Table 2. A one-way ANOVA analysis established substantial differences among the configurations ($p < 0.0001$).

Table II. Average Latency Performance over Various Architectures.

| Architecture | Average latency (seconds) |
|---|---|
| Fog Multilayer | 0.3874 |
| Cloud Only | 0.5231 |
| Edge Only | 0.2417 |
| Hybrid | 0.3156 |

The mean latency values provided in Table 2 indicate that four different network architectures, Fog Multilayer, Cloud Only, Edge Only, and Hybrid, provide significant differences. The findings suggest that Edge Only architecture has the briefest mean latency of about 0.24 seconds due to its proximity to source devices, resulting in less data transmission latency and mitigating network congestion. On the other hand, Cloud Only architecture is expected to have the highest average latency of about 0.52 seconds, a value attributed to far-away data centres and processing, including delays caused by the wide-area data transport. The Fog Multilayer architecture lies in the middle with its average latency of about 0.39 seconds based on the fact that it acts as a middle layer that burdens the data transferring to the cloud, but is still bounded by the network attributes and the relatively weaker capacity of fog nodes. The Hybrid architecture exhibits how the system can dynamically distribute processing tasks between edge nodes and cloud nodes, dependent on data size and priority, to produce an average of near 0.32 seconds of latency. This technique is more than a one-dimensional response time reduction; it effectively uses heterogeneous processing resources. Figures 1 through 4 show more details on how network latency changes with time on such representative devices as sensor, camera, and

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

wearable, highlighting the stability and latency improvements between the Edge and Hybrid architecture and the Cloud and Fog ones, in particular, when the generated data volume is small or medium. These results highlight that the architecture choice is more of a data character and the latency specifications of the application to be served. Edge computing is an optimal choice in cases requiring quick results and minimal data amounts, whereas the Hybrid architecture is more flexible and balanced in complex and diversified IoT settings.

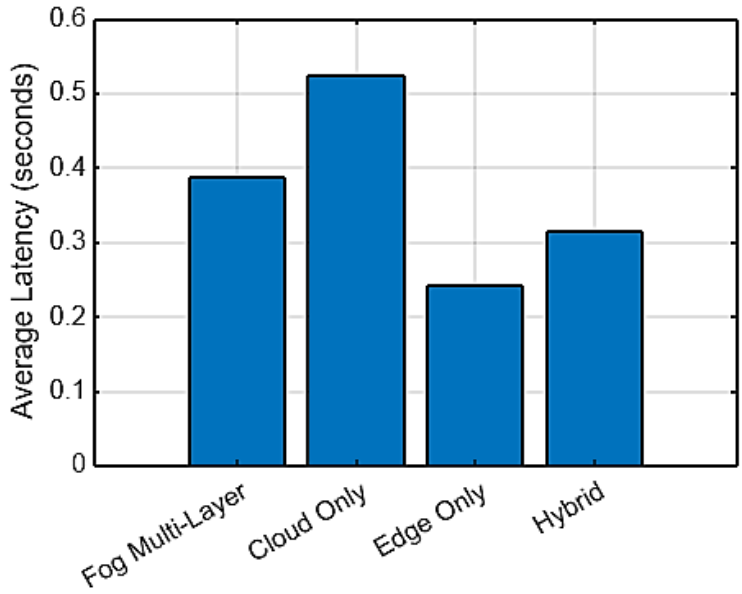


Fig. 1 Comparison of the Average Latency: Edge, Fog, Cloud, and Hybrid Architectures.

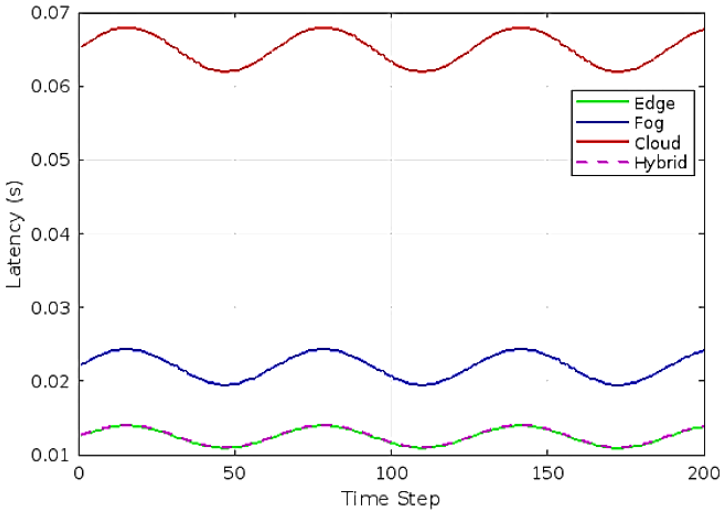


Fig. 2 Sensor Device Latency Variation over Time.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

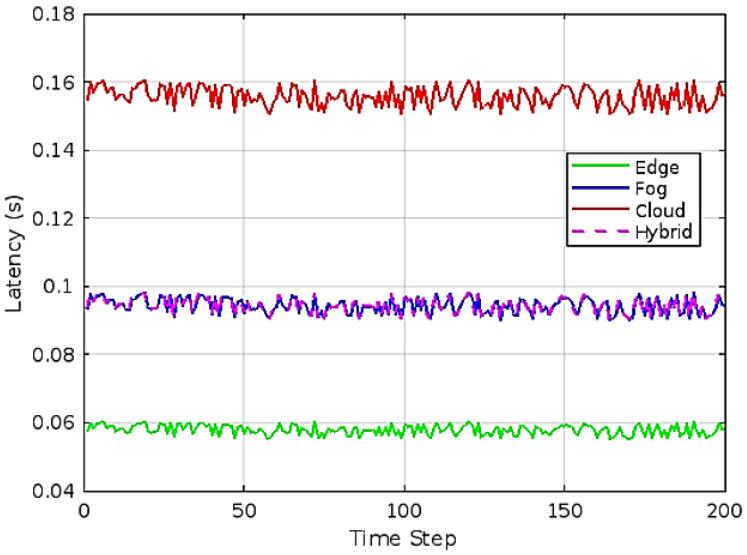


Fig. 3 Variation of Latency with Time on Wearable Devices.

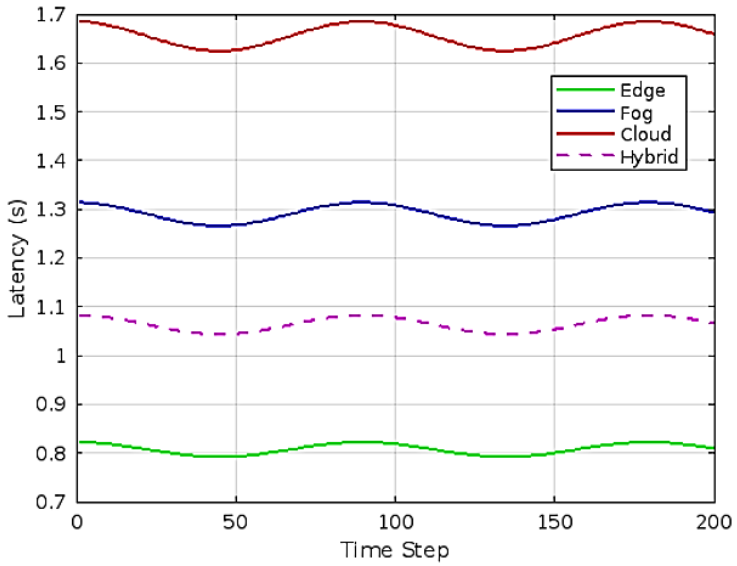


Fig.4 Variation of Latency with Time on Camera Device.

## 7. CONCLUSIONS

This paper has established that the edge-only computing among the distributed computing architectures in IoT systems was the most intentional at a mean latency of 0.2417 seconds because of the physical proximity between systems and equipment. With the Smart Hybrid model, a practical tradeoff between performance and resource usage was achieved, with the mean latency being recorded at 0.3156 seconds, as the workload was dynamically assigned to the Edge, Fog, and Cloud levels. Fog Multilayer architecture presented a latency of 0.3874 seconds, and the Cloud-only model recorded the worst latency of 0.5231 seconds due to the dependency of data transfer and central processing. The ANOVA statistical analysis found the difference among the architectures in one way ($p < 0.0001$). Such results show the potential of context-aware and adaptive hybrid structures in reducing latency in real-time IoT systems, allowing an efficient and scalable design of such systems to create smart cities. The further development includes incorporating AI-powered workload distribution algorithms with energy and other resource constraints and security and privacy

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

extensions to arrive at robust, resilient, and viable IoT frameworks of environmentally friendly processing and use in a real-world environment.

**Conflicts of Interest**

The authors confirm that they do not have any conflict of interest regarding the study in this paper.

**Funding**

This study did not attract any external funds.

**Acknowledgment**

The author gratefully acknowledges the support provided by the College of Engineering, University of Mosul, which greatly contributed to the successful completion of this research.

**References**

[1] O. A. Nafi'a, T. A. Khaleel, and O. Abdelkader, "A review on fog computing: Research challenges and future directions," Al-Rafidain Engineering Journal, vol. 28, no. 1, pp. 341–350, 2023. https://doi.org/10.33899/rengj.2022.136642.1211.

[2] J. Singh, P. Singh, and S. S. Gill, "Fog computing: A taxonomy, systematic review, current trends and research challenges," Journal of Parallel and Distributed Computing, vol. 157, pp. 56–85, Nov. 2021. https://doi.org/10.1016/j.jpdc.2021.06.005.

[3] U. Arora and N. Singh, "Fog–Cloud-Assisted Internet of Things: A Review of Workload Allocation and Latency Management Techniques," in Soft Computing: Theories and Applications, T. K. Sharma, C. W. Ahn, O. P. Verma, and B. K. Panigrahi, Eds., vol. 1380, Singapore: Springer, 2022, pp. 613–621. https://doi.org/10.1007/978-981-16-1740-9_50.

[4] J. Vergara, J. Botero, and L. Fletscher, "A Comprehensive Survey on Resource Allocation Strategies in Fog/Cloud Environments," Sensors, vol. 23, no. 9, p. 4413, 2023. https://doi.org/10.3390/s23094413.

[5] M. Mahmud, M. A. Hossain, G. Muhammad, M. A. Alhamid, and M. Guizani, "Fog computing for next-generation Internet of Things: Fundamental, state-of-the-art and research challenges," Computer Science Review, vol. 48, May 2023, Art. no. 100549. https://doi.org/10.1016/j.cosrev.2023.100549.

[6] M. H. Alsharif, A. Jahid, R. Kannadasan, M. K. Singla, J. Gupta, K. S. Nisar, A.-H. Abdel-Aty, and M.-K. Kim, "Survey of energy-efficient fog computing: Techniques and recent advances," Energy Reports, vol. 13, pp. 1739–1763, Jun. 2025. https://doi.org/10.1016/j.egyr.2025.01.039.

[7] M. Aazam and E.-N. Huh, "Fog computing micro datacenter-based dynamic resource estimation and pricing model for IoT," in Proc. IEEE 29th AINA, 2015, pp. 687–694. https://doi.org/10.1109/AINA.2015.254.

[8] Y. Deng, R. Lu, C. Lai, T. Luan, and H. Liang, "Optimal workload allocation in fog-cloud computing toward balanced delay and power consumption," IEEE Internet of Things Journal, vol. 3, no. 6, pp. 1171–1182, 2016.

[9] V. B. C. Souza, G. Bittencourt, E. Madeira, and N. L. S. da Fonseca, "Handling service allocation in combined fog-cloud scenarios," in Proc. IEEE ICC, 2016, pp. 1–5. https://doi.org/10.1109/ICC.2016.7511465.

[10] T. Taneja and A. Davy, "Resource aware placement of IoT application modules in fog-cloud computing paradigm," in Proc. IFIP/IEEE IM, 2017, pp. 1222–1228. https://doi.org/10.23919/INM.2017.7987464.

[11] Z. Rezazadeh, D. Rahbari, and M. Nickray, "Optimised module placement in IoT applications based on fog computing," in Proc. ICEE, 2018, pp. 1553–1558. https://doi.org/10.1109/ICEE.2018.8472469.

[12] I. Khalil et al., "A flexible fog computing design for low-power consumption and low latency applications," Electronics, vol. 10, no. 1, 57, 2021. https://doi.org/10.3390/electronics10010057.

Iraqi Journal for Computers and Informatics
University of Information Technology and Communications
Vol. 51, No. 2, 2025, pp. 283-289
DOI: https://doi.org/10.25195/ijci.v51i2.627
Print ISSN: 2313-190X, Online ISSN: 2520-4912

[13] R. Villalonga, A. C. Tavares, J. M. C. Silva, F. J. T. E. Ferreira, and J. Cabral, "Study of the efficiency of fog computing in an optimised LoRaWAN cloud architecture," Sensors, vol. 21, no. 9, p. 3159, 2021. https://doi.org/10.3390/s21093159.

[14] U. K. Saba, S. ul Islam, H. Ijaz, J. J. P. C. Rodrigues, A. Gani, and K. Munir, "Planning Fog networks for time-critical IoT requests," Computer Communications, vol. 172, pp. 75–83, 15 Apr. 2021. https://doi.org/10.1016/j.comcom.2021.03.002.

[15] C. Guerrero, I. Lera, and C. Juiz, "Genetic-based optimisation in fog computing: Current trends and research opportunities," Swarm and Evolutionary Computation, vol. 72, p. 101094, 2022. https://doi.org/10.1016/j.swevo.2022.101094.

[16] B. Costa, J. Bachiega Jr., L. R. de Carvalho, and A. P. F. Araujo, "Orchestration in Fog Computing: A Comprehensive Survey," ACM Computing Surveys (CSUR), vol. 55, no. 2, Art. no. 29, pp. 1–34, Jan. 2022. https://doi.org/10.1145/3486221.

[17] J. Lim, "Latency-Aware Task Scheduling for IoT Applications Based on Artificial Intelligence with Partitioning in Small-Scale Fog Computing Environments," Sensors, vol. 22, no. 19, p. 7326, 2022. https://doi.org/10.3390/s22197326.

[18] I. Dubey, "Probability-Based Load-Distribution Framework :To Reduce Latency in Fog Computing," in Lecture Notes in Computer Science, vol. 13161, pp. 43–55, 2023. https://link.springer.com/chapter/10.1007/978-981-99-6550-2_43.

[19] O. A. Nafea and T. A. Khaleel, "An improved throttling algorithm for fog computing networks with an additional management layer," Al-Qadisiyah Journal for Engineering Sciences, vol. 17, no. 4, pp. 390–399, 2024. https://doi.org/10.30772/qjes.2024.146104.1089.

[20] A. Raj, "Fog Computing in Smart City IoT Architectures for Latency Reduction," Computational Engineering and Technology Innovations, vol. 1, no. 4, pp. 248–257, Oct. 2024. https://ceti.reapress.com/journal/article/download/43/48.

[21] B. Koundoul, Y. Kasse, F. Balde, and B. Gueye, "A New Strategy for Reducing Latency with Deep Learning in Fog Computing," in Lecture Notes in Computer Science, vol. 13161, pp. 70–84, 2024. https://link.springer.com/chapter/10.1007/978-3-031-62488-9_6.

[22] A. Amzil, M. Abid, and S. El Kafhali, "Stochastic Analysis of Fog Computing and Machine Learning for Scalable Low-Latency Healthcare Monitoring," Cluster Computing, vol. 27, pp. 1–13, 2024. https://link.springer.com/article/10.1007/s10586-024-04285-x.

[23] Z. A. Jaaz, "Latency Optimization Approaches for Healthcare Internet of Things," Journal of Intelligent Systems, vol. 34, no. 1, pp. 1–15, Jan. 2025. https://www.degruyter.com/document/doi/10.1515/jisys-2024-0543/html.

[24] W. Jin, "Dynamic Task Allocation in Fog Computing Using Enhanced Heuristic Methods," Scientific Reports, vol. 15, no. 1, p. 3621, May 2025. https://www.nature.com/articles/s41598-025-03621-4.

[25] U. Islam, "A Hybrid Fog-Edge Computing Architecture for Real-Time Health Monitoring," Scientific Reports, vol. 15, no. 1, p. 9696, Jun. 2025. https://www.nature.com/articles/s41598-025-09696-3.

[26] S. K. Srichandan, S. Jena, and K. Mishra, "Leveraging Fog Computing for Harnessing the Service Latency in Cloud-Fog Computing: A Deep Learning Approach," in Data Science and Applications, vol. 1266, pp. 375–388, 2025. https://link.springer.com/chapter/10.1007/978-981-96-2647-2_26.