www.ijecs.in
International Journal Of Engineering And Computer Science
Volume 13 Issue 09 September 2024, Page No. 26597-26607
ISSN: 2319-7242 DOI: 10.18535/ijecs/v13i09.4926

Emergent Architectures in Edge Computing for Low-Latency Application

Gireesh Kambala

MD, CMS Engineer, Lead, Teach for America, USA.

Abstract
Emerging as a necessary paradigm for meeting the low-latency needs of contemporary real-time applications is edge computing. By distributing computation, storage, and network resources, edge designs reduce data transfer latency and increase system responsiveness. They are indispensible in fields including smart cities, autonomous systems, and healthcare. This study explores evolving architectural paradigms in edge computing—including layered hierarchies, microservices, and serverless computing—as well as their interfaces with technologies including 5G, IoT, and artificial intelligence. Extensive studies reveal how they allow scalable, modular, resource-efficient solutions for activities sensitive to latency. Even in this regard, security, interoperability, and resource allocation demand constant innovation notwithstanding their changing power. The work also looks at innovative ideas addressing these issues and highlights possible opportunities such federated learning and quantum computing. The outcomes highlight how crucial emergent edge computing architectures are in enabling ultra-low-latency applications and redefining operational efficiencies in many industries.

Keywords: Edge Computing, Low-Latency Applications, 5G Integration, Architectural Paradigms

I. Introduction

By moving data storage and processing closer to the point of origin, a novel paradigm known as "edge computing" has lately evolved to satisfy the growing demand for low-latency applications. Traditional cloud computing architectures aren't fast enough to meet the real-time processing needs of applications like smart grids, augmented reality, telemedicine, autonomous vehicles, and industrial internet of things (IoT), notwithstanding its scalability and adaptability. By spreading computation over a distributed network of nodes, which enhances data security, lowers bandwidth usage, and accelerates reaction times, new edge computing architectures offer inventive responses to these challenges. These systems combine several technologies and methodologies, including microdata centres, fog computing, and multi-access edge computing (MEC), so optimising resource allocation and allowing seamless interactions between edge devices and the cloud[1]–[3]. By moving processing duties from centralised infrastructure to local base stations or nodes in the network, MEC offers low-latency services at the periphery of a network thanks to 5G advances. New opportunities abound from this, including interactive gaming and real-time video analytics. Fog computing advances this concept by stressing proximity and scalability by creating a hierarchical ecosystem of resources including sensors, actuators, intermediate nodes, and the cloud. New edge architectures ensure real-time insights by depending more on artificial intelligence and machine learning models to do inferencing on-device and less on cloud processing[4].

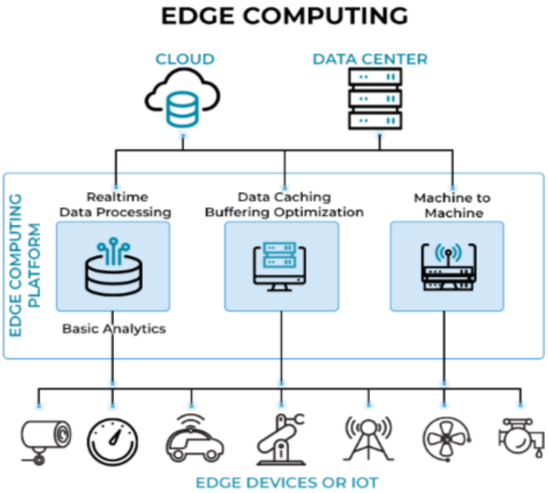

Figure 1 Edge Computing [5]

This trend has been sparked by the development of specialised hardware enhancing computational efficiency—edge accelerators and TPUs. Since they offer scalable and light-weight deployment options that fit evolving application requirements, serverless computing have become increasingly crucial in edge systems. Blockchain technology is also rewriting data integrity and security in edge networks by means of distributed trust systems and safe data exchanges. Particularly with edge devices with limited resources, main design concerns for these systems are dependability, energy economy, and interoperability[6]–[8]. Using hybrid edge-cloud models lets non-essential tasks be handled in the cloud while crucial calculations take place close to the edge helps to control workloads. Moreover customised to manage containerised workloads in edge environments, Kubernetes ensures fault tolerance by means of efficient resource allocation and fault tolerance. By allowing communication among heterogeneous devices, MQTT, CoAP, OPC UA among other standards and protocols serve to assure easy interoperability in many edge environments. As these architectures evolve, data consistency, security in widely dispersed networks, and fulfilling the different QoS needs of apps must all be managed. Cooperative research and development between academia and industry is producing edge computing ecosystems including energy-efficient algorithms, edge-native application development frameworks, and real-time analytics platforms. Emerging designs in edge computing moving forward will specify further automation, AI-driven resource management, and the convergence of technologies like 6G, quantum computing, and federated learning, so opening the route for a highly linked and intelligent world. These advances will expand the possibilities of edge computing, hence enabling next-generation applications needing exceptional efficiency and responsiveness[9], [10].

II. Literature Review

Diro 2024 With smart devices linking effortlessly to the Internet, technological developments in wireless communications and electronics have sped the expansion of the Internet of Things (IoT). Traditional networks struggle, meanwhile, in controlling these devices in various ways: scalability, real-time data transmission, programmability, and mobility. Integration of Software Defined Networking (SDN) with Fog computing has become clear as a workable answer to these problems. While fog computing analyses data at the network edge, guaranteeing minimal latency, SDN offers a centralised network control plane for traffic management and resource allocation. Notwithstanding the advantages, flow entry installation delays and constraints in conventional data and control space architectures cause latency overheads to remain. This work presents a convergent SDN and Fog architecture with differential flow space allocation for heterogeneous IoT applications, so boosting the efficiency of key flow classes while preserving fairness for non-critical ones and so mitigating these[11].

Makondo 2024 Ultra-Reliable Low-Latency Communications (uRLLCs) presented by Makondo 2024 5G technology support uses requiring ultra-low latency and great dependability. For conventional cloud-based systems, which struggle to satisfy the real-time needs of uRLLCs, reaching these strict latency criteria is a difficulty. Edge computing presents a good answer in this situation since it brings computational resources closer to end users. This work suggests a latency-optimized design using the control and user plane separation (CUPS) approach to progressively bring the user plane function (UPF) closer to users thereby lowering latency. Furthermore using Software Defined Networking (SDN) in the rear haul network improves dynamic scalability and efficiency, therefore allowing operators to provide a range of services. Experimental results from a 5G testbed show that this design lowers latency by 60% and raises average throughput by 40%, therefore confirming its efficiency in satisfying the latency and performance criteria of 5G networks[12].

Golpayegani 2024 To lower latency, improve quality of service, and enable mobility in dynamic contexts, edge computing brings processing resources nearer to end users. Edge-enabled systems running on limited resources depend on their capacity to adapt and self-organise in response to changing conditions. Reviewing current research on adaptive edge computing, this work uses a well accepted taxonomy to classify adaption techniques. It covers several facets of adaptation, including the causes of it, the degrees of it, and the methods of control. The article also points to research shortcomings, namely the dearth of proactive adaptation techniques and an emphasis on middleware, communication infrastructure, and context-based adjustments. [13].

Dazzi 2024 Especially for ultra-low latency, availability, resource management, decentralisation, self-organization, and security, Urgent Edge Computing (UEC) is a novel paradigm addressing the vital needs of time-sensitive applications in distributed edge environments. Designed to run in mission-critical events where speed is of the utmost, UEC is meant to operate in disaster response, environmental monitoring, and smart city management This work presents a conceptual architecture to serve the main use cases of UEC, therefore addressing the necessary criteria and challenges of this technology. To guarantee high availability and low latency for critical applications, the writers stress the need of strong security and effective resource management especially in distributed and self-organising edge contexts. Examined are possible uses of UEC that provide understanding of how this paradigm can affect edge computing for mission-critical applications going forward[14].

Nair 2024 Particularly for real-time applications needing minimal latency, the move from conventional cloud computing to edge computing represents a major change in data management. Combined with 5G and IoT, edge computing provides the basis of fog computing—which allows local data processing nearer the source. This solves problems with latency, data volume, security, and diversity, hence lessening dependency on central cloud data centres. Driven by rising edge device deployment in many different sectors, the edge computing market is expected to expand dramatically by 2026. Particularly in areas such health data monitoring, predictive maintenance, smart grids, and autonomous vehicles, these devices provide real-time data collecting, processing, and decision-making. [15].

Table.1 Literature Summary

Authors	**Methodology**	**Research gap**	**Findings**
Hartmann 2022[16]	surveys edge computing architectures and techniques for healthcare applications, focusing on classification	Previous surveys overlooked optimal computing techniques such as encryption,	Identifies challenges in privacy, data reduction, and computational complexity while emphasizing

	of health data like vital signs and fall detection	authentication, and classification in edge devices for healthcare.	future directions for improving healthcare quality via edge computing.
Liang 2022 [17]	The paper reviews Multi-access Edge Computing (MEC) advancements, exploring its framework, capabilities, and enabling technologies like SDN and NFV.	The article addresses gaps in MEC’s low-latency, high-bandwidth application support but does not fully explore scalability in real-world scenarios.	Summarizes MEC technologies and identifies key research challenges to optimize throughput, reliability, and latency for innovative applications.
Wang 2022 [18]	explores MEC for the Metaverse, focusing on technologies such as 3D rendering, AI, and Blockchain to enhance virtual reality	Limited attention to overcoming energy, latency, and storage issues that hinder the Metaverse’s development in real-time applications	Proposes MEC as a solution to improve performance by offloading tasks to the network edge, mitigating system challenges in Metaverse environments
Zhang 2022 [19]	Edge-as-a-Service (EaaS) to manage large-scale distributed intelligence for applications requiring low latency and high connectivity	Existing edge computing approaches are insufficient for meeting ultra-low latency and dynamic service needs in emerging applications	Proposes EaaS to enable flexible, large-scale deployment of edge services and addresses challenges like resource elasticity and edge autonomy in real-time applications.
Beylot 2022 [20]	the application of Multi-access Edge Computing (MEC) in vehicular networks (IoV), focusing on network availability, reliability, and task offloading	Previous studies have not sufficiently addressed the delay-sensitive nature of vehicular networks and the impact of offloading on real-time applications	MEC implementations for vehicular networks, highlighting improvements in communication, computation delays, and challenges

	performance		

III. Edge Computing And Its Low Latency Requirements

Edge computing is a transformative paradigm in modern computer architectures, meant to satisfy the rising demand for real-time processing, low-latency communication, and efficient resource use. Unlike traditional cloud computing, which concentrates data processing in remote data centres, edge computing decentralises computation, storage, and analytics by placing these capabilities near the data source, such as IoT devices, sensors, and user endpoints. Faster response times—which are critical for applications needing low latency—are made possible by this closeness helping to reduce the reliance on long-distance communication to central servers.

[21]. Edge computing has emerged from the explosive growth of IoT devices, the spread of data-intensive apps, and changes in networking technology such as 5G. The latency caused by data transfer over vast distances offers significant challenges with conventional centralised cloud systems for real-time applications such driverless automobiles, augmented reality (AR), virtual reality (VR), telemedicine, and industrial automation. Edge computing lowers round-trip time (RTT) required for data to flow between devices and data centres by processing data locally or regionally, therefore solving these issues. By filtering, aggregating, and analysing data at the edge before forwarding just the required information to the cloud for additional processing or storage, edge computing also maximises bandwidth use. This improves the general efficiency of data-intensive systems in addition to lightening the load on cloud infrastructure[22]–[24].

Low latency in edge computing is achieved through:

- **Proximity to Data Sources**: By deploying edge nodes close to the devices generating data, the physical distance for data travel is reduced, which minimizes latency.
- **Localized Processing**: Data is processed on edge devices or nearby edge servers, eliminating the need to transmit large volumes of raw data to distant cloud data centers.
- **Efficient Communication Protocols**: High-speed networking technologies, such as 5G, Wi-Fi 6, and software-defined networking (SDN), enable faster and more reliable communication.

IV. Importance Of Low Latency Applications

Low-latency applications are fundamental in modern computing settings, particularly in influence usability, user experience, or safety. Given the proliferation of real-time systems in fields such healthcare, autonomous systems, gaming, and financial services, minimising delays becomes increasingly important. Three main elements contribute to sum up the benefit of low-latency apps[25]:

Figure 2 Low Latency Applications[26]

A. Enhancing Real-Time Decision-Making and User Experiences

Low-latency applications process and react to data almost quickly, therefore enabling computers to make real-time choices. Low latency guarantees that important decisions, such obstacle avoidance or route optimisation, are performed within milliseconds in applications like autonomous cars, therefore greatly lowering the possibility of accidents. Analogous low latency is required in telemedicine systems to provide real-time communication and diagnostics between patients and healthcare practitioners, hence enabling potentially life-saving emergency response. Regarding user experience, reduced latency is exactly correlated with involvement and happiness. For gaming and virtual reality (VR), for example, even little delays can compromise the immersive experience and cause consumer unhappiness. Low-latency systems enable high-frequency trading by allowing fast execution of buy or sell orders, therefore helping financial services, where milliseconds might be the difference between profit and loss. Low-latency applications define the benchmark for performance in many real-time contexts by guaranteeing instantaneous response and smooth interactivity[27].

B. Driving Efficiency in Industrial and Mission-Critical Systems

Maximising industrial automation, smart factories, and mission-critical systems requires low-latency applications. Robotics and control systems used in industrial automation depend on instantaneous feedback to preserve accuracy and efficiency in manufacturing operations. Robotic arms, for instance, have to react instantly to sensor data in a production line to guarantee accurate completion of jobs including welding, assembling, and packaging free from delay. Likewise, low-latency communication is essential for mission-critical systems including air traffic control and emergency response to stop disastrous results. Any delay in data processing or transmission in these surroundings might cause problems compromising infrastructure or life. Low latency helps especially smart grids and energy management systems since they depend on real-time monitoring and control to effectively balance energy supply and demand. Low-latency applications reduce delays, therefore improving the dependability and performance of these vital systems and guaranteeing best operations and resilience[28].

C. Enabling Innovation in Emerging Technologies

Emerging technologies such 5G, edge computing, and AI-powered systems are built on low-latency applications and used mostly in this regard. Real-time processing and data interchange are what these technologies excel at, hence low latency is absolutely necessary for their success. Low latency, for example, lets devices in edge computing analyse data locally or regionally, therefore lowering the demand for long-distance transmission and supporting uses including smart cities, autonomous drones, and remote monitoring. Low-latency systems are fundamental for enabling real-time inference and decision-making in advancing artificial intelligence and machine learning. AI-powered diagnostic technologies, for instance, examine medical images and offer instantaneous insights in healthcare, therefore enhancing the speed and accuracy of therapies. Low-latency artificial intelligence technologies improve dynamic pricing techniques and tailored recommendations in the retail sector, therefore increasing consumer interaction and operational effectiveness. Adoption of 5G networks, with ultra-reliable low-latency communication (URLLC), increases the possibility for innovation even more. The low-latency features of 5G enable applications include connected cars, remote robotic surgery, and augmented reality (AR) overlays for repair professionals, hence stretching the bounds of what is theoretically achievable[29].

The importance of low-latency applications lies in their ability to enable real-time decision-making, improve efficiency in industrial and mission-critical systems, and foster innovation in emerging technologies. By minimizing delays, these applications enhance safety, reliability, and user experiences across diverse domains. As industries continue to adopt advanced technologies like edge computing, AI, and 5G, the demand for low-latency applications will only grow, solidifying their role as a foundational element of modern computing systems[30].

V. Role Of 5g, Iot, And Ai In Edge Computing

Emerging as a fundamental technology allowing real-time, low-latency applications is edge computing. By means of high-speed connectivity, ubiquitous data generation, and intelligent decision-making capacity, the synergistic combination of 5G, IoT, and AI magnifies the potential of edge computing. One might sum their roles as follows.

- 5G: High-Speed, Low-Latency Connectivity

The arrival of 5G networks has changed edge computing by offering ultra-reliable low-latency communication (URLLC) and high-speed data transport. Five Gbps data rates and latency as low as one millisecond let 5G enable edge devices to process and transfer data nearly instantly. Applications sensitive to latency such remote robotic surgery, augmented reality (AR), and autonomous autos rely on this ability. 5G's network slicing feature ensures consistent performance under various network demands and allows dedicated virtual networks for certain use cases, hence improving resource allocation for edge computing operations[31].

- IoT: Pervasive Data Generation at the Edge

Edge computing is built on the Internet of Things (IoT), which generates massive amounts of data at the edge of networks. IoT devices including sensors, cameras, and connected appliances deliver real-time data to edge computing platforms from data sources. Edge computing reduces bandwidth use via local data processing and provides faster reaction times, hence decreasing of the demand for transmission to centralised cloud servers. Applications include smart cities, industrial automation, and precision farming largely rely on this IoT-edge connection to acquire real-time data and operational efficiency[32].

- AI: Intelligent Decision-Making at the Edge

By means of real-time analytics and predictive modelling, artificial intelligence (AI) enhances edge computing, therefore enabling localised decision-making. Edge AI models—which analyse data directly on devices or edge nodes—minimize latency and increase privacy by reducing data transfer. AI-powered cameras, for example, may search video feeds to find objects or notice anomalies without needing cloud-based computing. Applications depend on this ability in predictive maintenance, security monitoring, and healthcare diagnostics[33].

VI. Architectural Paradigms In Edge Computing

Edge computing models have developed to meet low-latency, real-time data processing, scalability, energy economy, and demand. These paradigms are meant to distribute computing, storage, and network resources so that localised processing may be enabled and need on centralised data centres may be reduced[34]. The key architectural paradigms in edge computing can be categorized as follows:

A. Layered Architecture: Distributed Edge Hierarchy

Edge computing is split by layered architecture into several tiers—cloud, edge servers, end devices, each with unique purposes in data processing.

- Cloud Layer: Acts as a central centre for training machine learning models, high-capacity analytics, and long-term data storage.
- Edge Server Layer: Edge servers, positioned nearer the data source, minimise latency by processing, filtering, and storing important data. Often found at network base stations, data centres, or gateways, these servers guarantee fast response times for applications sensitive to latency.
- Device Layer: Comprising IoT devices, sensors, and actuators, this layer captures and sends raw data to edge servers or local processing.

This hierarchical model ensures an efficient division of computational tasks while optimizing resource allocation and reducing network congestion. For instance, a smart factory might rely on local edge nodes for immediate decision-making while sending aggregated data to the cloud for advanced analytics[35].

B. Microservices-Based Architecture: Modular and Scalable

Edge computing microservices-based architecture has a modular design whereby separate processes or services run independently. Edge nodes can be used for these light-weight, containerised services to handle certain tasks as analytics, data preparation, or artificial intelligence inference.

- Benefits of Modularity: One can upgrade, scale, or replace particular services without compromising the whole system.
- Scalability: Dynamic scaling depending on workload made possible by microservices lets the design fit changing needs.

In a smart city application, for instance, traffic monitoring systems might scale up during peak hours while other services—such as waste management—may remain normal. Using tools like Kubernetes for orchestration guarantees flawless deployment and administration of edge workloads[36].

C. Serverless Computing: Event-Driven Processing

Under serverless computing, developers create apps as functions that run in response to particular events, therefore removing the necessity for controlling underlying infrastructure.

- Edge Functions: Events like sensor readings or user interactions set off these lightweight features, allowing real-time processing with minimum latency.
- Benefits: Since functions consume compute power only when run, serverless computing maximises resource use. Abstracting hardware management helps to simplify development and lower costs as well.

Serverless architectures are quite successful in edge computing for dynamic, short-lived tasks like anomaly detection in IoT systems or tailored suggestions in retail. This paradigm's event-driven character fits very nicely the low-latency requirements of edge applications. Edge computing's architectural paradigms—layered hierarchy, microservices, serverless computing—offer the fundamental frameworks for enabling scalable, low-latency, efficient solutions. These paradigms are designed to satisfy the many needs of applications ranging from industrial automation to smart cities, therefore enabling edge computing to meet the challenges of contemporary technical requirements[37].

VII. Conclusion

Edge computing's emergent architectures mark a radical change in the design and execution of systems giving low latency, efficiency, and scalability top priority. Decentralising computing and reducing dependency on centralised cloud resources helps these architectures provide creative answers as the demand for real-time processing rises across sectors including healthcare, autonomous systems, smart cities, and industrial IoT. Edge computing has been greatly enabled by the integration of cutting-edge technologies such 5G, IoT, and artificial intelligence in order to satisfy low-latency application strict criteria. Layers of hierarchy, microservices, and serverless computing guarantee that systems are not just responsive but also flexible, scalable, and reasonably priced architectural paradigms. Edge computing systems lower network congestion, increase privacy, and increase resilience against failures by processing data nearer the source. Edge computing has great potential, but attaining ideal resource management, preserving interoperability across many devices and platforms, and guaranteeing strong security measures in distributed contexts remain difficult. Dealing with these difficulties will call for ongoing research and invention in fields such secure communication protocols, energy-efficient hardware, and AI-driven task. Edge computing's future resides in its capacity to change with newly developing technologies such federated learning, quantum computing, and next-generation networking options. These developments will improve the capacity of edge architectures even further, opening fresh opportunities for ultra-low-latency applications and enabling sectors to rethink operational efficiencies and user experiences. Edge computing's emergent designs lead front stage in a technology revolution ready to revolutionise data processing and application. Stakeholders that embrace these developments can fully utilise edge computing to satisfy the always increasing needs of contemporary, real-time applications.

References

1. H. Zhang, Y. Yang, X. Huang, C. Fang, and P. Zhang, “Ultra-Low Latency Multi-Task Offloading in

Mobile Edge Computing," *IEEE Access*, vol. 9, pp. 32569–32581, 2021, doi: 10.1109/ACCESS.2021.3061105.
2. C. Li, Y. Zhang, R. Xie, X. Hao, and T. Huang, "Integrating Edge Computing into Low Earth Orbit Satellite Networks: Architecture and Prototype," *IEEE Access*, vol. 9, pp. 39126–39137, 2021, doi: 10.1109/ACCESS.2021.3064397.
3. M. Kumar, K. Dubey, and R. Pandey, "Evolution of emerging computing paradigm cloud to fog: Applications, limitations and research challenges," *Proc. Conflu. 2021 11th Int. Conf. Cloud Comput. Data Sci. Eng.*, no. January, pp. 257–261, 2021, doi: 10.1109/Confluence51648.2021.9377050.
4. M. Rohith, A. Sunil, and Mohana, "Comparative Analysis of Edge Computing and Edge Devices: Key Technology in IoT and Computer Vision Applications," *2021 6th Int. Conf. Recent Trends Electron. Information, Commun. Technol. RTEICT 2021*, no. May, pp. 722–727, 2021, doi: 10.1109/RTEICT52294.2021.9573996.
5. M. J. P. Peixoto *et al.*, "Edge Intelligence: Paving the Last Mile of Artificial Intelligence With Edge Computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1778–1837, 2021.
6. J. Hu, C. Chen, L. Cai, M. R. Khosravi, Q. Pei, and S. Wan, "UAV-Assisted Vehicular Edge Computing for the 6G Internet of Vehicles: Architecture, Intelligence, and Challenges," *IEEE Commun. Stand. Mag.*, vol. 5, no. 2, pp. 12–18, 2021, doi: 10.1109/MCOMSTD.001.2000017.
7. D. Borsatti, G. Davoli, W. Cerroni, and C. Raffaelli, "Enabling Industrial IoT as a Service with Multi-Access Edge Computing," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 21–27, 2021, doi: 10.1109/MCOM.001.2100006.
8. I. Kovacevic, E. Harjula, S. Glisic, B. Lorenzo, and M. Ylianttila, "Cloud and Edge Computation Offloading for Latency Limited Services," *IEEE Access*, vol. 9, pp. 55764–55776, 2021, doi: 10.1109/ACCESS.2021.3071848.
9. Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A Survey on Mobile Augmented Reality with 5G Mobile Edge Computing: Architectures, Applications, and Technical Aspects," *IEEE Commun. Surv. Tutorials*, vol. 23, no. 2, pp. 1160–1192, 2021, doi: 10.1109/COMST.2021.3061981.
10. X. Chen *et al.*, "Achieving low tail-latency and high scalability for serializable transactions in edge computing," *EuroSys 2021 - Proc. 16th Eur. Conf. Comput. Syst.*, pp. 210–227, 2021, doi: 10.1145/3447786.3456238.
11. A. A. Diro, H. T. Reda, and N. Chilamkurti, "Differential flow space allocation scheme in SDN based fog computing for IoT applications," *J. Ambient Intell. Humaniz. Comput.*, vol. 15, no. 2, pp. 1353–1363, 2024, doi: 10.1007/s12652-017-0677-z.
12. N. Makondo, H. I. Kobo, T. E. Mathonsi, D. Du Plessis, T. M. Makhosa, and L. Mamushiane, "An Efficient Architecture for Latency Optimisation in 5G Using Edge Computing for uRLLC Use Cases," *7th Int. Conf. Artif. Intell. Big Data, Comput. Data Commun. Syst. icABCD 2024 - Proc.*, no. August, pp. 1–7, 2024, doi: 10.1109/icABCD62167.2024.10645277.
13. F. Golpayegani *et al.*, "Adaptation in Edge Computing: A review on design principles and research challenges," *ACM Trans. Auton. Adapt. Syst.*, vol. 19, no. 3, 2024, doi: 10.1145/3664200.
14. P. Dazzi *et al.*, "Urgent Edge Computing," *Fram. 2024 - Proc. 2024 Work. Flex. Resour. Appl. Manag. Edge, Part HPDC 2024 - 3rd Int. Symp. High-Performance Parallel Distrib. Comput.*, pp. 7–14, 2024, doi: 10.1145/3659994.3660315.
15. S. S. Nair, "Beyond the Cloud -Unraveling the Benefits of Edge Computing in Iot," *Int. J. Comput. Eng. Technol.*, vol. 14, no. 3, pp. 91–97, 2023.
16. M. Hartmann, U. S. Hashmi, and A. Imran, "Edge computing in smart health care systems: Review, challenges, and research directions," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 3, 2022, doi: 10.1002/ett.3710.
17. B. Liang, M. A. Gregory, and S. Li, "Multi-access Edge Computing fundamentals, services, enablers and challenges: A complete survey," *J. Netw. Comput. Appl.*, vol. 199, no. January, 2022, doi: 10.1016/j.jnca.2021.103308.
18. Y. Wang and J. Zhao, "A Survey of Mobile Edge Computing for the Metaverse: Architectures, Applications, and Challenges," *Proc. - 2022 IEEE 8th Int. Conf. Collab. Internet Comput. CIC 2022*, no. Cic, pp. 1–9, 2022, doi: 10.1109/CIC56439.2022.00011.

19. M. Zhang, J. Cao, Y. Sahni, Q. Chen, S. Jiang, and T. Wu, "EaaS: A Service-Oriented Edge Computing Framework Towards Distributed Intelligence," *Proc. - 16th IEEE Int. Conf. Serv. Syst. Eng. SOSE 2022*, pp. 165–175, 2022, doi: 10.1109/SOSE55356.2022.00026.
20. L. Bréhon–Grataloup, R. Kacimi, and A. L. Beylot, "Mobile edge computing for V2X architectures and applications: A survey," *Comput. Networks*, vol. 206, no. Mcc, 2022, doi: 10.1016/j.comnet.2022.108797.
21. R. Dave, N. Seliya, and N. Siddiqui, "The Benefits of Edge Computing in Healthcare, Smart Cities, and IoT," *J. Comput. Sci. Appl.*, vol. 9, no. 1, pp. 23–34, 2021, doi: 10.12691/jcsa-9-1-3.
22. L. Huang, A. Chandra, and J. Weissman, *Armada: A Robust Latency-Sensitive Edge Cloud in Heterogeneous Edge-Dense Environments*, vol. 1, no. 1. Association for Computing Machinery, 2021.
23. M. Ke *et al.*, "An Edge Computing Paradigm for Massive IoT Connectivity over High-Altitude Platform Networks," *IEEE Wirel. Commun.*, vol. 28, no. 5, pp. 102–109, 2021, doi: 10.1109/MWC.221.2100092.
24. S. C. Lin, K. C. Chen, and A. Karimoddini, "SDVEC: Software-Defined Vehicular Edge Computing with Ultra-Low Latency," *IEEE Commun. Mag.*, vol. 59, no. 12, pp. 66–72, 2021, doi: 10.1109/MCOM.004.2001124.
25. H. Rahimi, Y. Picaud, K. D. Singh, G. Madhusudan, S. Costanzo, and O. Boissier, "Design and Simulation of a Hybrid Architecture for Edge Computing in 5G and beyond," *IEEE Trans. Comput.*, vol. 70, no. 8, pp. 1213–1224, 2021, doi: 10.1109/TC.2021.3066579.
26. C. Cicconetti, M. Conti, and A. Passarella, "Architecture and performance evaluation of distributed computation offloading in edge computing," *Simul. Model. Pract. Theory*, vol. 101, 2020, doi: 10.1016/j.simpat.2019.102007.
27. L. Liu, C. Chen, Q. Pei, S. Maharjan, and Y. Zhang, "Vehicular Edge Computing and Networking: A Survey," *Mob. Networks Appl.*, vol. 26, no. 3, pp. 1145–1168, 2021, doi: 10.1007/s11036-020-01624-1.
28. L. Yang, X. Chen, S. M. Perlaza, and J. Zhang, "Special Issue on Artificial-Intelligence-Powered Edge Computing for Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9224–9226, 2020, doi: 10.1109/JIOT.2020.3019948.
29. S. D. A. Shah, M. A. Gregory, S. Li, and R. D. R. Fontes, "SDN enhanced multi-access edge computing (MEC) for E2E mobility and QoS management," *IEEE Access*, vol. 8, pp. 77459–77469, 2020, doi: 10.1109/ACCESS.2020.2990292.
30. G. Caiza, M. Saeteros, W. Oñate, and M. V. Garcia, "Fog computing at industrial level, architecture, latency, energy, and security: A review," *Heliyon*, vol. 6, no. 4, p. e03706, 2020, doi: 10.1016/j.heliyon.2020.e03706.
31. I. Pelle, F. Paolucci, B. Sonkoly, and F. Cugini, "Telemetry-driven optical 5G serverless architecture for latency-sensitive edge computing," *Opt. InfoBase Conf. Pap.*, vol. Part F174-OFC 2020, no. January, pp. 1–4, 2020, doi: 10.1364/ofc.2020.m1a.1.
32. S. Das, F. Slyne, A. Kaszubowska, and M. Ruffini, "Virtualized EAST-WEST PON architecture supporting low-latency communication for mobile functional split based on multiaccess edge computing," *J. Opt. Commun. Netw.*, vol. 12, no. 10, pp. D109–D119, 2020, doi: 10.1364/JOCN.391929.
33. C. S. M. Babou *et al.*, "Hierarchical Load Balancing and Clustering Technique for Home Edge Computing," *IEEE Access*, vol. 8, pp. 127593–127607, 2020, doi: 10.1109/ACCESS.2020.3007944.
34. F. Wang, M. Zhang, X. Wang, X. Ma, and J. Liu, "Deep Learning for Edge Computing Applications: A State-of-the-Art Survey," *IEEE Access*, vol. 8, pp. 58322–58336, 2020, doi: 10.1109/ACCESS.2020.2982411.
35. S. Baidya, Y. J. Ku, H. Zhao, J. Zhao, and S. Dey, "Vehicular and edge computing for emerging connected and autonomous vehicle applications," *Proc. - Des. Autom. Conf.*, vol. 2020-July, 2020, doi: 10.1109/DAC18072.2020.9218618.
36. F. Vhora and J. Gandhi, "A Comprehensive Survey on Mobile Edge Computing: Challenges, Tools, Applications," *Proc. 4th Int. Conf. Comput. Methodol. Commun. ICCMC 2020*, no. March, pp. 49–55, 2020, doi: 10.1109/ICCMC48092.2020.ICCMC-0009.
37. S. Al Harbi, T. Halabi, and M. Bellaiche, "Fog Computing Security Assessment for Device

Authentication in the Internet of Things," *Proc. - 2020 IEEE 22nd Int. Conf. High Perform. Comput. Commun. IEEE 18th Int. Conf. Smart City IEEE 6th Int. Conf. Data Sci. Syst. HPCC-SmartCity-DSS 2020*, no. December, pp. 1219–1224, 2020, doi: 10.1109/HPCC-SmartCity-DSS50907.2020.00202.

.