

Edge Computing: Enhancing IoT Performance through Decentralized Data Processing and Real-Time Decision-Making

Tamasa priyadarsini¹ Binayak Sahoo²

¹Dept. of CSE, GIET Ghangapatna, ²Dept. of CSE MITM Bhubaneswar

¹tamasa.priyadarsini@gmail.com, ²sahoobinayak@gmail.com

Abstract:

The growing proliferation of Internet of Things (IoT) devices has led to the generation of massive volumes of data, which are traditionally transferred to centralized cloud data centers for storage and processing. However, this approach often leads to network performance challenges, including increased latency, bandwidth consumption, and migration costs. Edge computing, a decentralized architectural strategy, has emerged as a solution to these issues by placing computation and storage resources closer to the end users and devices, minimizing data travel and optimizing network efficiency. This paper explores the role of edge computing in mitigating the limitations of cloud-based solutions, highlighting its benefits such as reduced latency, bandwidth optimization, enhanced privacy and security, real-time decision-making, scalability, and improved application performance. Key applications of edge computing in areas such as traffic control, content caching, healthcare, and autonomous driving are discussed, demonstrating its ability to support real-time, compute-intensive applications. Additionally, the architecture of edge computing, including cloudlets, multi-access edge computing (MEC), and the integration with 5G networks, is outlined as essential to enabling future advancements in IoT and real-time systems.

Keywords: IoT, Cloud Computing, Edge Computing.

I. Introduction

Massive volumes of data have been produced in recent years by the Internet of Things (IoT). Large amounts of data are being transferred to the cloud by IoT devices [1]. Data from Internet of Things devices is stored in a sizable data centre on the cloud. IoT gadgets are intelligent gadgets. The cloud is unable to react when all IoT devices submit massive amounts of data to it. The network's performance is impacted by excessive latency, which can be caused by overloading the cloud or heavy traffic, and it can also result in increased migration costs. Thus, edge computing offers a solution to this issue. "Technological developments that enable computation at the edge network are referred to as edge computing. It includes processing upstream data for IoT services and downstream data for cloud services [2].

The phrase "edge computing" refers to a group of architectural strategies that aim to place storage and processing resources in close proximity to end users and devices [4]. IoT devices transmit data to edge devices in edge computing, which receives all of the data from IoT devices and transmits critical data to cloud computing [1]. By lowering latency and bandwidth

consumption, improving connectivity, delivering services quickly, scalability, speeding up computation and storage, and enhancing overall application performance, edge computing seeks to maximise network performance while saving money on bandwidth. The goal of edge computing is to allow real-time compute-intensive applications to be run directly at the network edge by billions of linked edge devices. Edge computing stands out for its proximity to end users, support for mobility, and dense geographic deployment [3]. However, because of the physical distance separating users and the centralised cloud infrastructure, it cannot support real-time Internet of Things applications. Edge computing, as defined by Cisco, is a concept that moves processing power away from public clouds or central data centres and onto devices that are embedded at the edge of service provider networks [4]. The processing happens at the network edge rather than sending the data to centralised cloud infrastructures. A novel architectural element known as cloudlets was designed to meet the low-latency requirements of resource-intensive applications [5]. Edge computing was identified in 2018 as one of the key technologies that will propel future enterprises,

especially those connected to IoT networks where obtaining fast responses is a key goal [6].

The remainder of the paper's sections are organised as follows: In Section II, we described the Advantages and Requirements of edge computing, while Section III covers the Purpose and Uses of Edge Computing. Section IV describes the Architecture of Edge Computing with a diagram. In Section V, we have discussed the 5G network-enabled mobile edge computing architecture and in Section V we summarized Applications of Edge Computing. Finally, concludes the paper in Section VI.

II Advantages and Requirements:

Edge computing is a decentralised computing strategy that locates storage and processing power close to the source of data generation, such as the devices or the network edge, where it is needed. The following are some main justifications for edge computing's requirements and advantages in a variety of fields:

II.1 Latency: Reducing latency is the main reason to put an application at the network edge [7] [8]. The goal of edge computing is to reduce the amount of time it takes for data to travel between devices and centralised cloud servers.

II.2 Efficient use of Bandwidth: A higher bandwidth can reduce transmission time from the standpoint of latency, especially for big amounts of data like video. Edge computing minimises the need to send massive amounts of data to centralised cloud servers, which maximises the use of available network bandwidth. When dealing with situations where connectivity is scarce, such remote or Internet of Things devices with restricted network capabilities, this optimisation of network capacity is especially helpful.

II.3 Improved Privacy and Security of Data: Converting to edge computing from cloud-based paradigms improves privacy and lowers security risks [9]. By allowing data to be handled locally, edge computing keeps sensitive data contained within the edge network and lessens its exposure to the cloud or other networks. Effectively managing a sizable amount of varied edge devices presents issues for trust-based classical security procedures like policy and SLA verification [10]. As a result, using a security method like white-listing features that makes use of fewer authentication principles

may be a workable approach [11]. Therefore, edge computing helps allay worries about security and privacy by storing sensitive data near to where it originates rather than sending it to distant cloud servers.

II.4 The instantaneous Decision-Making: Real-time or nearly real-time decision-making is required for some applications. Decisions can be made independently without depending on sending data to the cloud and having it returned for analysis by processing data at the edge. Rapid analysis and prompt action based on data insights are essential for operational efficiency and competitive advantage in applications like autonomous vehicles, smart grids, and industrial automation, where quick responses are required for optimal performance and safety [12].

II.5 Distributed Architecture and Scalability: With edge computing, computing power is dispersed among edge devices, gateways, and local servers in a scalable architecture. Adding more devices to the edge network permits horizontal scaling and the effective distribution of resources. Additionally, it lessens the strain on centralised cloud servers, strengthening the system's overall resilience and capacity to manage extensive deployments.

III. The Purpose and Uses of Edge Computing

The primary causes of the enormous wide-area latency experienced while utilising cloud services and the notable rise in compute-intensive, delay-sensitive applications like virtual reality at the beginning of MEC. The rationale for implementing MEC: For other examples of MEC, consider the different application situations in various spheres of existence.

III.1 Traffic Control

There are several uses for wireless sensor networks in cities [13, 14]. These sensor nodes could monitor traffic patterns and other aspects of the urban environment. Imagine a traffic management scenario where a resource-rich platform is needed to analyse traffic-related data, such as accidents and traffic congestion, and identify alternate routes. Because these duties include real-time computing and information dissemination, using a cloud server is not a practical choice due to the high latency of the vast area network. Consequently, the roadside devices' edge servers handle these computationally

demanding and time-sensitive activities. Drivers can receive the calculated data from the MEC server in time to choose better alternative routes.

III.2 Content Caching on the Edge

Mobile device users can benefit from improved web content access by enabling content caching on the base station. When a mobile user retrieves content from a web server, the content can be cached on a base station, which can shorten the time it takes for other users to access the same content later. In addition to enhancing customer satisfaction, edge content caching lowers network costs for content providers.

III.3 Distribution of local content

Imagine an alternative situation where a single mobile phone is used to stream live footage from a music concert to its users. The service provider may supplement the video with commercials and other pertinent information before the distribution of these contents. In this case, the service provider can add the information to the video stream by utilising the processing power of the MEC server installed on the Cell base station. Without sending the video to a different server for processing and sending it back to the base station, the information is added straight to the video on the base station. As a result, the processed video has a shorter transmission latency.

IV Architecture

Forecasts indicate that by 2025, there will be 77.44 billion IoT devices due to the growth of IoT applications [3]. Numerous IoT architectures have been developed by various organisations, considering diverse views, to accommodate the enormous number of IoT devices. It has been determined that edge computing is essential to the support of IoT systems. Figure 1 depicts the edge computing architecture.

IV.1 Cloud layer

It is made up of a sizable data centre that stores data that is passed through edge devices, the fog layer, and the edge layer of the edge computing architecture. Cloud data centres are large, highly accessible groups of virtualised resources that may be dynamically rearranged to meet different workload demands. Cloud services, which usually have a pay-as-you-go cost model, benefit greatly from this flexibility [8]. Infrastructure and space can be rented within a cloud data centre. These sizable

data centres' security and compliance are guaranteed by cloud providers.

IV.2 Fog layer

Fog computing is brought about by the fog layer. Big data, or the enormous amount of data produced by billions of sensors, presents difficulties when processing and transporting it entirely on the cloud. Moreover, some Internet of Things applications need processing that is quicker than what the cloud can offer right now [16]. The Fog computing paradigm is used to solve these problems. It makes use of nearby devices' idle computing resources by leveraging their processing capabilities.

IV.3 Server for Edge Gateways

A network device or server that serves as a bridge between the local network and the larger internet or cloud infrastructure is called an edge gateway server sometimes referred to as an edge server or an edge computing server. An edge gateway server's primary function is to optimise and secure traffic from the external network to the local network.

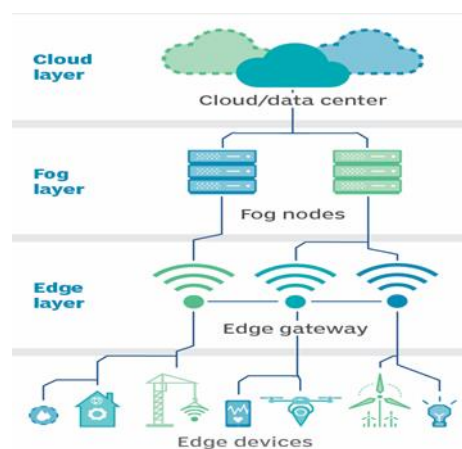


Fig. 1: The Architecture of Edge Computing

Cloudlet: In 2009, Satyanarayanan et al.[17] introduced the concept of cloudlets. According to the author, it is a reliable, resource-rich computer or cluster of computers that is accessible to mobile devices in the vicinity and has strong Internet connectivity, providing computation and storage capacity [18]. It is a small, regional data centre that increases the potential for cloud computing nearer the edge of the network. It is sometimes referred to as a small-scale cloud computing infrastructure or mini-cloud. Cloudlets, which offer low-latency computing resources and services, are usually

deployed in close proximity to end users or Internet of Things devices. They served as nodes in the middle that connected end devices and the cloud. It transfers some computation from the cloud to the edge, allowing for quicker processing and less network congestion.

An expansion of MCC, **multi-access edge** computing puts cloud-based resources and services close to users. One of FC's subsets, MEC offers virtualised MEC stands for multi-access edge computing, which extends beyond mobile devices and finds application in use cases like virtual reality (VR) and augmented reality (AR).

Multi-Access Edge Computing

Computer resources nearby the mobile devices with limited resources that have middling skills. These resources, known as MEC resources, are installed at a mobile location like a car or at a stationary place like a train station, stadium, or shopping centre. The servers for these resources are co-located with the base station. Ultra-low latency, much quicker bandwidth (up to 10 Gbps), and a more connected world are all made possible by the synergy between 5G and MEC. Many MEC models are the result of the development of Radio Access Network (RAN)2 architecture. RANs used all-in-one boxes with analogue, digital, and power functions for 1G and 2G cellular networks. These boxes were housed in a special room equipped with air conditioning and power supplies.

MEC makes use of the idea of edge servers, sometimes referred to as edge nodes, which are placed at the network's edge. Servers power a number of services and applications, including content distribution, augmented reality, real-time analytics, video streaming, and Internet of Things data processing. It also improves the effective use of network resources by moving some tasks from the edge to the central cloud. This tactic improves the effectiveness of the network as a whole and raises the standard of service for mobile users. Low latency, optimised bandwidth, improved user experience, scalability, and edge intelligence are the advantages of MEC.

IV. 5G network-enabled mobile edge computing architecture

One goal of smart solutions is coordinating 5G with new technologies for spectrum management. Emerging 5G communication technologies against the growing customer and organisation traffic. Future 5G designs and efficient network maintenance techniques will determine further organisation limitations. Once more, demands and dynamic management occur to further create network boundaries when they are truly required [18]. The European Telecommunication Standards Institute (ETSI) proposed the definition of systematic Mobile Edge Computing (MEC).The operators can use this cutting-edge technology to run their wireless communication edge over authorised third agents [19]. As was previously said, 5G network solutions can facilitate large device communication and high-speed internet access. Network Function Virtualisation (NFV) and Software Defined Networking (SDN), two of 5G's technology enablers, are used to set up virtualised network resources for constructing an entire optimised system for computing and storage applications. However, 5G functionalities have high latency, making them unsuitable for the newest IOT This integrated framework is capable of supporting cloud and edge computing tasks. The design assigns MEC computing assets to super 5G Base Stations (BS), LTE eNodeBs, 5G gNodeBs, and the edge of backbone networks for storage service provisioning and processing. This facilitates improvements in system capabilities and performance.

The most suitable conveyance frequency must be chosen in light of the internet of things' quick progress and the arrival of fifth-generation technologies. In the high frequency range, fifth generation networks can be used, however it's important to prevent problems with signal attenuation, faster and easier obstruction, and coverage distance.

V Applications of Edge Computing

There are a tonne of new and forthcoming gadgets that are online, all of which depend on technological advancements. Applications can frequently be connected with edge computing technologies through the IT service environment to

increase service quality, save operating costs, and use less energy.

Healthcare systems, entertainment and interactive media devices, virtual, augmented, and mixed reality technologies, tactile network communications, Ultra Reliable Low Latency Communications (URLLC), Internet of Things (IOT) applications, Intelligent Transportation Systems (ITS), and smart systems are just a few of the many services and applications that make full use of edge computing technology [20]. With the use of autonomous driving technology, cars may travel safely without a human driver. Because AI is used to apply this technology based on precise data that has been data, it would be quite advantageous to monitor and evaluate large-scale data. Edge computing is the ideal paradigm to accomplish quick computational procedures to manage the vehicle precisely. An example of an autonomous driving system is shown in Figure 3. Using edge computing. Information gathered from each sensor attached to an autonomous car is utilised to manage the various kinds of accidents that occur when operating a vehicle. This sensor gathers data by observing the surroundings on the ground traffic and the state of the roads. Instead of contacting the cloud server, the edge computing system handles processing such real-time tasks and intermediate answers [21]. Medical Applications: State-of-the-art high-performance computing, particularly for image merging in the medical technology domain. In the meantime, lowering the danger of data privacy on site, the precise medication may be implemented via the real-time monitoring system, and the early pathologies of the patient's medical data can be investigated. The vast quantity of data being processed at the edge and then only pertinent information being sent to the cloud could violate patient privacy due to this alteration and analysis of data. By restricting access for specific numbers of users, high-performance computing in the medical field along with the elimination of latency on cloud computing resources reduces the digital footprint of patient data.

VI Conclusion:

Edge computing represents a critical technological shift in the management and processing of data generated by IoT devices. By bringing computation

closer to the data source, edge computing overcomes the inherent limitations of centralized cloud infrastructures, reducing latency, bandwidth usage, and enhancing data privacy and security. It offers a scalable and flexible architecture, enabling real-time decision-making for mission-critical applications such as autonomous driving, healthcare, and smart cities. As IoT devices continue to grow exponentially, the integration of edge computing with emerging technologies like 5G and MEC will be pivotal in meeting the demands for low-latency, high-performance computing environments. This decentralized model is expected to play a key role in the future of IoT ecosystems, ensuring efficient, secure, and fast processing of vast amounts of data at the network edge.

References:

- [1] A.N. Lone, S. Mustajab, and M. Alam, "A comprehensive study on cybersecurity challenges and opportunities in the IoT world". *Security and Privacy*, e318, Apr. 2023
- [2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.
- [3] A. Ahmed, E. Ahmed, A survey on mobile edge computing, in: *Proceedings of 10th International Conference on Intelligent Systems and Control, ISCO'16, India, 2016.*
- [4] M. Satyanarayanan, "The emergence of edge computing," *Computer* vol. 50, no. pp. 30–39, 2017
- [5] Y. Ai, M. Peng, and K. Zhang, "Edge Computing Technologies for Internet of Things: A Primer," *Digit. Commun. Networks*, vol. 4, no. 2, pp. 77–86, Apr. 2018.
- [6] N. A. Sulieman, L. C. Ricciardi, W. Li, A. Zomaya, and M. Villari, "Edge-oriented computing: A survey on research and use cases," *Energies*, vol. 15, no. 2, pp.452, Jan. 2022.
- [7] M. Talebkhah, A. Sali, M. Marjani, M. Gordan, S.J. Hashim, and F.Z. Rokhani, "Edge computing: architecture, applications and future perspectives,". In *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAJET)* (pp. 1-6). Sep. 2020. IEEE.

- [8] H. Chang, A. Hari, S. Mukherjee, and T.V. Lakshman, "Bringing the cloud to the edge, in: 2014 IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPs, pp. 346–351, Apr. 2014.
- [9] Y. Mansouri, and M. A. Babar, "A review of edge computing: Features and resource virtualization," *Journal of Parallel and Distributed Computing* vol. 150, pp. 155–183, 2021.
- [10] I. Khalil, A. Khreishah, and M. Azeem, "Cloud computing security: A survey," *Computers*, vol. 3, no. 1, pp. 1–35, Feb. 2014.
- [11] C.S. Gates, N. Li, J. Chen, R.W. Proctor, Codeshield: towards personalized application whitelisting, in: R.H. Zakon (Ed.), 28th Annual Computer Security Applications Conference, ACSAC 2012, Orlando, FL, USA, 3–7 Dec. 2012, ACM, 2012, pp. 279–288.
- [12] Q. V. Pham, F. Fang, V.N. Ha, M.J. Piran, M. Le, L.B. Le, ... and Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE access*, vol. 8, pp. 116974-117017, Jun. 2020.
- [13] B. Rashid, M.H. Rehmani, Applications of wireless sensor networks for urban areas: a survey, *J. Netw. Comput. Appl.* 60 (2016) 192–219.
- [14] E. Ahmed, S. Ali, A. Akheenzada, I. Yaqoob, Cognitive radio sensor networks: bridging the gap for network, in: *Cognitive Radio Sensor Networks: Applications, Architectures, and Challenges: Applications, Architectures, and Challenges*. IGI Global, 2014, p. 160.
- [15] M. Alam, and M. I Siddiqui, "Effective framework to tackle urban unemployment by e-government: an IoT solution for smart/metro cities in developing nation," *Journal of Science Technology Policy Management*, vol. 14, no. 1, pp. 213-238, Feb. 2023.
- [16] C. Puliafito, E. Mingozzi, F. Longo, A. Puliafito, and O. Rana, "Fog computing for the Internet of Things: A survey," *ACM Trans. Internet Technol.* vol.19, no.2, pp. 18:1–18:41, 2019.
- [17] M. Satyanarayanan, P. Bahl, R. Cáceres, N. Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervasive Comput*, vol. 8, no. 4, pp. 14–23, Oct. 2009.
- [18] Thayanathan, V, Surekha Lanka (2019). An overview of 5G and Beyond chapter 1 (pp. 1-22) in *Advanced security issues of lot based 5G plus wireless communication for industry 4.0*. Nova Science Publishers.
- [19] K. Cao, Y. Liu, G. Meng and Q. Sun, "An Overview on Edge Computing Research," in *IEEE Access*, vol. 8, pp. 85714-85728, 2020, doi: 10.1109/ACCESS.2020.2991734.
- [20] Hassan, Najmul & Yau, Kok-Lim & Wu, Celimuge. (2019)" Edge Computing in 5G: A Review" . *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2019.2938534
- [21] Neural Network Algorithm." *Journal of Artificial Intelligence* 3, no. 01 (2021): 43-52.