

International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 12, Issue 4, July-August 2025

Impact Factor: 8.152



Edge vs. Cloud Computing: A Comparative Review of Architectures, Use Cases, and Future Directions

Nisarga N, Naveen Pawar, Moumita Roy

Department of MCA, CMR Institute of Technology, Bengaluru, India

Department of MCA, CMR Institute of Technology, Bengaluru, India

Assistant Professor, CMR Institute of Technology, Bengaluru, India

ABSTRACT: The availability of remote services in the form of cloud computing has brought a fundamental change in how data is stored, processed, and accessed. It has made these processes much more scalable, flexible, and economical. Regrettably, intelligent devices and automated systems that need data processing appears at an increasingly rapid pace. The necessity to process data in real-time with minimal delays rises exponentially. This latter need is addressed by edge computing which, need to process data nearer to its source, in scenarios where timecritical applications require quick responses to data processing. This review paper tackles the comparison between edge and cloud computing focusing on architecture, use cases, performance analysis and limitations. Critical parameters which are of concern include: bandwidth, lag periods, scalability, and the need for data security. Besides the paper seeks to address newer trends such as hybrid computing models and Edge AI which are enabled by the revolution in 5ger technology. This research study outlines resource allocation and infrastructure strategy with relevant literature and adoption data insights.

KEYWORDS: Edge and Cloud Computing Latency Bandwidth Scalability Hybrid Architecture Real-time Processing Edge AI 5G Networks.

I. INTRODUCTION

The rapid expansion of real-time applications and IoT devices has resulted in an unprecedented surge in data generation and processing requirements. This growth has necessitated more resilient and adaptive computing models. Historically, cloud computing emerged as the primary model for delivering IT services, offering scalable and on-demand computational resources via remote data centers [1]. While cloud infrastructures have consistently yielded excellent results in a broad spectrum of applications, they are often hindered by latency and network dependency in scenarios requiring instantaneous feedback and control.

Emerging application areas such as autonomous vehicles, industrial automation, smart surveillance systems, and remote healthcare demand extremely low latency, high bandwidth efficiency, and uninterrupted connectivity [2] [4]. In such time-sensitive environments, cloud-based systems may fall short due to inherent delays caused by data transmission to and from centralized servers. Edge computing overcomes these drawbacks by relocating processing data where it is collected, thereby enabling real-time responses, reducing bandwidth consumption, and improving operational reliability [3] [8].

Even though numerous studies have thoroughly investigated the distinct advantages of cloud and edge computing, a comprehensive comparative evaluation of both paradigms— particularly when considering hybrid systems, edge AI, and 5G integration—remains underdeveloped. Moreover, several open challenges such as interoperability, security, and cost-effectiveness in multi-layered architectures continue to be under-addressed in current literature [5] [7] [10]. It has become crucial to synthesize these gaps to guide the development of more resilient computing strategies.

This article adds to the body of literature by offering a balanced and critical contrast between cloud and edge computing. It summarizes their fundamental architecture, examines performance trade-offs, finds main benefits, and challenges, and explores application-specific use cases. In addition, it points out new trends like hybrid architecture and edge AI enabled by 5G networks, giving an understanding of how these technologies are being used in practice [6] [9].

By virtue of this comparative methodology, the research hopes to help researchers, system architects, and industry professionals align technological decisions with application requirements.

The remainder of this paper is structured as follows: We begin in Section II by describing the focus area of research, presenting the development, extent, and important concepts of cloud and edge paradigms of computing. In Section III, we address the relevant literature in a concise manner, presenting studies conducted so far, technological advancements, and areas of research gaps in the field. Section IV shows our comparative framework, with focus on architectural differences, performance factors, and recent developments like hybrid computing and edge AI. Lastly, Section V wraps up the paper with a summary of main findings and directions for future research and practical implementation.

II. RESEARCH AREA

Cloud computing has been a leading paradigm for providing computing offerings like data storage and computational resources, and software solutions through centralized data centers for many years. Its advantages include elastic scalability, cost savings, remote access, and on-demand service provisioning [1]. These are the features that make it ideal for use in applications such as big data analytics, enterprise apps, and AI workloads. But in real-time interaction-driven scenarios, the high-speed internet and central processing-based dependence of the cloud poses a major bottleneck. Peculiarities such as higher latency, delay in data transmission, and susceptibility to security attacks owing to central storage are major drawbacks, particularly in mission-critical and latency sensitive applications [4] [5].

A new development is edge computing, which serves as a complementary solution to mitigate these limitations. It pushes computing tasks closer toward the data source—such as sensors, user devices, or local edge servers—allowing for rapid processing and low-latency responses [2] [3]. This paradigm is particularly relevant for scenarios such as self-driving cars and automated manufacturing, remote health monitoring, and real-time surveillance. By cutting down on the need to send data to centralized cloud infrastructure, edge computing alleviates network congestion and enhances privacy by limiting data exposure.

Nevertheless, edge computing introduces its own set of challenges. These include high initial deployment costs for edge infrastructure, limited resource scalability, and increased operational complexity due to the distributed nature of the system [4] [7]. Unlike cloud platforms that are maintained by centralized service providers, edge environments often require localized updates, security management, and consistent performance across heterogeneous hardware.

This area of research revolves around the trade-offs among cloud and edge computing with focus placed on performance criteria such as latency, bandwidth, scalability, energy efficiency, cost, and security. With the new computing demands evolving with the emergence of real-time and intelligent systems, there exists a greater requirement to assess a blended approach to computing that unites both paradigms. This paper provides an exploration of these trends, showcasing the architectural changes and market-driven implementations of edge, cloud, and hybrid models in many sectors.

III. RELATED WORK

Numerous studies have explored the capabilities and the challenges associated with both edge and cloud computing in diverse domains. Numerous studies have focused on how cloud computing contributes to supporting large-scale data processing, artificial intelligence (AI) training, and enterprise application hosting [1] [4]. Buyya et al. emphasized how cloud platforms such as Amazon Web Services (AWS) and Google Cloud have enabled researchers and organizations to conduct complex computational tasks without the high cost of initial infrastructure [1]. Similarly, Khan et al. discussed how cloud infrastructures facilitate enterprise agility and resource management through specialized services tailored to industry-specific needs [4].

Conversely, an increasing number of studies have shifted their focus to the benefits of edge computing, especially in latency-sensitive use cases. Shi et al. outlined the opportunities provided by edge computing in enabling real-time responsiveness, it also reduces bandwidth usage by processing data where it originates. [2]. Satyanarayanan highlighted its relevance in healthcare, particularly for wearable devices that demand continuous and immediate processing of health data [3]. Furthermore, Li et al. demonstrated the importance of edge solutions in autonomous systems, such as self-driving cars, where even minimal latency can significantly affect safety and performance [8].

Security concerns have been another prominent topic in both domains. Ghosh et al. examined data privacy risks associated with centralized cloud storage, especially in multitenant environments [5]. In contrast, edge computing's distributed architecture has been studied by Raza et al., who identified the physical security vulnerabilities of edge nodes and proposed a strategy that uses federated learning to improving data confidentiality [7]. Zhang et al. further analyzed the risks and challenges of securing real-time applications deployed on shared infrastructure in both edge and cloud settings [9].

Despite these contributions, several critical areas remain underexplored. The seamless blending of cloud and edge computing into hybrid models, performance benchmarking across use cases, and strategies for resource optimization are often discussed in isolation. There is limited literature that systematically compares both paradigms from architectural, operational, and cost perspectives while also considering emerging enablers such as 5G and Edge AI. This gap underlines a call for a thorough examination that consolidates existing findings and provides actionable insights for both academic and industry stakeholders.

IV. PROPOSED WORK

This review presents a structured comparative framework to evaluate cloud and edge computing paradigms across multiple dimensions including architecture, latency, scalability, bandwidth efficiency, security, energy consumption, cost, and application suitability. By bringing together ideas from previous research, real-world case studies, and emerging research trends, this work aims to offer a consolidated perspective that aids in selecting the optimal computing approach for varying use cases.

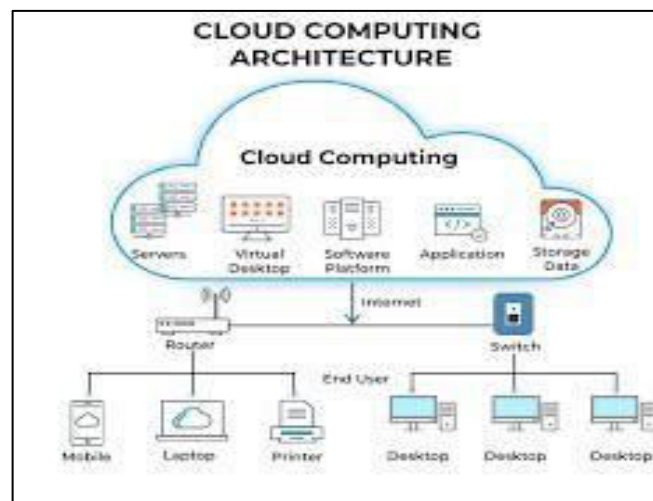


Figure1. Architecture of Cloud Computing

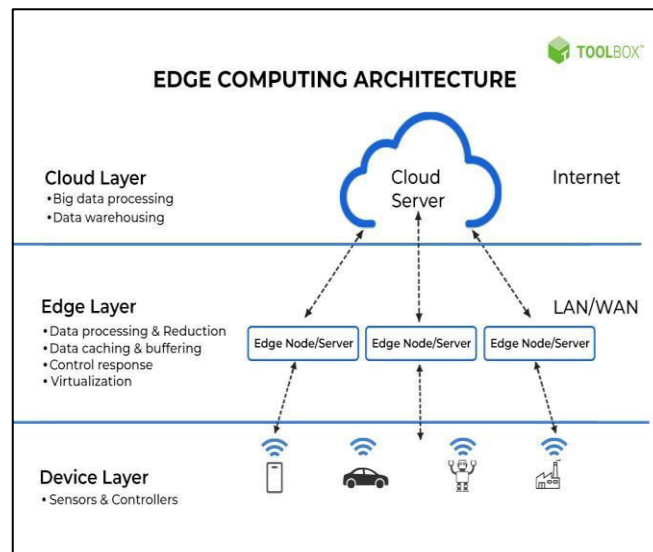


Figure2. Architecture of Edge Computing

As illustrated in Figures 1 and 2, the architectural models of edge and cloud computing differ fundamentally regarding the placement of data computation and response latency. Cloud uses a central infrastructure, whereas edge computing moves those functions closer to where the data is created for real-time responsiveness.

The architectural differences are fundamental: the approach taken by cloud computing is a centralized model where resources are provisioned from distant data centers, whereas edge computing distributes bringing processing nearer to the data's source. Our framework compares these models in realtime performance, with edge computing demonstrating lower latency and higher responsiveness in time-critical environments such as self-driving cars, alongside healthcare monitoring systems [3][8]. Conversely, cloud computing remains more effective for batch processing, AI model training, and large-scale analytics [1][4].

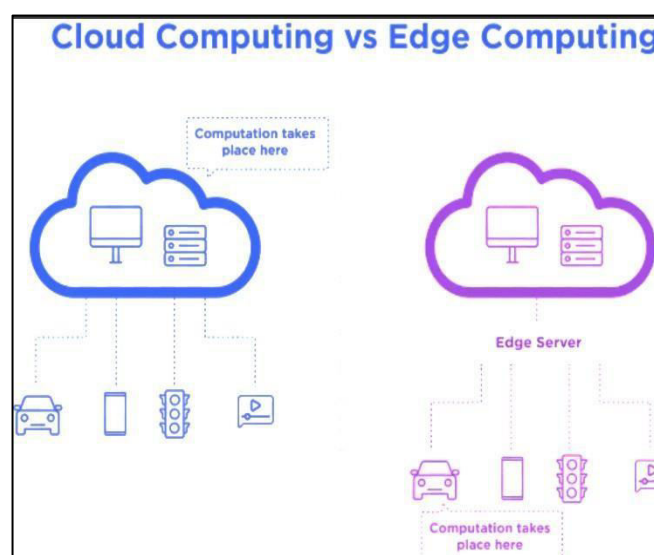


Figure3. Cloud vs. Edge Computing: Difference in Computation Location

Security and improved energy conservation also form key points of divergence. While centralized cloud systems benefit from unified control and standardized protocols, they remain vulnerable to large-scale breaches and single

points of failure [5]. Edge systems, although locally more secure, require distributed security strategies and physical protection at multiple endpoints [7]. Likewise, energy efficiency varies depending on usage — cloud environments optimize energy at scale, whereas edge computing reduces transmission power by processing locally [10][19].

Emerging trends are also addressed. The way edge computing is merging with 5G networks is enabling lowlatency, high-bandwidth applications in smart cities, industrial IoT, automated and connected transport infrastructure [8]. Additionally, the blending of AI with edge computing allows for real-time decision making without cloud dependency, a vital feature for applications in healthcare, surveillance, and manufacturing [9] [10]. Hybrid computing models that balance cloud scalability with edge responsiveness are being used by a growing number of industries to optimize performance, cost, and reliability. In this paper, we also provide a side-by-side tabular comparison of key operational parameters and a discussion of industry real-world examples in fields such as healthcare, autonomous systems, and smart manufacturing. This multifaceted comparison not only informs decision-making but also lays the groundwork for further research into integrated and adaptive computing systems.

Table 1: Comparative Analysis of Cloud and Edge Computing

Aspect	Cloud Computing	Edge Computing
Architecture	Centralized computing hubs [1][4]	Distributed computing near data source [2][3]
Latency [2][3][6]	Higher due to distance from data source	Very low; real-time processing at the edge [2][8][9]
Scalability	Highly scalable with virtually unlimited resources [1][14]	Limited by local infrastructure and hardware constraints [4][18]
Bandwidth Usage	High; requires continuous data transfer [2][4]	Low; processes most data locally [2][13][19]
Security	Vulnerable due to centralized storage; single point of failure [5][17]	More secure locally, but harder to manage across distributed nodes [5][13][20]
Energy Efficiency	Optimized in large-scale data centers [16]	Saves energy by avoiding cloud transfers [10][19]
Maintenance	Managed by providers; less user overhead [1][4]	Requires local updates and infrastructure upkeep [4][15]
Cost Structure	Cost-effective at scale and long term [1][14]	Higher upfront cost for edge deployment [13][18]
Use Case Fit	Ideal for batch tasks, analytics, and training AI models [1][14][16]	Suited for real-time, latency-sensitive apps like autonomous systems [2][8][13]
Internet Dependency	Fully dependent on continuous network access [1][4]	Operates with partial or no cloud connectivity [3][9][10]

V. CONCLUSIONS AND FUTURE RESEARCH OPPORTUNITIES

Our paper provides an in-depth comparative study of cloud and edge paradigms, including their architectural models, operational trade-offs, application spaces, and future trends. Cloud computing offers adaptable, elastic capabilities and economical solutions to centralized data storage and processing, a task ideal for enterprise systems, big data processing, and training AI models. But its limitations in latency, security, and uninterrupted connectivity are hurdles in real-time and mission-critical applications.

Edge computing overcomes much of these limitations by localizing processing right where the data is collected. This facilitates quick decision-making, less bandwidth consumption, as well as enhanced responsiveness for uses in areas like autonomous cars, remote tracking of health data, and industrial IoT. However, the decentralized structure of edge platforms introduces new challenges, such as infrastructure cost, reduced scalability, and higher maintenance

complexity. With growing computing requirements, hybrid models that integrate both cloud and edge architectures are becoming a strategic answer to balance latency, scale, and intelligence. Edge AI, 5G, and blockchain technologies are further augmenting the capabilities of these models, catalyzing innovation in areas requiring real-time, secure, and adaptive computing environments.

Future research needs to concentrate on the creation of standard frameworks for hybrid system integration, dynamic workload distribution, and smart resource management. Beyond that, security protocols designed for heterogeneous and distributed settings, and cost-optimization mechanisms for hybrid deployment, are also areas in need of ongoing investigation. Addressing these challenges will be key to achieving the full potential of edge-cloud convergence in the next decade.

REFERENCES

- [1] R. Buyya, C. Vecchiola, and S. T. Selvi, *Mastering Cloud Computing: Foundations and Applications Programming*. Morgan Kaufmann, 2013.
- [2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, 2016.
- [3] M. Satyanarayanan, "The Emergence of Edge Computing," *Computer*, IEEE, 2017.
- [4] L. U. Khan et al., "Edge Computing: A Survey," *Future Generation Computer Systems*, 2020.
- [5] A. Ghosh, S. Ghosh, and N. Chaki, *Security Issues in Cloud and Edge Computing*. Springer, 2021.
- [6] A. Javed et al., "On the growing role of hybrid edge-cloud architectures in real-time IoT systems," 2022.
- [7] S. Raza et al., "A Detailed Survey on Security Challenges in Edge Environments, Including Federated Learning," 2023.
- [8] Y. Li et al., "Study on 5G-enabled Edge AI, Especially for Autonomous Systems and Smart Cities," 2023.
- [9] Y. Zhang et al., "Implementation of AI-Driven Edge Computing with Low-Latency Applications," 2022.
- [10] T. Nguyen et al., "Review of Blockchain-Based Security for Edge Computing," 2024.
- [11] Y. Mao, J. Zhang, and K. B. Letaief, "Dynamic Computation Offloading for Mobile-Edge Computing with Energy Harvesting Devices," *IEEE Journal on Selected Areas in Communications*, 2017.
- [12] X. Chen et al., "Energy-Efficient Resource Allocation in Mobile Edge Computing with Collaborative D2D Communication," *IEEE Transactions on Wireless Communications*, 2019.
- [13] P. Hou, Q. Lv, and Y. Li, "Survey on Edge Computing and its Applications in Cloud and IoT," *IEEE Access*, 2018.
- [14] B. Ottenwalder, L. Coviello, V. Stauffer, and R. Schoenen, "The Economics of Edge Computing," *Computer*, 2016.
- [15] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint Optimization of Radio and Computational Resources for Multicell Mobile-Edge Computing," *IEEE Transactions on Signal Processing*, 2019.
- [16] Y. Chen, C. Li, and Y. Zheng, "Machine Learning Techniques for Edge Computing: A Review," *IEEE Access*, 2020.
- [17] X. Li, H. Hu, and C. Chen, "A Survey on the Security of Edge Computing in 5G Networks," *IEEE Internet of Things Journal*, 2021.
- [18] H. Gupta, A. V. Dastjerdi, S. K. Ghosh, and R. Buyya, "Edge Computing: Opportunities and Challenges in Distributed IoT," *ACM SIGCOMM Computer Communication Review*, 2018.
- [19] S. Wang et al., "A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications," *IEEE Access*, 2019.
- [20] J. Ren, M. Peng, and C. Wang, "A Survey on SoftwareDefined Networking and Edge Computing for IoT," *IEEE Communications Surveys & Tutorials*, 2020.

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 8.152