Volume 2 No 6, 2026
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Review Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Load Balancing in Edge-Cloud Ecosystems: A Survey of Hybrid Orchestration Models

Dr. Bal Krishna Sharma
Professor
Department of Computer Sciences and Applications
Mandsaur University, Mandsaur
bksharma7426@gmail.com

***Abstract*—Edge-cloud computing has emerged as a promising paradigm for supporting the growing demands of Internet of Things (IoT) applications, intelligent services, and data-intensive workloads. By integrating distributed edge resources with centralized cloud infrastructures, edge-cloud ecosystems enable low-latency processing, improved scalability, and efficient resource utilization. However, the dynamic and heterogeneous nature of these environments introduces significant challenges in workload distribution, resource allocation, and service orchestration. Consequently, load balancing and orchestration mechanisms have become essential for maintaining Quality of Service (QoS) and ensuring efficient system operation. This survey presents a comprehensive review of load-balancing techniques and hybrid orchestration models in edge-cloud ecosystems. A taxonomy of load-balancing approaches is discussed based on decision location, orchestration strategy, and system architecture. The study further examines recent advances in distributed computing frameworks, Software-Defined Networking (SDN), Network Function Virtualization (NFV), workload scheduling, collaborative resource management, and intelligent orchestration techniques. A comparative analysis of existing research is provided to identify key strengths, limitations, and emerging trends. The findings indicate that hybrid orchestration models, which combine centralized coordination with decentralized decision-making, offer significant advantages in terms of scalability, adaptability, resource utilization, and service performance. Additionally, critical challenges and future research directions are highlighted to support the development of efficient, secure, and intelligent edge-cloud ecosystems.**

Keywords—Edge-cloud computing, load balancing, hybrid orchestration, Resource Management, Workload scheduling, Internet of Things (iot), Quality of Service (qos), distributed computing.

I. Introduction

The rapid proliferation of Internet of Things (IoT) devices, autonomous systems, and intelligent applications has significantly increased the demand for scalable, efficient, and intelligent distributed computing infrastructures [1]. Traditional cloud-centric architectures provide extensive computational and storage resources; however, they often face challenges related to latency, bandwidth consumption, and real-time responsiveness. These limitations become critical for emerging applications such as smart cities, healthcare systems, industrial automation, autonomous vehicles, and augmented reality services, where timely processing and decision-making are essential [2].

To overcome these challenges, edge computing has emerged as a complementary paradigm that extends computational and storage capabilities closer to end users and data sources. By processing data at the network edge, edge computing reduces communication delays and network congestion while improving service responsiveness [3][4]. The integration of edge and cloud computing has led to the development of edge-cloud ecosystems, where computing resources are distributed across edge devices, edge servers, and centralized cloud data centers to support diverse application requirements [5].

As edge-cloud environments continue to grow in scale and complexity, efficient resource management becomes increasingly important [6][7][8]. Dynamic workloads, heterogeneous resources, varying network conditions, and user mobility can lead to uneven resource utilization and performance degradation. Therefore, load balancing plays a crucial role in distributing workloads across available resources to improve system performance, reduce latency, enhance resource utilization, and maintain Quality of Service (QoS). In parallel, orchestration mechanisms are required to coordinate resources and services across multiple layers of the edge-cloud infrastructure [9].

Recent research has focused on hybrid orchestration models that combine the strengths of centralized cloud management and decentralized edge intelligence. These models enable adaptive workload placement, resource scheduling, and service orchestration across the edge-cloud continuum, providing greater scalability, flexibility, and reliability. However, challenges related to interoperability, energy efficiency, security, privacy, and real-time decision-making remain significant barriers to widespread deployment [10] [11]. This survey presents a comprehensive review of load-balancing strategies and hybrid orchestration models in edge-cloud ecosystems. It examines existing architectures, resource management techniques, scheduling approaches, and orchestration frameworks while highlighting current challenges, research gaps, and future directions for building efficient and scalable edge-cloud computing environments.

A. *Structure of the paper*

This paper is organized as follows Section II overview of Taxonomy of Load Balancing Techniques. Section III Hybrid Orchestration Models in Edge-Cloud. Section IV Challenges and Research Gaps. Section V Research on Literature Review Section VI Conclusions and Future Direction.

II. Overview Of Taxonomy Of Load Balancing Techniques

Load balancing in edge-cloud ecosystems can be classified across multiple dimensions, reflecting where decisions are made, how orchestration is carried out, and the architectural setup of the system. This taxonomy facilitates a better understanding of existing approaches and aids in selecting suitable strategies for diverse IoT and cloud-edge use cases [12] [13].

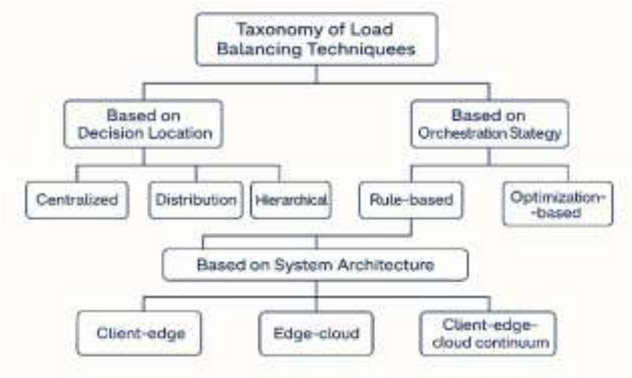

Fig. 1. Taxonomy Load Balancing

Figure 1 presents a taxonomy of load-balancing techniques in edge-cloud environments. The classification is organized according to decision location, orchestration strategy, and system architecture, highlighting centralized, distributed, hierarchical, rule-based, and optimization-based approaches across client-edge, edge-cloud, and edge-cloud continuum computing infrastructures.

A. *Decision Location*

Load-balancing techniques can be classified according to the location where load-distribution decisions are made within the computing infrastructure. The decision-making entity may be centralized, distributed across multiple nodes, or organized hierarchically across different layers of the system. This classification influences system scalability, fault tolerance, communication overhead, and overall resource management efficiency in edge-cloud environments [14]. And there are three types of Decision Location are as follows:

1) *Centralized*

In centralized load balancing, a single controller or orchestrator collects global information about system resources, demand patterns, and network conditions to make load distribution decisions [15]. While this provides a holistic view and can optimize globally, it suffers from scalability issues and is vulnerable to single points of failure.

2) *Distribution*

Distributed load balancing decentralizes the decision-making process across multiple nodes. Each node either cooperates with others or makes autonomous decisions based on local knowledge [16]. This improves scalability and fault tolerance but may result in sub-optimal global performance due to limited information.

3) *Hierarchical*

Hierarchical models combine the benefits of centralized and distributed approaches by structuring decision-making into layers, such as client-edge-cloud or regional-global [17]. Upper tiers oversee broader policies, while lower tiers handle real-time local balancing, enabling both scalability and coordination.

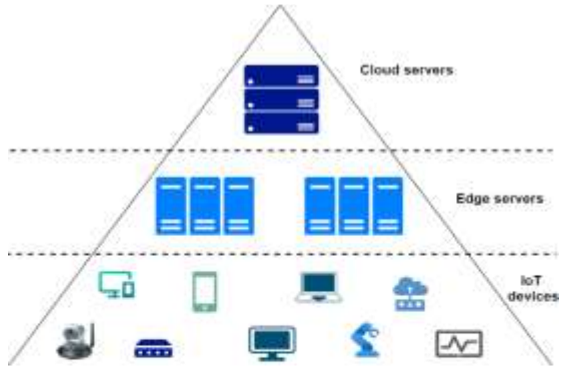

Fig. 2. Hierarchical models

Figure 2 illustrates a hierarchical edge-cloud computing architecture consisting of three layers: IoT devices, edge servers, and cloud servers. IoT devices generate data at the network edge, edge servers perform intermediate processing and resource management, while cloud servers provide large-scale computation, storage, and centralized service coordination.

B. *Orchestration Strategy*

Load-balancing techniques can also be classified according to the orchestration strategy used to manage workload distribution and resource allocation. The orchestration mechanism determines how decisions are made, coordinated, and executed across the edge-cloud infrastructure. Depending on the level of intelligence and adaptability required, orchestration strategies are commonly categorized as rule-based[18], optimization-based, and AI/ML-based approaches, each offering different trade-offs in terms of complexity, scalability, responsiveness, and performance.

1) *Rule-Driven Orchestration*

Rule-based load balancing follows predefined heuristics or static policies, such as round-robin, least-connections, or fixed thresholds [19]. While simple and computationally lightweight, these methods often lack adaptability to dynamic network states or traffic fluctuations.

2) *Optimization-Driven Orchestration*

These techniques model load balancing as a mathematical optimization problem (e.g. linear programming, convex optimization) aiming to minimize latency, energy, or cost. They provide more precise control but are computationally expensive and may not scale well in real-time environments.

3) *Intelligent Orchestration (AI/ML)*

AI/ML-based approaches learn from historical and real-time data to predict traffic patterns and make intelligent balancing decisions. These include reinforcement learning, deep learning, and federated learning techniques that adapt over time, improving performance in complex, dynamic scenarios [20].

C. *System Architecture on cloud edge*

Edge-cloud computing integrates edge resources with centralized cloud infrastructures to support efficient workload execution and resource management [21]. By distributing computation across multiple layers, these architectures reduce latency, improve scalability, and enhance Quality of Service

(QoS). Depending on the location of computational resources and workload distribution mechanisms, edge-cloud systems can be organized into different architectural models [22].

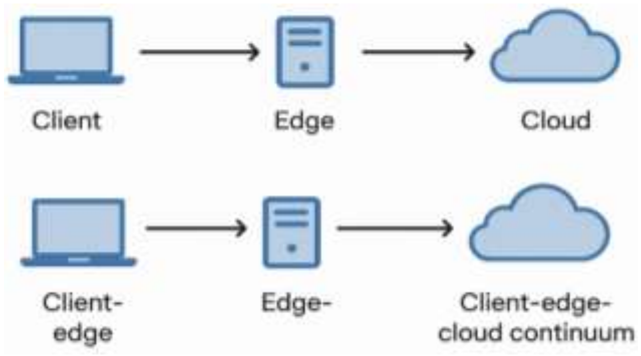

Fig. 3. System Architecture of Cloud Edge

Figure 3 illustrates a cloud-edge system architecture comprising client devices, edge servers, and cloud resources. The architecture distributes workloads across computing layers, enabling low-latency processing at the edge while leveraging cloud resources for large-scale computation, scalability, and efficient resource management.

1) Client-Edge

In this architecture, balancing is performed between client devices and edge nodes. Load is distributed to nearby edge nodes based on proximity, client demand, or edge resource availability [23]. This reduces latency and conserves bandwidth by processing data near the source.

2) Edge-Cloud

Load balancing decisions occur between edge servers and centralized cloud infrastructure. Tasks may be offloaded to the cloud when edge resources are constrained, or to leverage the cloud's superior computational capacity [24]. This setup enables elastic scalability and supports intensive applications.

3) Client-Edge-Cloud Continuum

This full-spectrum architecture supports dynamic load balancing across all layers—client, edge, and cloud [25]. It enables seamless offloading and task migration across the continuum, using context-aware and adaptive strategies to maintain QoS and optimize resource utilization.

III. Hybrid Orchestration Models In Edge-Cloud

Hybrid orchestration models in edge-cloud ecosystems combine elements of both centralized and decentralized control to strike a balance between global optimization and local responsiveness. In these models, orchestration logic is distributed across multiple layers—such as the client, edge, and cloud—based on workload requirements, latency sensitivity, data locality, and real-time constraints. The growing heterogeneity of devices and services, along with dynamic workload patterns and network variability, makes purely centralized or purely decentralized orchestration insufficient [26]. Hybrid models are thus essential to support adaptive, scalable, and context-aware orchestration strategies that can dynamically switch control or collaborate across layers to optimize performance, cost, and energy efficiency [27].

A. Hybrid Approaches

1) Decentralized Decision-Making with Centralized Coordination

In this approach, local nodes (e.g. edge devices) make fast, autonomous decisions based on local metrics and resource availability. However, their actions are periodically aligned with a global controller (often in the cloud) that maintains a broader system-wide view. This setup reduces latency while retaining strategic global optimization through periodic synchronization. For example, edge nodes may handle task offloading locally, but periodically receive updated policies from a centralized orchestrator.

2) Federated and Collaborative Scheduling

Federated scheduling distributes decision-making among independent edge/cloud nodes that share a common objective but operate autonomously. These nodes may collaborate via peer-to-peer communication or hierarchical structures to allocate resources and schedule tasks. Collaborative scheduling enables multi-stakeholder environments (e.g. multiple service providers) to share edge resources while maintaining autonomy, privacy, and fairness.

3) Betweenness and Graph-Theory-Based Routing

Inspired by complex systems and network science, some hybrid models use graph-theoretic measures, such as betweenness centrality, to identify critical nodes and links in the network. These insights are used to guide dynamic load balancing and task routing decisions in a decentralized manner, with occasional global feedback. For instance, in an IoT network, tasks might be offloaded based on topological importance of nodes (e.g. those with high betweenness values) to optimize load and avoid bottlenecks.

B. Real-Time Decision Making

One of the most critical challenges in hybrid edge-cloud orchestration is achieving low-latency, real-time decision-making. Applications like autonomous vehicles, remote surgery, and smart manufacturing demand near-instantaneous processing. However, orchestration frameworks often suffer from delays due to complex decision processes, communication overhead, and data synchronization between distributed nodes. Current AI/ML-based strategies are either too computationally intensive for edge deployment or lack sufficient adaptability in dynamic environments. There is a pressing need for lightweight, real-time-aware orchestration models that balance responsiveness with decision accuracy.

C. Security and Privacy Constraints

In edge-cloud systems, data often travels through and resides in multiple layers, from user devices to edge nodes to cloud platforms. This multi-tiered structure introduces significant attack surfaces and privacy leakage risks [28] [29] [30]. While traditional encryption and access control models exist, they may not suit the constrained resources of edge devices. Furthermore, federated learning and collaborative orchestration methods introduce new privacy concerns. Homomorphic encryption, differential privacy, and zero-trust architectures are promising, but require optimization for hybrid environments. This domain remains underexplored for real-world deployments.

D. Interoperability and Standardization

The edge-cloud ecosystem comprises heterogeneous devices, platforms, and orchestration frameworks, often developed by different vendors with varying standards. This leads to poor interoperability and fragmented architectures, making orchestration across multiple domains extremely challenging. The lack of unified APIs, data models, and service descriptors limits the portability and composability of services. While initiatives like ETSI MEC and OpenFog Consortium attempt standardization, real-world adoption is slow. Research is needed in cross-platform orchestration protocols, semantic interoperability layers, and open orchestration interfaces.

E. Resource-Constrained Edge Devices

Unlike cloud servers, edge devices have limited CPU, memory, storage, and battery life. Orchestration mechanisms that assume abundant resources can overwhelm these devices, causing service degradation or failures. Even running lightweight containerized services or AI models at the edge can strain devices. There's a critical need for resource-aware orchestration strategies that dynamically assess and allocate workloads based on device profiles, mobility patterns, and contextual awareness. Research should also focus on collaborative edge federations to pool resources while maintaining autonomy.

IV. LITERATURE REVIEW

This section reviews recent research on load balancing and orchestration techniques in edge-cloud ecosystems. The selected studies investigate distributed computing architectures, resource management strategies, workload scheduling mechanisms, and intelligent orchestration approaches aimed at improving scalability, resource utilization, latency, and Quality of Service (QoS).

H. F. Shahid et al. (2026) analyzes the strengths and limitations of distributed computing frameworks, including centralized IoT architectures, fog computing, edge computing, local-edge architectures, and the three-tier edge–cloud continuum. Particular emphasis is placed on the three-tier edge–cloud continuum, which addresses key limitations of earlier frameworks by enabling efficient IoT service orchestration through distributed resources near end users while utilizing centralized cloud resources. The frameworks are evaluated based on scalability, resource efficiency, adaptability, resource availability, security and privacy, and robustness [31].

B. U. Kazi et al. (2025) The explosion of connected devices and data transmission in the Internet of Things (IoT) era places a substantial burden on cloud computing capabilities. Moreover, IoT devices are typically deployed at the network edge and have limited resources. To address these challenges, edge-cloud distributed computing networks have emerged. Due to their distributed nature, technologies such as Software-Defined Networking (SDN) and Network Function Virtualization (NFV) play key roles in resource management, orchestration, and load balancing. This article surveys these technologies, focusing on SDN controllers, orchestration, and the role of Artificial Intelligence (AI) in enhancing controller capabilities within edge-cloud computing networks [32]

M. S. Aslanpour et al. (2024) proposes a performance-driven weight-tuning approach for effective load balancing based on node characteristics and capabilities. By profiling nodes, performance metrics such as throughput, energy efficiency, response time, AI accuracy, and cost are evaluated. A weighted round-robin strategy is introduced to optimize load distribution according to the significance of these metrics. To address multiple objectives, a multi-objective approach balances different performance goals simultaneously. The study also explores a coordinated distributed approach to overcome centralized load-balancing limitations. Furthermore, Hedgi, a heterogeneous serverless edge architecture, is introduced to support and validate the proposed load-balancing policies. Extensive experiments using a real-time object detection application demonstrate the effectiveness of the proposed approach [33].

M. N. Jamil et al. (2023) Multi-access Edge Computing (MEC) is an edge computing architecture designed to support resource-intensive and latency-sensitive applications through computation offloading. However, dynamic and unpredictable user demands can lead to non-optimal task allocation, increasing latency and task failures. To address this challenge, a two-stage Belief Rule-Based (BRB) workload orchestrator is proposed to distribute workloads to optimal computing units while satisfying Quality of Service (QoS) requirements. The orchestrator determines task execution locations based on network conditions, resource availability, and task requirements. Simulation experiments using EdgeCloudSim compare the proposed approach with four existing orchestration methods [34].

M. Raeisi-Varzaneh et al. (2023) Edge servers can effectively utilize computational and storage resources for task execution. However, offloading all computing tasks to edge servers may lead to increased processing delays and energy consumption. In addition, underutilized edge devices and cloud data centers can result in resource waste. Therefore, collaborative scheduling among edge devices, edge servers, and cloud centers is essential for efficient edge computing operations. This survey reviews edge computing architectures, resource scheduling techniques under different collaboration modes, and task offloading schemes. It also examines fairness and load-balancing metrics used in scheduling to improve resource utilization and system performance [35].

A. Paszkiewicz et al. (2022) This contribution concerns load balancing, based on mechanisms from complex systems theory dedicated to IoT solutions within the Edge-Cloud continuum. The basis of considered mechanisms is the betweenness analysis applied to distributed nodes in a wireless IoT network. A high value of this parameter can indicate the key role of a given node, which is often reflected in its high load. In addition, both the distance and the error rate for connections between nodes are considered. The proposed solution aims at providing path redundancy in the wireless network and enabling efficient distribution of the network traffic load [36].

Table I compares recent load-balancing and orchestration approaches in edge-cloud and IoT environments, highlighting their focus areas, key findings, challenges, and contributions toward improving resource utilization, scalability, and QoS.

TABLE I. COMPARATIVE ANALYSIS OF LOAD BALANCING AND ORCHESTRATION TECHNIQUES IN EDGE-CLOUD AND IOT NETWORKS

Author's	Focus Area	key findings	Challenges	Key Contributions
H. F. Shahid et al. (2026)	Distributed computing frameworks for IoT	Three-tier edge–cloud continuum improves service orchestration and resource utilization	Scalability, security, privacy, and robustness issues across heterogeneous environments	Comparative evaluation of centralized, fog, edge, local-edge, and edge–cloud continuum architectures

B. U. Kazi et al. (2025)	SDN, NFV, and AI in edge-cloud networks	SDN and NFV enhance resource management and load balancing; AI improves controller intelligence	Complex orchestration and management of distributed resources	Comprehensive survey of SDN/NFV-based orchestration and AI-enabled network control
M. S. Aslanpour et al. (2024)	Performance-aware load balancing	Weighted round-robin based on throughput, latency, energy, accuracy, and cost improves performance	Multi-objective optimization and centralized bottlenecks	Proposed adaptive weight-tuning strategy and Hedgi heterogeneous edge architecture
M. N. Jamil et al. (2023)	MEC workload orchestration	BRB-based orchestrator improves task allocation and QoS satisfaction	Dynamic workloads and resource uncertainty	Two-stage workload orchestration considering network and resource conditions
M. Raeisi-Varzaneh et al. (2023)	Resource scheduling and task offloading	Collaborative scheduling enhances resource utilization and system efficiency	Balancing latency, energy consumption, and fairness	Survey of scheduling techniques, collaboration modes, and load-balancing metrics
A. Paszkiewicz et al. (2022)	IoT load balancing in edge-cloud continuum	Betweenness-based routing improves traffic distribution and path redundancy	Network congestion and uneven node load	Introduced complex-system-based load balancing using network topology metrics

V. Conclusion And Future Work

Edge-cloud computing has emerged as an effective paradigm for supporting modern distributed applications that require low latency, scalability, and efficient resource utilization. In such environments, load balancing and orchestration play a vital role in ensuring optimal workload distribution and service performance. This survey presented a comprehensive review of load-balancing techniques and hybrid orchestration models in edge-cloud ecosystems. A taxonomy based on decision location, orchestration strategy, and system architecture was discussed, followed by an analysis of recent research contributions in distributed computing frameworks, workload scheduling, and intelligent resource management. The comparative study revealed that hybrid orchestration models, which combine centralized coordination with decentralized decision-making, provide improved adaptability, scalability, and QoS. Additionally, key challenges including real-time decision-making, security, interoperability, and resource constraints were identified. The findings highlight the importance of hybrid orchestration as a promising approach for managing increasingly complex and heterogeneous edge-cloud environments.

Future research should focus on developing intelligent, secure, and resource-aware orchestration frameworks capable of supporting dynamic edge-cloud environments. Integrating AI-driven decision-making, energy-efficient scheduling, interoperability standards, and privacy-preserving mechanisms will be essential for improving performance and enabling large-scale deployment of next-generation edge-cloud ecosystems.

References

[1] S. S. Nasrin, “Nebula Core : A Scalable Multimodal Framework for Distributed Intelligence in Edge-Cloud Systems Methodology :,” vol. 1, pp. 2–5, 2025, doi: 10.13140/RG.2.2.13170.36807.

[2] X. Xu, S. Zang, M. Bilal, X. Xu, and W. Dou, “Intelligent architecture and platforms for private edge cloud systems: A review,” *Futur. Gener. Comput. Syst.*, vol. 160, no. June, pp. 457–471, 2024, doi: 10.1016/j.future.2024.06.024.

[3] A. Joon, B. K. R. Janumpally, A. Gogineni, and P. Chatterjee, “Efficient Large-Scale Intrusion Identification and Prevention in Distributed Cloud Networks Using Artificial Intelligence,” in *2025 5th International Conference on Intelligent Technologies (CONIT)*, HUBBALI, India: IEEE, 2025, pp. 1–8, September. doi: 10.1109/CONIT65521.2025.11167760.

[4] S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, “Advanced Machine Learning Techniques for Prediction of Customer Churn in Telecommunication Sector,” in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, Rourkela, India: IEEE, 2025, pp. 1–6, August. doi: 10.1109/INDISCON66021.2025.11252233.

[5] F. C. Andriulo, M. Fiore, M. Mongiello, E. Traversa, and V. Zizzo, “Edge Computing and Cloud Computing for Internet of Things: A Review,” *Informatics*, vol. 11, no. 4, Sep. 2024, doi: 10.3390/informatics11040071.

[6] K. Jangiti, “Design and Validation of a Machine Identity Governance Framework for AI Agents in Multi-Cloud Environments,” in *SoutheastCon 2026*, IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/SoutheastCon63549.2026.11476363.

[7] R. K. Gadiraju, “Artificial Intelligence for Resource Optimization in Cloud Computing Environments,” *J. Electr. Syst.*, vol. 20, no. 6, pp. 3164–3174, March, 2024.

[8] S. Sen, “Data Stewardship: How AI Agents Form the Pillars for Effective Data and AI Governance,” *Int. J. Emerg. Trends Comput. Sci. Inf. Technol.*, vol. 5, no. 4, pp. 01–04, December, 2024, doi: https://doi.org/10.63282/3050-9246.

[9] T. P. Patel, “Adaptive Token Routing for Heterogeneous LLM Inference in Edge-Cloud Continuum,” in *SoutheastCon 2026*, Huntsville, AL, USA: IEEE, 2026, pp. 1–7, April. doi: 10.1109/SoutheastCon63549.2026.11476596.

[10] M. Trigka and E. Dritsas, “Edge and Cloud Computing in Smart Cities,” *Futur. Internet*, vol. 17, no. 3, 2025, doi: 10.3390/fi17030118.

[11] I. Korontanis, A. Makris, and K. Tserpes, “A Survey on Modeling Languages for Applications Hosted on Cloud-Edge Computing Environments,” *Appl. Sci.*, vol. 14, no. 6, 2024, doi: 10.3390/app14062311.

[12] R. Vaño, I. Lacalle, P. Sowiński, R. S-Julián, and C. E. Palau, “Cloud-Native Workload Orchestration at the Edge: A Deployment Review and Future Directions,” 2023. doi: 10.3390/s23042215.

[13] V. K. Sharma and K. S. Abhilash, “Latency-Aware Edge-Cloud Architecture for 5G IoT Integration,” in *2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, Sep. 2025, pp. 1398–1405. doi: 10.1109/ICESC65114.2025.11212232.

[14] A. Ullah *et al.*, “Orchestration in the Cloud-to-Things compute continuum: taxonomy, survey and future directions,” 2023. doi: 10.1186/s13677-023-00516-5.

[15] Y. Chiang *et al.*, “Management and Orchestration of Edge Computing for IoT: A Comprehensive Survey,” *IEEE Internet Things J.*, 2023, doi: 10.1109/JIOT.2023.3245611.

[16] F. Al-Doghman, N. Moustafa, I. Khalil, N. Sohrabi, Z. Tari, and A. Y. Zomaya, “AI-Enabled Secure Microservices in Edge Computing: Opportunities and Challenges,” *IEEE Trans. Serv. Comput.*, 2023, doi: 10.1109/TSC.2022.3155447.

[17] B. Kar, W. Yahya, Y. D. Lin, and A. Ali, “Offloading Using Traditional Optimization and Machine Learning in Federated Cloud-Edge-Fog Systems: A Survey,” *IEEE Commun. Surv. Tutorials*, 2023, doi: 10.1109/COMST.2023.3239579.

[18] S. Sen, “AI-Enabled Substation Architectures for Autonomous Power Systems: Reliability, Asset Intelligence, and Grid-Edge Analytics,” *ijcttjournal*, vol. 74, no. 2, pp. 01–04, Feb, 2026.

[19] M. Gaglianese, J. Soldani, S. Forti, and A. Brogi, “Green Orchestration of Cloud-Edge Applications: State of the Art and Open Challenges,” in *Proceedings - 17th IEEE International Conference on Service-Oriented System Engineering, SOSE 2023*, 2023. doi: 10.1109/SOSE58276.2023.00036.

[20] Z. A. Khan, I. A. Aziz, N. A. B. Osman, and I. Ullah, “A Review on Task Scheduling Techniques in Cloud and Fog Computing: Taxonomy, Tools, Open Issues, Challenges, and Future Directions,” *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3343877.

[21] H. N. Dholariya, “Regulatory-Grade Autonomous Data Modernization: The RAMA Framework for Compliance-Aware AI-Native Cloud Architectures,” *J. Comput. Anal. Appl.*, vol. 35, no. 1, pp. 883–898, Jan, Jan. 2026, doi: 10.48047/jocaaa.2026.35.01.40.

[22] A. Furutanpey *et al.*, “Architectural Vision for Quantum Computing in the Edge-Cloud Continuum,” in *Proceedings - 2023 IEEE International Conference on Quantum Software, QSW 2023*, 2023. doi: 10.1109/QSW59989.2023.00021.

[23] H. J. Lee, S. Noghabi, B. Noble, M. Furlong, and L. P. Cox, “BumbleBee: Application-aware adaptation for edge-cloud orchestration,” in *Proceedings - 2022 IEEE/ACM 7th Symposium on Edge Computing, SEC 2022*, 2022. doi: 10.1109/SEC54971.2022.00017.

[24] J. Dogani, R. Namvar, and F. Khunjush, “Auto-scaling techniques in container-based cloud and edge/fog computing: Taxonomy and survey,” 2023. doi: 10.1016/j.comcom.2023.06.010.

[25] J. Zhang, F. Keramat, X. Yu, D. M. Hernandez, J. P. Queralta, and T. Westerlund, “Distributed Robotic Systems in the Edge-Cloud Continuum with ROS 2: a Review on Novel Architectures and Technology Readiness,” in *2022 7th International Conference on Fog and Mobile Edge Computing, FMEC 2022*, 2022. doi: 10.1109/FMEC57183.2022.10062523.

[26] G. K. Walia, M. Kumar, and S. S. Gill, “AI-Empowered Fog/Edge Resource Management for IoT Applications: A Comprehensive Review, Research Challenges, and Future Perspectives,” *IEEE Commun. Surv. Tutorials*, 2024, doi: 10.1109/COMST.2023.3338015.

[27] J. B. Mehta, “Autonomous Workload Right-Sizing for Multi-Cloud Cost Optimization,” in *2026 International Conference on Artificial Intelligence, Systems, and Emerging Technologies (ICAISET)*, Cairo, Egypt: IEEE, 2026, pp. 1–11, June. doi: 10.1109/ICAISET66439.2026.11541899.

[28] N. R. Barot, “Transparency-Driven Operational Intelligence: A New Data Governance Model for High-Risk Industrial Automation,” *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 63s, pp. 1019–1028, Dec. 2025, doi: 10.52783/jisem.v10i63s.13975.

[29] N. D. Bhandarwar, “A Mathematical Framework For Explainable And Adversarially Robust Ids Using Ml For Large-Scale Enterprise And Cloud Systems,” *Int. J. Appl. Math.*, vol. 39, no. 1s, pp. 1200–1212, Jan. 2026, doi: 10.12732/ijam.v39i1s.1910.

[30] R. rao Thallada and N. Alapati, “Privacy and Cybersecurity Convergence: GRC Controls for Data Protection,” *J. Bus. Manag. Stud.*, vol. 8, no. 5, pp. 42–48, March, 2026, doi: 10.32996/jbms.

[31] H. F. Shahid, B. Akdemir, J. Islam, I. Ahmad, and E. Harjula, “IoT Service Orchestration in Edge–Cloud Continuum With 6G: A Review,” in *IEEE Internet of Things Journal*, IEEE, 2026, pp. 20339–20358, Februrary. doi: 10.1109/JIOT.2026.3663918.

[32] B. U. Kazi, M. K. Islam, M. M. H. Siddiqui, and M. Jaseemuddin, “A Survey on Software Defined Network-Enabled Edge Cloud Networks: Challenges and Future Research Directions,” *Network*, vol. 5, no. 2, p. 16, May 2025, doi: 10.3390/network5020016.

[33] M. S. Aslanpour, A. N. Toosi, M. A. Cheema, M. B. Chhetri, and M. A. Salehi, “Load balancing for heterogeneous serverless edge computing: A performance-driven and empirical approach,” *Futur. Gener. Comput. Syst.*, 2024, doi: 10.1016/j.future.2024.01.020.

[34] M. N. Jamil, M. S. Hossain, R. U. Islam, and K. Andersson, “Workload Orchestration in Multi-Access Edge Computing Using Belief Rule-Based Approach,” *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3326244.

[35] M. Raeisi-Varzaneh, O. Dakkak, A. Habbal, and B. S. Kim, “Resource Scheduling in Edge Computing: Architecture, Taxonomy, Open Issues and Future Research Directions,” *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3256522.

[36] A. Paszkiewicz *et al.*, “Network Load Balancing for Edge-Cloud Continuum Ecosystems,” in *Lecture Notes in Electrical Engineering*, 2022. doi: 10.1007/978-981-19-1677-9_56.

