Hybrid Edge–Cloud Models for IoT: A Systematic Review

Harpreet Kaur[1], Komalpreet Kaur[2]

[1]Student, Global Group of Institutes, Amritsar, Punjab, India
[2]Assistant Professor, Department of Computer Application, Global Group of Institutes, Amritsar, Punjab, India

Abstract

The rapid proliferation of Internet of Things (IoT) devices across diverse domains such as smart cities, healthcare, industrial automation, and intelligent transportation systems has led to an exponential growth in data generation and real-time processing demands. Traditional cloud-centric computing architectures, although powerful, face significant challenges in meeting the stringent requirements of modern IoT applications, particularly in terms of latency, bandwidth utilization, scalability, energy efficiency, and data privacy. To address these limitations, hybrid edge–cloud computing has emerged as a promising paradigm that synergistically integrates the global processing capabilities of cloud computing with the localized, low-latency intelligence of edge computing. This paper presents a comprehensive review of hybrid edge–cloud computing models for IoT systems, focusing on architectural designs, enabling technologies, task offloading strategies, scheduling mechanisms, and security frameworks. The role of key enabling technologies such as 5G/6G networks, virtualization, containerization, artificial intelligence, and software-defined networking in enhancing hybrid deployments is also discussed in detail. Additionally, this review examines task scheduling and resource management techniques that aim to optimize energy consumption, reduce response time, and improve system throughput.

Keywords: Hybrid Edge–Cloud Computing, Internet of Things (IoT), Edge Computing, Cloud Computing, Task Offloading, Resource Scheduling, Low Latency, Scalability, Data Security, 5G Networks, Artificial Intelligence.

1. Introduction

IoT devices are generating enormous amounts of distributed, heterogeneous data. Traditional cloud infrastructures struggle to meet stringent requirements for real-time responsiveness and low-latency communication [1]. These limitations are resolved through edge computing, which processes data located near the source, minimizing round-trip delays and improving QoS (Quality of Service) [2]. However, edge nodes typically have limited resource capacity. A **hybrid edge–cloud model** combines both paradigms, enabling flexible workload distribution, enhanced scalability, and improved resilience [3].

2. Architecture of Hybrid Edge–Cloud Models

2.1 Multi-Tier Layered Architecture

Most hybrid architectures adopt a **three-tier model** consisting of:

- **IoT/Device Layer:** Sensors and actuators that generate raw data.

- **Edge Layer**: Gateways, micro-data centers, routers, and fog infrastructure nodes performing pre-processing.
- **Cloud Layer:** Centralized servers providing large-scale analytics and storage.

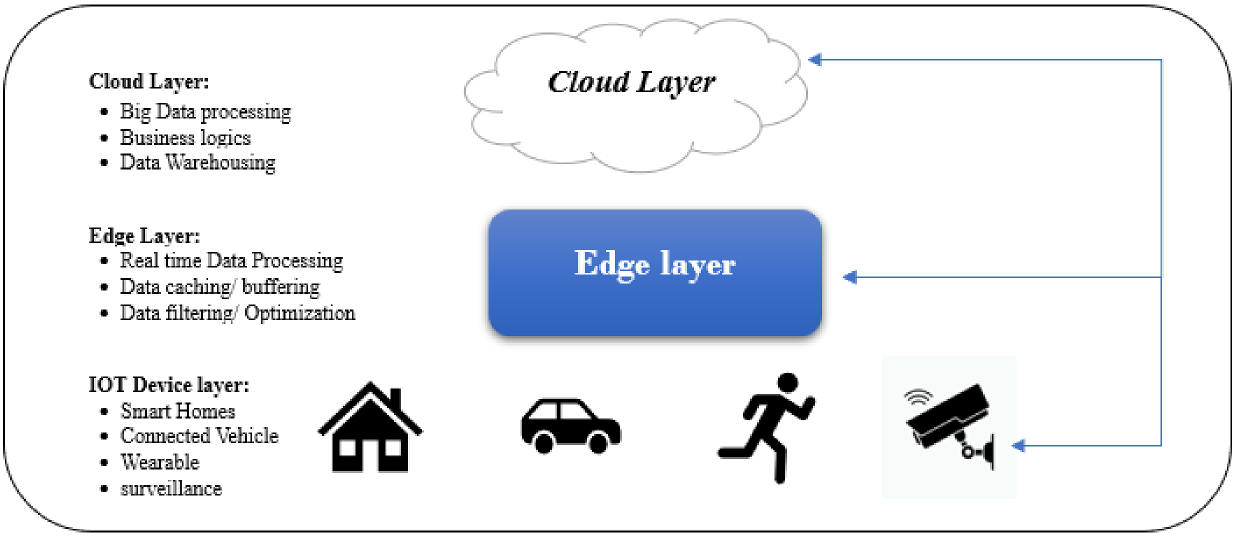

Figure 1: Hybrid Edge-cloud Model

This design is widely supported due to its modularity and adaptability [4]. Edge tier design reduces workload pressure on cloud infrastructure [5].

2.2 Fog-Enhanced Edge–Cloud Architecture

Fog computing implements an intermediate layer between edge and cloud, enabling distributed intelligence across multiple network nodes [6]. It improves contextual awareness and supports delay-sensitive IoT applications such as healthcare and autonomous systems [7].

3. Workload Allocation Strategies in Hybrid Systems

3.1 Latency-Driven Offloading

Many models use **latency thresholds** to determine whether tasks remain at the edge or are migrated to the cloud Techniques like dynamic task partitioning and QoS-aware schedulers contribute significant performance gains [8].

3.2 Resource-Aware Task Scheduling

Hybrid architectures often employ scheduling algorithms such as:

- **Reinforcement learning-based offloading** [9]
- **Genetic algorithms** for multi-objective optimization [10]
- **Heuristic-based resource allocation** [11]

These approaches aim to balance computation load, minimize energy consumption, and avoid network bottlenecks.

3.3 Collaborative Edge–Cloud Learning

Federated learning is increasingly integrated into hybrid models to allow distributed model training without sending pushing data uploaded to the cloud [12]. This boosts privacy and reduces bandwidth consumption.

4. Key Enabling Technologies

4.1 Containerization and Virtualization

Containers such as Docker improve workload mobility between edge and cloud nodes [13]. Lightweight virtualization enhances system flexibility and supports dynamic service deployment.

4.2 5G and Beyond Connectivity

5G networks offer ultra-low latency, network slicing, and massive IoT connectivity—critical for hybrid computing [14]. Multi-access edge computing (MEC) further enhances distributed processing capabilities [15].

4.3 AI-Driven Optimization

AI techniques enable:

- Prediction-based load balancing
- Adaptive scheduling
- Failure detection at edge nodes

These techniques help maintain QoS in heterogeneous environments [16].

5. Security and Privacy Implications

5.1 Data Protection and Encryption

Because IoT data may pass through multiple distributed nodes, security frameworks with **end-to-end encryption** and **lightweight cryptography** are essential [17].

5.2 Trust Management in Hybrid Environments

Hybrid models introduce new trust challenges as data moves between edge nodes and cloud servers. Blockchain-based frameworks can ensure transparency and integrity [18].

5.3 Access Control and Authentication

Role-based and attribute-based access control models have been adapted for resource-constrained edge environments [19].

6. Applications of Hybrid Edge–Cloud Models

6.1 Smart Cities

Traffic management, environmental tracking, and public safety systems need real-time responsiveness and massive data analytics—enabled by hybrid models [20].

6.2 Healthcare IoT

Hybrid systems allow real-time patient monitoring at the edge while maintaining historical data and analytics in the cloud [21].

6.3 Industrial IoT (IIoT)

Predictive maintenance, robotics, and automation require low latency and high reliability, making hybrid processing ideal [22].

6.4 Autonomous Vehicles

Edge nodes support ultra-fast decision-making, while cloud servers handle global mapping and large-scale AI training [23].

7. Challenges and Future Research Opportunities

7.1 Standardization Issues

There is no universal standard for hybrid architectures or interfaces, causing interoperability issues across vendors and platforms [24].

7.2 Scalability Concerns

Edge nodes, being resource constrained, struggle to support increasing device density. Distributed orchestration frameworks are needed [25].

7.3 Energy Optimization

Energy-aware hybrid frameworks are required for sustainable IoT deployment, especially in battery-powered devices [26].

7.4 AI Automation and Self-Managing Systems

Future hybrid models will leverage autonomous orchestration, self-healing networks, and intelligent routing to optimize performance dynamically [27].

8. Conclusion

Hybrid edge–cloud computing has emerged as a powerful and flexible architecture for modern IoT systems by combining the strengths of cloud and edge computing. It effectively addresses latency, bandwidth, energy efficiency, security, and scalability challenges faced by traditional cloud-based IoT frameworks. This paper presented a comprehensive review of hybrid models, system architectures, enabling technologies, scheduling mechanisms, security issues, applications, and open research challenges. The study highlights the growing importance of hybrid computing in enabling next-generation intelligent IoT services. Future advancements in AI, networking, and security technologies are expected to further strengthen hybrid edge–cloud computing as a backbone for smart digital ecosystems.

References

1. Satyanarayanan, M. "The emergence of edge computing." *IEEE Computer*, 2017.
2. Shi, W. et al. “Edge Computing: Vision and Challenges.” *IEEE IoT Journal*, 2016.
3. Bonomi, F. et al. " Fog computing and its role in IoT." MCC *Workshop*, 2012.
4. Chiang, M., Zhang, T. “Fog and IoT: An Overview.” *IEEE Internet of Things Journal*, 2016.
5. Roman, R. et al. “Mobile edge computing in smart environments.” *Future Generation Computer Systems*, 2018.
6. Yi, S. et al. "Fog computing introduction." *ACM MCC*, 2015.
7. Mukherjee, M. et al. “Fog computing for IoT healthcare.” *IEEE Communications*, 2018.
8. Mao, Y. et al. "Mobile edge computing offloading." *IEEE Communications Surveys*, 2017.
9. Chen, X. “RL-based task offloading.” *IEEE Transactions on Mobile Computing*, 2019.
10. Tuli, S. et al. “Multi-objective offloading in fog.” *Journal of Network and Computer Applications*, 2020.
11. Deng, R. et al. “Latency and resource optimization.” *IEEE IoT Journal*, 2016.
12. Kairouz, P. et al. “Federated Learning Advances.” *Foundations & Trends in ML*, 2021.
13. Morabito, R. “Containerized cloudlet applications.” *IEEE Cloud Computing*, 2017.
14. Zhang, H. et al. “5G and MEC for IoT.” *IEEE Network*, 2019.
15. Taleb, T. et al. “Multi-access edge orchestration.” *IEEE Communications*, 2017.
16. Xu, X. et al. “AI for edge–cloud optimization.” *IEEE Transactions on Industrial Informatics*, 2021.
17. Alrawais, A. et al. “IoT security challenges.” *IEEE Access*, 2017.
18. Dorri, A. et al. “Blockchain for IoT.” *IEEE IoT Journal*, 2017.
19. Ouaddah, A. et al. “Access control in IoT.” *IEEE Communications Surveys*, 2017.

20. Zanella, A. et al. “Smart city IoT systems.” *IEEE IoT Journal*, 2014.
21. Aceto, G. et al. “IoT in healthcare.” *IEEE Communications Surveys*, 2020.
22. Lee, J. et al. “IIoT and predictive maintenance.” *Manufacturing Letters*, 2017.
23. Chen, T. et al. “Edge–cloud architecture for autonomous vehicles.” *IEEE Vehicular Technology*, 2020.
24. Rahman, A., et al. “Standardization in fog and edge computing.” *ITU Journal*, 2019.
25. Chiang, M. “Fog scalability challenges.” *Proceedings of the IEEE*, 2016.
26. Sookhak, M. et al. “Energy-aware IoT edge computing.” *Journal of Network and Computer Applications*, 2019.
27. Zhou, Z. et al. “Self-managing IoT edge systems.” *IEEE Transactions on Network & Service Management*, 2021.