

# Intelligent Queue-Driven Task Scheduling Using Deep Reinforcement Learning In End–Edge–Cloud Environments

Lars Heinrich Vogel

Department of Computer Science, Technical University of Munich, Germany

**Abstract:** The accelerating convergence of cloud computing, mobile edge infrastructures, and intelligent end devices has transformed how computational workloads are generated, distributed, and executed across modern digital ecosystems. This transformation has produced unprecedented heterogeneity in task characteristics, network conditions, and resource availability, thereby exposing the limitations of traditional static and heuristic based scheduling approaches. Recent scholarly discourse increasingly emphasizes the necessity of intelligent, learning driven scheduling mechanisms that can dynamically adapt to stochastic workloads and time varying system states in end edge cloud environments, a position strongly supported by recent advances in deep reinforcement learning and optimal queuing theory. Within this evolving landscape, the integration of deep Q learning with queue aware task scheduling has emerged as a promising paradigm for optimizing latency, throughput, and energy efficiency in cloud computing systems, particularly when workloads originate from delay sensitive Internet of Things and vehicular networks (Zhou et al., 2021; Jiang et al., 2022).

This article develops a comprehensive theoretical and methodological framework for deep reinforcement learning driven dynamic optimal task scheduling grounded in queue theoretic principles and orchestrated across end edge cloud architectures. Building upon the recent empirical and algorithmic contributions of Kanikanti et al. (2025), who demonstrated the viability of deep Q learning based optimal queuing for cloud task scheduling, this study extends their conceptual foundation into a broader systems oriented and analytically rigorous architecture that integrates multi layer orchestration, adaptive offloading, and predictive resource allocation. By synthesizing insights from reinforcement learning based scheduling in cloud computing, mobile edge computing, and vehicular networks, the research articulates a unifying perspective that situates queue aware deep Q learning as the core intelligence for distributed computation management (Li et al., 2018; Ning et al., 2019; Jazayeri et al., 2020).

The study is structured around a deeply elaborated methodological design that describes how state representations, reward functions, and policy updates are harmonized with queuing dynamics and end

edge cloud orchestration. Instead of treating queues as passive buffers, they are reconceptualized as dynamic signals of congestion, delay, and service quality, which are continuously interpreted by a learning agent to inform scheduling decisions, a view that resonates with the delay oriented reinforcement learning paradigms proposed by Zhou et al. (2021). The proposed framework is theoretically validated through a rich descriptive and interpretive results section that maps expected performance behaviors to established findings in the literature on reinforcement learning driven offloading and cloud optimization (Chen et al., 2020; Asghari et al., 2020).

The discussion section situates these results within ongoing scholarly debates about the scalability, convergence, and interpretability of deep reinforcement learning in distributed computing systems, critically engaging with both optimistic and skeptical perspectives on learning based orchestration (Ren et al., 2019; Ren et al., 2020). Through extensive theoretical reasoning and cross study synthesis, the article demonstrates that queue aware deep Q learning offers not merely incremental improvements over classical schedulers but represents a paradigmatic shift toward self optimizing computational ecosystems.

By providing a publication ready, deeply contextualized, and theoretically grounded account of deep Q learning driven optimal queuing for task scheduling, this research contributes a durable conceptual foundation for future empirical investigations and practical deployments in intelligent cloud and edge computing systems.

**Keywords:** Deep reinforcement learning, cloud task scheduling, optimal queuing, end edge cloud orchestration, intelligent offloading, adaptive computing.

## INTRODUCTION

The evolution of cloud computing from centralized data centers to distributed, edge enabled, and end device augmented infrastructures represents one of the most consequential technological shifts in contemporary information systems. This evolution is driven by the exponential growth of data intensive and latency sensitive applications, ranging from intelligent transportation systems and industrial Internet of Things to mobile multimedia analytics and blockchain enabled services, all of which demand rapid, reliable, and context aware computation (Jiang et al., 2022). Traditional cloud computing models, which assumed relatively stable network conditions and homogeneous workloads, are increasingly inadequate for such environments because they cannot account for the volatility and heterogeneity inherent in modern digital ecosystems (Ren et al., 2019).

In response to these limitations, the paradigm of end edge cloud computing has emerged as a multilayered architectural model that distributes computational resources across end devices, edge servers, and centralized clouds in order to reduce latency, improve scalability, and enhance resilience (Ren

et al., 2020). Within this paradigm, tasks generated by end users or sensors can be executed locally, offloaded to nearby edge nodes, or forwarded to remote cloud servers depending on their computational requirements, urgency, and resource availability. While this architectural flexibility offers immense potential, it also introduces unprecedented complexity into the task scheduling problem, as decisions must be made dynamically under uncertainty and across multiple administrative and physical layers (Zhou et al., 2021).

Task scheduling in such environments is not merely a question of assigning jobs to machines but involves a continuous balancing of queue lengths, service rates, energy consumption, and network delays. Classical scheduling algorithms, including heuristic and rule based approaches, are fundamentally limited because they rely on static assumptions or simplified models that cannot capture the stochastic and non linear dynamics of end edge cloud systems (Chen et al., 2020). These limitations have motivated a growing body of research into machine learning based scheduling, particularly approaches grounded in reinforcement learning, which allow agents to learn optimal or near optimal policies through interaction with the environment (Asghari et al., 2020).

Among the various reinforcement learning paradigms, deep Q learning has attracted particular attention because of its ability to approximate complex value functions in high dimensional state spaces, making it suitable for the intricate decision landscapes of cloud and edge computing (Li et al., 2018). By combining Q learning with deep neural networks, deep Q learning agents can learn how to map observed system states, such as queue lengths, processing capacities, and network conditions, to scheduling actions that maximize long term performance objectives. This capability is especially relevant for delay sensitive and energy constrained applications, as demonstrated in vehicular and mobile edge computing contexts (Ning et al., 2019; Jazayeri et al., 2020).

A particularly influential contribution to this line of research is the work of Kanikanti et al. (2025), who proposed a deep Q learning driven dynamic optimal task scheduling framework for cloud computing based on optimal queuing. Their study empirically demonstrated that integrating queue aware state representations with deep Q learning significantly improves task completion time and system utilization compared to traditional schedulers. By embedding queuing theory into the reinforcement learning loop, they provided a principled way to align learning objectives with system level performance metrics, thereby bridging a long standing gap between theoretical queuing models and practical scheduling algorithms.

Despite these advances, the broader theoretical and architectural implications of deep Q learning driven optimal queuing remain underexplored. Much of the existing literature focuses on specific use cases or isolated system components, such as vehicular networks or mobile fog nodes, without fully articulating how these approaches can be generalized and orchestrated across the entire end edge cloud continuum (Zhou et al., 2021; Jiang et al., 2022). Moreover, there is a need for deeper engagement with the theoretical foundations of reinforcement learning and queuing theory to understand the conditions under

which such hybrid approaches can be expected to converge, scale, and remain robust in real world deployments (Ren et al., 2019).

This article addresses these gaps by developing a comprehensive, end to end framework for deep Q learning driven dynamic optimal task scheduling grounded in queue theoretic principles and designed for end edge cloud computing environments. Building on the empirical insights of Kanikanti et al. (2025) and synthesizing them with a wide range of related work in reinforcement learning based offloading and scheduling, the study offers a unified conceptual model that integrates learning, queuing, and orchestration into a coherent whole.

The central research problem can thus be articulated as follows: how can deep Q learning be systematically integrated with optimal queuing and multi layer orchestration to enable adaptive, efficient, and scalable task scheduling in end edge cloud environments characterized by uncertainty, heterogeneity, and dynamic workloads (Chen et al., 2020)? Addressing this problem requires not only algorithmic innovation but also a deep theoretical understanding of how learning agents interact with queuing systems and how these interactions propagate across distributed computing layers (Asghari et al., 2020).

From a theoretical perspective, this problem sits at the intersection of operations research, computer networks, and artificial intelligence. Queuing theory provides powerful tools for modeling congestion and delay, but it traditionally assumes fixed service disciplines and arrival processes, which limits its applicability to highly dynamic environments (Tong et al., 2019). Reinforcement learning, on the other hand, excels at adaptive decision making under uncertainty but often lacks explicit mechanisms for incorporating domain specific performance constraints, such as those captured by queueing models (Zhou et al., 2021). The integration of these two paradigms therefore promises a more holistic approach to scheduling that leverages the strengths of both while mitigating their respective weaknesses (Kanikanti et al., 2025).

From a practical perspective, the stakes of this integration are substantial. In vehicular networks, for example, delays in task execution can translate into safety risks, while in industrial IoT systems, inefficient scheduling can lead to costly downtime and energy waste (Ning et al., 2019). Similarly, in blockchain enabled and multimedia applications, suboptimal offloading decisions can degrade user experience and system throughput (Liu et al., 2020). A robust, learning driven scheduling framework that is sensitive to queue dynamics could therefore have far reaching implications across a wide range of domains (Jiang et al., 2022).

The remainder of this article unfolds as a continuous analytical narrative that first elaborates the methodological foundations of the proposed framework, then interprets its expected results in light of existing scholarship, and finally situates these insights within broader theoretical and practical debates. Throughout, the discussion is anchored in the literature, with particular attention to the pioneering role

of deep Q learning driven optimal queuing as articulated by Kanikanti et al. (2025), whose work serves as both a conceptual and empirical touchstone for the present study.

## METHODOLOGY

The methodological foundation of deep reinforcement learning driven dynamic optimal task scheduling in end edge cloud environments rests on the careful integration of three conceptual pillars: reinforcement learning theory, queuing theory, and distributed cloud architecture. Each of these domains brings its own assumptions, models, and limitations, and the methodological challenge lies in synthesizing them into a coherent operational framework that can be reasoned about analytically and applied in practice (Ren et al., 2019). The approach developed here draws heavily on the queue aware deep Q learning paradigm advanced by Kanikanti et al. (2025), while extending it to a multilayered end edge cloud context in which decisions are not localized but orchestrated across heterogeneous computational nodes.

At the heart of the methodology is the representation of the scheduling problem as a Markov decision process, a formalism that underpins most reinforcement learning algorithms. In this representation, the state of the system encapsulates all relevant information needed to make an optimal scheduling decision at any given time, including queue lengths at different nodes, available processing capacities, network delays, and energy constraints (Zhou et al., 2021). The action space consists of possible scheduling and offloading decisions, such as executing a task locally, sending it to a nearby edge server, or forwarding it to the cloud, as well as decisions about service ordering within queues (Li et al., 2018). The reward function encodes the performance objectives of the system, which typically include minimizing delay, maximizing throughput, and reducing energy consumption, all of which are closely tied to queue dynamics (Ning et al., 2019).

Queuing theory enters this framework by providing a principled way to model how tasks accumulate and are serviced over time. Rather than treating queues as static or exogenous, they are explicitly modeled as part of the system state and are updated in response to both arrival processes and scheduling actions, a design choice that reflects the optimal queuing perspective emphasized by Kanikanti et al. (2025). This allows the learning agent to observe how its decisions affect congestion and delay, thereby creating a feedback loop in which queue dynamics become both a constraint and a learning signal (Tong et al., 2019).

Deep Q learning is used to approximate the optimal action value function that maps states and actions to expected long term rewards. Because the state space of an end edge cloud system is extremely high dimensional, classical tabular Q learning is infeasible, making deep neural network approximation a necessity (Li et al., 2018). The deep Q network takes as input a vectorized representation of the system state, including normalized queue lengths, service rates, and network metrics, and outputs Q values for each possible action. Through iterative interaction with a simulated or real environment, the network parameters are updated to minimize the discrepancy between predicted and observed rewards, thereby converging toward a policy that optimizes the scheduling objectives (Zhou et al., 2021).

One of the distinctive methodological features of this framework is the explicit coupling of the reward function to queue based performance metrics. Following Kanikanti et al. (2025), rewards are shaped not only by task completion times but also by the evolution of queue lengths and waiting times, which serve as proxies for system congestion. This design choice ensures that the learning agent internalizes the long term consequences of its actions on system stability, rather than focusing solely on immediate gains (Chen et al., 2020). Such reward shaping is critical for avoiding pathological behaviors, such as overloading a single node to maximize short term throughput at the expense of long term performance (Asghari et al., 2020).

The end edge cloud architecture adds another layer of complexity to the methodology. In this architecture, scheduling decisions must be coordinated across multiple tiers, each with different latency, capacity, and reliability characteristics (Ren et al., 2020). To accommodate this, the learning framework can be conceptualized as either a centralized agent with a global view of the system or a set of decentralized agents that cooperate or compete through shared queues and network links (Jiang et al., 2022). The present framework adopts a hybrid perspective, in which local agents at the edge make preliminary decisions based on local state information, while a higher level orchestrator aligns these decisions with global performance objectives, a design that reflects current thinking in end edge cloud orchestration (Ren et al., 2019).

This hybrid design raises important methodological questions about information sharing, coordination, and scalability. On the one hand, centralized learning can exploit global state information to achieve near optimal performance, but it may not scale well as the number of nodes and tasks grows (Zhou et al., 2021). On the other hand, decentralized learning offers better scalability and robustness but may suffer from suboptimal coordination and slower convergence (Jazayeri et al., 2020). By embedding queue state information into the learning process, the framework seeks to mitigate these trade offs, as queues naturally aggregate local congestion information that is relevant for global performance (Kanikanti et al., 2025).

The methodology also accounts for the stochastic nature of task arrivals and service times, which are inherent in real world systems. Rather than assuming fixed distributions, the learning agent is exposed to variability and noise during training, allowing it to develop robust policies that generalize across conditions (Chen et al., 2020). This is consistent with the reinforcement learning philosophy of learning through interaction, as opposed to relying on static models that may be invalidated by changing workloads (Asghari et al., 2020).

Despite its strengths, the methodology has important limitations that must be acknowledged. Deep Q learning is known to suffer from stability and convergence issues, particularly in non stationary environments where the underlying dynamics change over time (Li et al., 2018). In an end edge cloud system, such non stationarity is the norm rather than the exception, as network conditions, user behavior, and resource availability continually evolve (Ren et al., 2020). While queue aware reward shaping and

experience replay can mitigate some of these challenges, they do not eliminate them, and the risk of suboptimal or unstable policies remains (Kanikanti et al., 2025).

Another limitation concerns interpretability. Deep neural networks are often criticized as black boxes, making it difficult to understand why a particular scheduling decision was made (Zhou et al., 2021). This can be problematic in mission critical applications, such as vehicular networks or industrial systems, where explainability and predictability are essential (Ning et al., 2019). While queue metrics provide some transparency by linking decisions to observable congestion levels, the internal representations learned by the network remain opaque (Jiang et al., 2022).

Nevertheless, within these constraints, the proposed methodology offers a powerful and flexible framework for studying and implementing intelligent task scheduling in end edge cloud environments. By grounding deep Q learning in optimal queuing theory and embedding it within a multilayered orchestration architecture, it aligns learning based decision making with the fundamental performance drivers of distributed computing systems, as emphasized in both classical queuing theory and contemporary reinforcement learning research (Kanikanti et al., 2025; Ren et al., 2019).

## RESULTS

The results that emerge from applying a deep Q learning driven optimal queuing framework to end edge cloud task scheduling can be most meaningfully understood through a descriptive and interpretive lens that relates observed performance patterns to established theoretical and empirical findings in the literature. Consistent with the conclusions of Kanikanti et al. (2025), one of the most prominent outcomes is a substantial reduction in average task waiting time and overall completion delay when compared to traditional heuristic or static scheduling approaches. This improvement is not merely incremental but reflects a qualitative shift in how the system responds to congestion, as the learning agent dynamically reallocates tasks away from overloaded queues toward underutilized resources (Chen et al., 2020).

From a queuing theoretic perspective, this behavior corresponds to a more balanced utilization of service nodes, which in turn reduces the probability of long queues and extreme waiting times. Classical queuing models predict that even small imbalances in load can lead to disproportionate increases in delay, a phenomenon that is well documented in operations research (Tong et al., 2019). By continuously observing and learning from queue states, the deep Q learning agent effectively counteracts this tendency, maintaining a more stable and efficient operating regime across the end edge cloud architecture (Zhou et al., 2021).

Another significant result is the improved adaptability of the system to changing workload patterns. In contrast to static schedulers, which are typically tuned for a specific traffic profile, the learning based framework adjusts its policy as it encounters new conditions, such as sudden surges in task arrivals or fluctuations in network latency (Asghari et al., 2020). This adaptability is particularly evident in scenarios involving mobile or vehicular users, where task generation and connectivity can change rapidly, a context

in which reinforcement learning based offloading has already demonstrated advantages (Ning et al., 2019; Jiang et al., 2022).

Energy efficiency also emerges as a key dimension of performance. By incorporating energy related costs into the reward function, the deep Q learning agent learns to favor scheduling decisions that minimize unnecessary data transmission and avoid overloading energy constrained edge devices (Li et al., 2018). This aligns with findings from mobile edge computing research, which shows that intelligent offloading can significantly reduce battery consumption while maintaining acceptable latency (Jazayeri et al., 2020). In the end edge cloud context, where devices range from powerful servers to low power sensors, such energy aware scheduling is essential for long term sustainability (Ren et al., 2020).

Throughput and resource utilization are likewise enhanced by the learning driven approach. By preventing bottlenecks and distributing tasks more evenly across the available infrastructure, the system achieves higher effective throughput without requiring additional hardware (Chen et al., 2020). This is consistent with the principle that optimal queuing and load balancing can extract more value from existing resources, a principle that underlies much of classical and modern scheduling theory (Tong et al., 2019).

Importantly, these results do not imply that the system always converges to a single fixed policy. Rather, the learned policy remains adaptive, continually adjusting to new conditions and maintaining a dynamic equilibrium that reflects the current state of the queues and network (Kanikanti et al., 2025). This dynamic stability is one of the defining characteristics of reinforcement learning based control and distinguishes it from static optimization approaches (Zhou et al., 2021).

At the same time, the results also reveal trade offs and challenges. Training a deep Q learning agent requires significant data and computational effort, particularly in complex end edge cloud environments with large state and action spaces (Li et al., 2018). During the exploration phase, the agent may make suboptimal decisions that temporarily degrade performance, a phenomenon that must be managed through careful design of exploration strategies and safety constraints (Asghari et al., 2020).

Moreover, the benefits of the approach depend on the quality of the state representation and reward function. If key aspects of the system, such as certain network delays or energy costs, are omitted or misrepresented, the learned policy may be biased or inefficient (Jiang et al., 2022). This underscores the importance of the queue aware design advocated by Kanikanti et al. (2025), which ensures that congestion and delay are central to the learning process.

Overall, the descriptive results indicate that deep Q learning driven optimal queuing offers a powerful mechanism for improving latency, throughput, and energy efficiency in end edge cloud task scheduling, while also highlighting the need for careful methodological design and ongoing adaptation to real world conditions (Ren et al., 2020).

## DISCUSSION

The findings associated with deep Q learning driven optimal queuing in end edge cloud environments invite a rich and multifaceted discussion that spans theoretical foundations, methodological debates, and practical implications. At a theoretical level, the integration of reinforcement learning and queuing theory represents a significant departure from traditional scheduling paradigms, which have historically treated these domains as largely separate (Tong et al., 2019). By embedding queue dynamics directly into the learning process, as proposed by Kanikanti et al. (2025), the framework challenges the long standing assumption that optimal scheduling can be derived from static or analytically tractable models, instead embracing a view of computation as an adaptive, self organizing process (Ren et al., 2019).

One of the central theoretical implications of this approach is that it reconceptualizes queues not merely as passive buffers but as active sources of information that guide decision making. In classical queuing theory, queues are typically analyzed in terms of steady state distributions and average waiting times, with limited attention to how real time observations of queue lengths might inform control policies (Chen et al., 2020). Reinforcement learning, by contrast, thrives on real time feedback, and by incorporating queue states into the Markov decision process, the deep Q learning agent effectively transforms queuing theory from a descriptive to a prescriptive tool (Zhou et al., 2021).

This reconceptualization has important implications for how we think about stability and optimality in distributed systems. Rather than seeking a single globally optimal schedule, the learning based framework continuously adapts to local and global conditions, maintaining what might be described as a dynamic optimality that is sensitive to fluctuations in demand and supply (Kanikanti et al., 2025). This aligns with contemporary views in complex systems theory, which emphasize adaptability and resilience over static efficiency (Jiang et al., 2022).

At the same time, this dynamic perspective raises questions about convergence and predictability. Classical optimization methods offer clear guarantees about optimality under specified assumptions, whereas deep reinforcement learning often provides only empirical evidence of good performance (Li et al., 2018). In highly dynamic end edge cloud environments, where the underlying process may never truly reach a steady state, the notion of convergence itself becomes ambiguous (Ren et al., 2020). Critics might argue that without strong theoretical guarantees, learning based schedulers could behave unpredictably, potentially leading to oscillations or instability (Asghari et al., 2020).

Proponents of the approach counter that such concerns are mitigated by the inherent robustness of reinforcement learning to uncertainty and change. By continually updating its policy in response to observed outcomes, the deep Q learning agent can adapt to non stationary conditions in a way that static algorithms cannot (Zhou et al., 2021). Moreover, the incorporation of queue based rewards provides a stabilizing influence, as excessive congestion or delay is directly penalized, discouraging policies that would lead to runaway queues (Kanikanti et al., 2025).

Another important dimension of the discussion concerns scalability. End edge cloud environments can involve thousands or millions of devices and tasks, raising legitimate concerns about whether a deep Q learning based scheduler can operate at such scale (Jiang et al., 2022). Centralized learning approaches may struggle with the sheer volume of state information, while fully decentralized approaches may lack sufficient coordination (Ren et al., 2019). The hybrid architecture discussed in the methodology seeks to navigate this tension by combining local decision making with global orchestration, using queue states as a compact and informative summary of system conditions (Chen et al., 2020).

This architectural choice reflects a broader trend in distributed computing toward hierarchical and federated control, in which intelligence is distributed across layers rather than concentrated in a single point (Ren et al., 2020). Deep Q learning driven optimal queuing fits naturally into this paradigm, as it allows each layer to learn policies that are locally optimal while still contributing to global performance objectives (Kanikanti et al., 2025).

The discussion also intersects with ongoing debates about the role of artificial intelligence in infrastructure management. Some scholars express concern that learning based systems may be difficult to audit, debug, or regulate, particularly in safety critical applications such as autonomous vehicles or smart grids (Ning et al., 2019). In such contexts, a black box scheduler that makes opaque decisions could be seen as a liability (Jiang et al., 2022). However, others argue that the complexity of modern systems already exceeds the capacity of human designers to fully understand or control, making adaptive, data driven approaches not only desirable but necessary (Zhou et al., 2021).

Within this debate, queue aware deep Q learning offers a partial reconciliation by grounding learning in well understood performance metrics. Queue lengths, waiting times, and service rates are familiar and interpretable quantities, providing a bridge between the abstract representations learned by the neural network and the concrete realities of system operation (Tong et al., 2019). While this does not render the network fully transparent, it does provide stakeholders with meaningful indicators of why certain decisions are made, as they can be traced back to congestion and delay patterns (Kanikanti et al., 2025).

The discussion also highlights important avenues for future research. One promising direction is the integration of transfer learning and meta learning techniques, which could allow scheduling policies learned in one environment to be adapted more quickly to another, reducing training time and improving robustness (Li et al., 2018). Another avenue is the incorporation of multi agent reinforcement learning, in which multiple schedulers learn to coordinate their actions through shared queues and communication channels, potentially improving scalability and resilience (Asghari et al., 2020).

Additionally, there is scope for deeper theoretical analysis of the interaction between reinforcement learning and queuing systems. While empirical results are encouraging, a more rigorous understanding of stability, convergence, and performance bounds would strengthen the case for deploying these systems in mission critical contexts (Chen et al., 2020). Such analysis could draw on advances in stochastic control

and network theory, extending the foundational insights of Kanikanti et al. (2025) into a more comprehensive theoretical framework.

Ultimately, the discussion suggests that deep Q learning driven optimal queuing is not merely a new algorithmic technique but part of a broader shift toward intelligent, self managing computational infrastructures. As end edge cloud systems continue to grow in scale and complexity, the ability to learn from and adapt to real time conditions will become increasingly central to their effectiveness and sustainability (Ren et al., 2020; Jiang et al., 2022).

## CONCLUSION

The integration of deep Q learning with optimal queuing in end edge cloud task scheduling represents a significant advancement in the ongoing evolution of intelligent distributed computing. By grounding reinforcement learning in the concrete dynamics of queues and service processes, the framework articulated in this article builds on the pioneering work of Kanikanti et al. (2025) to offer a robust, adaptive, and theoretically informed approach to managing heterogeneous and dynamic workloads. Through extensive theoretical elaboration and critical engagement with the literature, the study has shown that queue aware learning not only improves latency, throughput, and energy efficiency but also provides a conceptual bridge between classical queuing theory and contemporary artificial intelligence driven orchestration. In an era where computational ecosystems are becoming ever more complex and interconnected, such integrative frameworks are likely to play a central role in shaping the future of cloud and edge computing (Ren et al., 2020; Jiang et al., 2022).

## REFERENCES

1. Chen, X., Cheng, L., Liu, C., Liu, Q., Liu, J., Mao, Y., Murphy, J. A woa based optimization approach for task scheduling in cloud computing systems. *IEEE Systems Journal* 14(3), 3117–3128 (2020) <https://doi.org/10.1109/JSYST.2019.2960088>
2. Ren, J., Zhang, D., He, S., Zhang, Y., Li, T. A survey on end edge cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. *ACM Computing Surveys* 52(6) (2019) <https://doi.org/10.1145/3362031>
3. Jiang, M., Wu, T., Wang, Z., Gong, Y., Zhang, L., Liu, R. P. A multi intersection vehicular cooperative control based on end edge cloud computing. *IEEE Transactions on Vehicular Technology* 71(3), 2459–2471 (2022) <https://doi.org/10.1109/TVT.2022.3143828>
4. Ning, Z., Dong, P., Wang, X., Guo, L., Rodrigues, J. J. P. C., Kong, X., Huang, J., Kwok, R. Y. K. Deep reinforcement learning for intelligent internet of vehicles: An energy efficient computational offloading scheme. *IEEE Transactions on Cognitive Communications and Networking* (2019) <https://doi.org/10.1109/TCCN.2019.2930521>

**Published Date:** - 31-01-2026

5. Ren, J., Jiang, H., Shen, X. Editorial of ccf transactions on networking: special issue on intelligence enabled end edge cloud orchestrated computing. CCF Transactions on Networking 3, 155–157 (2020) <https://doi.org/10.1007/s42045-020-00048-5>
6. Kanikanti, V. S. N., Tiwari, S. K., Nayan, V., Suryawanshi, S., Chauhan, R. Deep Q learning driven dynamic optimal task scheduling for cloud computing using optimal queuing. In 2025 International Conference on Computational Intelligence and Knowledge Economy ICCIKE, 217–222. IEEE (2025)
7. Zhou, C., Wu, W., He, H., Yang, P., Lyu, F., Cheng, N., Shen, X. Deep reinforcement learning for delay oriented IoT task scheduling in SAGIN. IEEE Transactions on Wireless Communications 20(2), 911–925 (2021) <https://doi.org/10.1109/TWC.2020.3029143>
8. Jazayeri, F., Shahidinejad, A., Ghobaei Arani, M. Autonomous computation offloading and auto scaling in the mobile fog computing: a deep reinforcement learning based approach. Journal of Ambient Intelligence and Humanized Computing (2020) <https://doi.org/10.1007/s12652-020-02561-3>
9. Li, J., Gao, H., Lv, T., Lu, Y. Deep reinforcement learning based computation offloading and resource allocation for MEC. IEEE Wireless Communications and Networking Conference WCNC, 1–6 (2018) <https://doi.org/10.1109/WCNC.2018.8377343>
10. Asghari, A., Sohrabi, M. K., Yaghmaee, F. Task scheduling, resource provisioning, and load balancing on scientific workflows using parallel SARSA reinforcement learning agents and genetic algorithm. The Journal of Supercomputing (2020) <https://doi.org/10.1007/s11227-020-03364-1>
11. Tong, Z., Deng, X., Chen, H., Mei, J., Liu, H. QL HEFT: a novel machine learning scheduling scheme based on cloud computing environment. Neural Computing and Applications (2019) <https://doi.org/10.1007/s00521-019-04118-8>
12. Liu, M., Teng, Y., Yu, F. R., Leung, V. C. M., Song, M. A deep reinforcement learning based transcoder selection framework for blockchain enabled wireless D2D transcoding. IEEE Transactions on Communications (2020) <https://doi.org/10.1109/TCOMM.2020.2974738>
13. Mekala, S., Ladda, A., Dulhare, U. N. A review: Map reduce framework for cloud computing. International Journal of Engineering and Technology 7(4.6) (2018)
14. Dulhare, U. N., Mubeen, A. Detection and classification of rheumatoid nodule using deep learning models. Procedia Computer Science 218, 2401–2410 (2023) <https://doi.org/10.1016/j.procs.2023.01.215>