

## Queuing-Integrated Deep Reinforcement Learning For Adaptive Task Scheduling In Cloud Data Centers

**Dr. Adrian Keller**

Technical University of Munich, Germany

Article Received: 05/12/2025, Article Revised: 25/12/2025, Article Accepted: 10/01/2026, Article Published: 25/01/2026

© 2026 Authors retain the copyright of their manuscripts, and all Open Access articles are disseminated under the terms of the [Creative Commons Attribution License 4.0 \(CC-BY\)](https://creativecommons.org/licenses/by/4.0/), which licenses unrestricted use, distribution, and reproduction in any medium, provided that the original work is appropriately cited.

---

### ABSTRACT

The accelerating digitalization of economic, industrial, and social systems has rendered cloud computing the backbone of contemporary information infrastructure. Yet, the unprecedented growth in computational demand, heterogeneity of workloads, and volatility of user requirements have exposed deep limitations in classical task scheduling and resource management paradigms. Static or heuristics-based schedulers, which historically dominated cloud environments, are increasingly unable to cope with highly dynamic and stochastic workloads, fluctuating service-level requirements, and the imperative for energy-efficient operations. This study advances a comprehensive theoretical and analytical investigation of deep reinforcement learning–driven dynamic task scheduling in cloud computing, with particular emphasis on queuing-aware optimal decision making. Building on the methodological foundation established by Kanikanti et al. (2025), who demonstrated the effectiveness of deep Q-learning combined with optimal queuing theory for cloud task scheduling, this research situates their contribution within a broader interdisciplinary framework that spans energy-aware systems, multi-agent learning, and cyber-physical digital twins. The article develops a unifying perspective that integrates insights from reinforcement learning theory, stochastic queuing models, energy management in cyber-physical systems, and adaptive control of autonomous agents. By synthesizing developments in microgrid energy management, underwater robotics, autonomous vehicle control, and digital twin–based production systems, the study demonstrates that the core challenge of cloud scheduling is not merely computational efficiency but the orchestration of learning-driven decisions across uncertain, delayed, and resource-constrained environments. In this sense, cloud data centers resemble complex adaptive systems in which computing tasks compete for shared resources in a manner analogous to energy flows in microgrids or coordinated actions in multi-robot systems.

Methodologically, the research adopts a text-based analytical design that combines formal reinforcement learning principles derived from Markov decision processes with queuing-theoretic interpretations of cloud workloads. The deep Q-learning framework of Kanikanti et al. (2025) is critically analyzed and extended conceptually through comparative evaluation against SARSA-based, actor–critic, and deep deterministic policy gradient approaches reported in the broader literature. Particular attention is devoted to how state abstraction, reward shaping, and queue length feedback enable schedulers to balance latency, throughput, and energy consumption simultaneously.

The results of this study are presented in a descriptive and interpretive manner grounded in the comparative literature. They indicate that deep Q-learning–based dynamic schedulers consistently outperform rule-based and shallow reinforcement learning approaches in terms of adaptive responsiveness, queue stability, and energy-aware decision making, as supported by studies in cloud computing, microgrids, and robotic coordination. The discussion further reveals that queuing-informed deep reinforcement learning architectures provide a theoretically robust mechanism for mitigating congestion collapse, improving quality of service, and aligning cloud operations with sustainability goals.

By offering an extensive theoretical elaboration and critical synthesis of existing research, this article contributes a unified conceptual framework for understanding and advancing learning-driven cloud task scheduling. It concludes that the convergence of deep reinforcement learning and optimal queuing theory, as exemplified by Kanikanti et al. (2025), represents not a marginal technical improvement but a paradigm shift in how future cloud ecosystems will be designed, governed, and optimized.

## KEYWORDS

Deep Q-learning, cloud task scheduling, optimal queuing, energy-aware computing, reinforcement learning, dynamic resource management.

## INTRODUCTION

Cloud computing has evolved from a relatively simple paradigm of remote resource provisioning into a globally distributed, highly heterogeneous, and deeply interconnected digital ecosystem. Modern cloud infrastructures support everything from social media platforms and financial trading systems to autonomous vehicle coordination and large-scale scientific workflows. This dramatic expansion has generated unprecedented challenges for task scheduling, resource provisioning, and energy management, because cloud data centers must simultaneously satisfy performance guarantees, economic constraints, and sustainability requirements. Classical scheduling algorithms, which rely on fixed heuristics or deterministic optimization, were developed for comparatively stable and predictable environments and thus struggle to cope with the volatility, scale, and complexity of present-day cloud workloads (Ding et al., 2020).

The emergence of reinforcement learning has transformed how researchers conceptualize decision making under uncertainty in complex systems. Reinforcement learning, as formalized through Markov decision processes, enables an agent to learn optimal policies by interacting with an environment and receiving feedback in the form of rewards (Puterman, 1994; Sutton and Barto, 1998). In the context of cloud computing, the environment consists of servers, virtual machines, queues of tasks, and fluctuating user demands, while the agent corresponds to a scheduler that must decide how to allocate tasks across available resources. Unlike static heuristics, reinforcement learning-based schedulers can adapt to changing conditions, discover non-intuitive strategies, and continuously improve performance over time (Guo et al., 2020).

Recent years have witnessed a convergence of reinforcement learning and deep learning, giving rise to deep reinforcement learning architectures that can handle high-dimensional state spaces and complex, nonlinear dynamics. Deep Q-learning, in particular, has emerged as a powerful approach for approximating action-value functions using neural networks, enabling agents to learn sophisticated control policies in domains ranging from robotics to energy systems (Carlucho et al., 2018; Ariche et al., 2025). Within cloud computing, deep Q-learning has been applied to dynamic task scheduling, load balancing, and energy-efficient resource management, with promising results that consistently outperform traditional methods (Zhu et al., 2020; Ding et al., 2020).

A particularly significant contribution in this trajectory is

the work of Kanikanti et al. (2025), who proposed a deep Q-learning-driven dynamic optimal task scheduling framework explicitly grounded in optimal queuing theory. Their approach recognizes that cloud workloads are fundamentally queuing systems in which tasks arrive stochastically, wait for service, and experience delays depending on system congestion. By integrating queue length and waiting time information into the state representation and reward structure of a deep Q-learning agent, Kanikanti et al. (2025) demonstrated that scheduling decisions could be made not only with respect to immediate resource availability but also with foresight into congestion dynamics and long-term system stability.

The importance of this queuing-aware perspective cannot be overstated. In classical cloud scheduling, tasks are often treated as independent jobs whose placement decisions can be optimized in isolation. However, queuing theory teaches that the performance of a system emerges from the collective interaction of tasks competing for shared resources, and that small changes in arrival or service rates can lead to nonlinear and sometimes catastrophic effects on delay and throughput. By embedding queuing information into a learning-based scheduler, the approach of Kanikanti et al. (2025) aligns cloud computing with decades of research in stochastic systems and operations research, thereby providing a theoretically grounded and practically robust solution to dynamic scheduling.

The relevance of reinforcement learning-based control extends far beyond cloud computing. In energy systems, for example, Q-learning and its variants have been used to manage microgrids, electric vehicles, and hybrid power systems by learning optimal strategies for balancing supply, demand, and storage under uncertainty (Ramesh et al., 2025; Xu and Li, 2025). In robotics, deep reinforcement learning has enabled autonomous underwater vehicles, multi-robot cargo systems, and unmanned aerial vehicles to coordinate their actions in complex, dynamic environments (Sufan and Troni, 2025; Rybak et al., 2020; Liang et al., 2023). These domains share a common structure: a set of agents or components must make sequential decisions in the presence of uncertainty, delayed feedback, and resource constraints.

Cloud data centers can therefore be understood as cyber-physical systems whose computational resources are analogous to energy or mechanical resources in physical systems. Just as a microgrid must allocate power to competing loads while maintaining stability, a cloud scheduler must allocate CPU, memory, and network bandwidth to competing tasks while maintaining quality

of service and minimizing energy consumption (Ariche et al., 2025). The analogy becomes even more compelling when one considers the rise of digital twins, which replicate physical systems in virtual environments and use reinforcement learning to optimize operations, as in vertical farm production systems (Luo and Ball, 2025). Cloud computing not only hosts such digital twins but also embodies similar dynamics within its own infrastructure.

Despite this rich body of research, significant gaps remain in the theoretical integration of queuing theory and deep reinforcement learning for cloud scheduling. Many existing studies focus on algorithmic performance metrics such as makespan or throughput without adequately accounting for the stochastic nature of task arrivals and service times (Asghari et al., 2020; Wang et al., 2025). Others emphasize energy efficiency but treat scheduling and power management as separate problems, even though they are deeply interdependent (Mohan Sharma and Garg, 2020; Li, 2023). The work of Kanikanti et al. (2025) offers a promising pathway toward unification, yet its implications for broader cloud architectures, energy-aware computing, and multi-agent coordination have not been fully explored.

This article addresses this gap by providing an extensive theoretical and analytical examination of deep Q-learning-based dynamic task scheduling within queuing-aware cloud environments. Drawing on a diverse set of references from cloud computing, energy systems, robotics, and control theory, the study seeks to develop a unified conceptual framework that explains why and how reinforcement learning-driven schedulers outperform traditional approaches, and under what conditions their advantages are most pronounced. In doing so, it aims not merely to summarize existing results but to engage in critical scholarly debate, exploring counterarguments, limitations, and future directions.

The central research problem can be articulated as follows: how can cloud computing systems be designed to autonomously and adaptively schedule tasks in a manner that simultaneously optimizes performance, stability, and energy efficiency in the face of stochastic demand and heterogeneous resources? The answer, as argued throughout this article, lies in the integration of deep reinforcement learning with optimal queuing theory, as exemplified by Kanikanti et al. (2025) and supported by a growing body of interdisciplinary research. By situating cloud scheduling within this broader intellectual landscape, the study provides both a rigorous theoretical foundation and a compelling vision for the future of intelligent cloud infrastructures.

## METHODOLOGY

The methodological orientation of this research is grounded in analytical synthesis rather than empirical

experimentation, reflecting the objective of developing a comprehensive theoretical framework for deep reinforcement learning-based cloud task scheduling. The methodological design is therefore informed by the principles of comparative literature analysis, conceptual modeling, and cross-domain theoretical integration, all of which are widely used in advanced research on complex adaptive systems (Puterman, 1994; Sutton and Barto, 1998). Within this framework, the methodological core is the formalization of cloud task scheduling as a Markov decision process enriched by queuing-theoretic constructs, following the conceptual pathway established by Kanikanti et al. (2025).

At the heart of this methodology lies the representation of a cloud computing environment as a stochastic decision system in which tasks arrive according to probabilistic processes, are placed in queues associated with different servers or virtual machines, and are processed with service times that vary depending on resource characteristics and workload intensity. In classical queuing theory, such systems are analyzed through arrival rates, service rates, queue lengths, and waiting time distributions. However, queuing theory alone does not prescribe how an intelligent agent should make real-time scheduling decisions in a dynamic environment. Reinforcement learning fills this gap by providing a mechanism through which a scheduler can learn to map observed system states to optimal actions that maximize long-term reward (Ding et al., 2020).

The deep Q-learning paradigm adopted by Kanikanti et al. (2025) extends this framework by using neural networks to approximate the action-value function, which estimates the expected cumulative reward of taking a given action in a given state. Methodologically, this involves defining a state space that captures relevant aspects of the cloud system, such as current queue lengths, resource utilization levels, and task characteristics, as well as an action space consisting of possible scheduling decisions, such as assigning a task to a particular server or delaying its execution. The reward function is then designed to reflect multiple objectives, including minimizing task completion time, reducing energy consumption, and maintaining system stability, which aligns with the multi-objective nature of cloud management (Mohan Sharma and Garg, 2020; Ramesh et al., 2025).

A key methodological contribution of the queuing-aware approach is the explicit incorporation of queue dynamics into both the state representation and the reward structure. In traditional reinforcement learning applications to cloud scheduling, states often include only coarse-grained metrics such as average CPU utilization or number of active tasks. By contrast, Kanikanti et al. (2025) demonstrated that including detailed queue information enables the agent to anticipate congestion and make proactive decisions that prevent

bottlenecks from forming. This methodological choice is consistent with insights from queuing theory, which emphasize that system performance is highly sensitive to the distribution of waiting tasks and not merely to instantaneous load.

To critically evaluate and contextualize this methodology, the present study compares it with alternative reinforcement learning approaches reported in the literature. SARSA-based algorithms, for example, have been used for cloud scheduling and resource provisioning, often in conjunction with genetic algorithms to enhance exploration and convergence (Asghari et al., 2020). While SARSA offers on-policy learning that can be more stable in certain environments, it is generally less capable of handling large state spaces than deep Q-learning, particularly when combined with function approximation (Qi et al., 2020). Actor-critic methods, such as deep deterministic policy gradient algorithms, have also been applied to scheduling problems in manufacturing and hybrid flow shops, demonstrating strong performance in continuous action spaces (Wang et al., 2025). However, their complexity and sensitivity to hyperparameters can limit their practical deployment in cloud environments, which must operate reliably at scale.

The methodology further draws on analogies from energy management and robotic control to enrich the conceptual model. In microgrid energy management, Q-learning has been used to dynamically adjust power flows based on fluctuating demand and renewable generation, with reward functions that balance cost, stability, and efficiency (Xu and Li, 2025; Ramesh et al., 2025). These studies demonstrate that reinforcement learning agents can successfully learn to manage queues of energy demand and supply in a manner that is mathematically and conceptually analogous to managing queues of computing tasks. Similarly, in autonomous underwater vehicles and multi-robot systems, deep reinforcement learning enables agents to coordinate their actions based on local observations and delayed feedback, highlighting the importance of state representation and reward shaping in complex, distributed systems (Carlucho et al., 2018; Rybak et al., 2020).

A critical methodological consideration in this research is the treatment of uncertainty and non-stationarity. Cloud environments are inherently non-stationary, as user demands, application mixes, and resource availability change over time. Reinforcement learning algorithms, including deep Q-learning, are traditionally formulated under the assumption of a stationary environment, which raises questions about their robustness and convergence properties in real-world clouds. Kanikanti et al. (2025) addressed this challenge by incorporating adaptive learning rates and experience replay mechanisms that allow the agent to continuously update its policy in response to changing conditions. This

methodological strategy is consistent with broader trends in deep reinforcement learning, where techniques such as replay buffers and target networks are used to stabilize learning in non-stationary settings (Sutton and Barto, 1998).

Another methodological dimension concerns the evaluation of scheduling policies. In empirical studies, performance is typically measured through metrics such as average response time, throughput, energy consumption, and makespan. However, in a purely analytical and theoretical study such as the present one, evaluation takes the form of comparative reasoning grounded in the literature. By systematically examining how different reinforcement learning-based schedulers perform across a range of domains, including cloud computing, energy systems, and robotics, the methodology enables a form of triangulation that strengthens the validity of the conclusions (Guo et al., 2020; Ariche et al., 2025).

The limitations of this methodology must also be acknowledged. Because the study relies on secondary sources rather than original experimental data, its conclusions are necessarily constrained by the quality, scope, and comparability of the existing literature. Moreover, differences in experimental setups, workload models, and evaluation metrics across studies can make direct comparisons challenging. Nonetheless, by focusing on underlying theoretical principles rather than specific numerical results, the methodology seeks to identify robust patterns and insights that transcend individual implementations.

In summary, the methodological approach of this research integrates queuing theory, deep reinforcement learning, and cross-domain analogies into a coherent analytical framework for understanding dynamic cloud task scheduling. By building on the queuing-aware deep Q-learning model of Kanikanti et al. (2025) and situating it within a broader interdisciplinary context, the methodology provides a rigorous and comprehensive basis for the subsequent analysis of results and discussion of theoretical implications.

## RESULTS

The results of this study emerge from a systematic and interpretive analysis of the comparative literature on reinforcement learning-based scheduling, energy management, and autonomous control, with particular emphasis on queuing-aware deep Q-learning as articulated by Kanikanti et al. (2025). Across this diverse body of work, a consistent pattern can be observed: learning-driven schedulers that incorporate detailed state representations and feedback mechanisms achieve superior performance, stability, and adaptability compared to static or heuristic-based approaches (Ding et al., 2020; Guo et al., 2020).

One of the most salient results is the demonstrated ability of deep Q-learning-based schedulers to maintain queue stability under highly variable workloads. In cloud computing environments, instability manifests as rapidly growing queues, excessive waiting times, and violations of service-level agreements. Kanikanti et al. (2025) showed that by using queue length and waiting time as part of the state input, their deep Q-learning agent learned to distribute tasks across servers in a way that prevented congestion from concentrating in particular nodes. This result is echoed in other domains where queuing dynamics are critical, such as microgrid energy management, where Q-learning-based controllers have been shown to balance energy flows in a manner that avoids overloading specific components (Ramesh et al., 2025; Xu and Li, 2025).

Another significant result concerns energy efficiency. Traditional cloud schedulers often prioritize performance metrics such as throughput or response time without explicitly considering energy consumption, leading to wasteful over-provisioning and inefficient resource utilization (Mohan Sharma and Garg, 2020). In contrast, reinforcement learning-based approaches can incorporate energy costs directly into their reward functions, enabling a trade-off between performance and sustainability. Studies on battery electric vehicles and microgrids have demonstrated that fuzzy Q-learning and deep reinforcement learning controllers can significantly reduce energy consumption while maintaining operational stability (Ariche et al., 2025; Dong et al., 2024). When these insights are transposed to cloud computing, as in Kanikanti et al. (2025), the result is a scheduler that not only minimizes delay but also reduces the energy footprint of data center operations.

The results also highlight the importance of learning speed and adaptability. Cloud workloads are notoriously dynamic, with sudden spikes in demand, shifts in application mix, and failures of individual components. Deep Q-learning agents, by virtue of their neural network-based function approximators, can generalize across similar states and thus adapt more quickly to new conditions than tabular or shallow reinforcement learning methods (Qi et al., 2020; Wang et al., 2025). Kanikanti et al. (2025) reported that their scheduler converged to effective policies more rapidly than traditional Q-learning approaches, which is consistent with findings in robotics and autonomous vehicle control, where deep reinforcement learning enables faster adaptation to complex environments (Sufan and Troni, 2025; Liang et al., 2023).

A further result pertains to the scalability of learning-based schedulers. As cloud systems grow in size and complexity, the state and action spaces of scheduling problems expand dramatically, posing challenges for any algorithm that relies on explicit enumeration or rule-based logic. Deep Q-learning mitigates this challenge by

using neural networks to approximate value functions, allowing the scheduler to operate effectively in high-dimensional spaces. This scalability has been demonstrated in multi-agent systems for cargo transportation and collaborative search, where reinforcement learning agents coordinate their actions across large and dynamic environments (Rybak et al., 2020; Yu et al., 2021). The queuing-aware deep Q-learning model of Kanikanti et al. (2025) extends this scalability to cloud scheduling by enabling the agent to reason about large numbers of tasks and resources in a unified framework.

The interpretive analysis also reveals that learning-based schedulers are particularly effective in environments characterized by delayed and stochastic feedback. In cloud computing, the consequences of a scheduling decision may not be immediately observable, as tasks take time to execute and their impact on queues and energy consumption unfolds over extended periods. Reinforcement learning, by optimizing long-term cumulative reward rather than immediate outcomes, is well suited to such delayed feedback environments (Sutton and Barto, 1998). Kanikanti et al. (2025) leveraged this property by designing reward functions that reflect not only immediate task completion but also future congestion and energy costs, thereby encouraging the agent to adopt foresighted scheduling strategies.

Finally, the results indicate that queuing-aware reinforcement learning provides a robust mechanism for balancing competing objectives. Cloud schedulers must often trade off between minimizing response time, maximizing throughput, and reducing energy consumption, objectives that can be mutually conflicting. By encoding these objectives into a composite reward function, deep Q-learning agents can learn policies that achieve a nuanced balance tailored to the priorities of the system (Ding et al., 2020; Ariche et al., 2025). The work of Kanikanti et al. (2025) exemplifies this approach, demonstrating that optimal queuing-based rewards lead to more balanced and stable scheduling outcomes than simpler, single-objective formulations.

Taken together, these results support the conclusion that deep Q-learning-driven, queuing-aware task scheduling represents a significant advancement over traditional cloud management approaches. The convergence of evidence from cloud computing, energy systems, and autonomous control suggests that the underlying principles of reinforcement learning and queuing theory are broadly applicable and mutually reinforcing, providing a strong foundation for the next generation of intelligent cloud infrastructures (Kanikanti et al., 2025; Ramesh et al., 2025; Carlucho et al., 2018).

## **DISCUSSION**

The results presented in this study invite a deeper

theoretical interpretation of why and how deep reinforcement learning, particularly in its queuing-aware form, transforms the problem of cloud task scheduling. At a fundamental level, cloud computing systems can be understood as large-scale, stochastic, and distributed decision environments in which the actions of a scheduler propagate through a web of interdependent queues, servers, and workloads. Classical scheduling algorithms implicitly assume that these interactions can be captured through fixed heuristics or deterministic optimization, yet decades of research in queuing theory and operations research have shown that such systems exhibit nonlinear and often counterintuitive dynamics (Puterman, 1994). The integration of deep Q-learning and optimal queuing theory, as demonstrated by Kanikanti et al. (2025), provides a conceptual bridge between the descriptive power of queuing models and the prescriptive power of learning-based control.

One of the central theoretical insights that emerges from this integration is the notion of anticipatory scheduling. In traditional cloud schedulers, decisions are typically based on current system state, such as available CPU capacity or the number of waiting tasks. However, queuing theory emphasizes that the future evolution of a system depends on arrival and service processes, not just instantaneous load. By embedding queue length, waiting time, and service rate information into the state representation of a deep Q-learning agent, Kanikanti et al. (2025) enabled the scheduler to anticipate future congestion and adjust its actions accordingly. This anticipatory capability aligns with the concept of foresighted control in reinforcement learning, where agents optimize long-term reward rather than short-term gain (Sutton and Barto, 1998).

From a scholarly perspective, this represents a significant departure from earlier reinforcement learning applications to cloud computing, which often focused on myopic performance metrics. Ding et al. (2020), for example, demonstrated that Q-learning-based dynamic scheduling could reduce energy consumption, but their state representations were relatively coarse, limiting the agent's ability to reason about complex queuing dynamics. By contrast, the queuing-aware approach of Kanikanti et al. (2025) offers a richer and more theoretically grounded framework, suggesting that the full potential of reinforcement learning in cloud environments can only be realized when it is coupled with domain-specific models of system dynamics.

The analogy with energy management systems further illuminates this point. In microgrids and battery electric vehicles, reinforcement learning agents must allocate energy flows over time while accounting for storage constraints, generation variability, and demand uncertainty (Ariche et al., 2025; Ramesh et al., 2025). These systems, like cloud data centers, can be modeled as networks of queues in which energy or power packets

wait to be processed or delivered. The success of Q-learning and deep reinforcement learning in these domains underscores the generality of the queuing-aware learning paradigm. It suggests that the principles articulated by Kanikanti et al. (2025) are not confined to cloud computing but reflect a broader class of optimal control problems in cyber-physical systems.

Another important dimension of the discussion concerns multi-agent and distributed learning. Modern cloud infrastructures are increasingly decentralized, with tasks and resources spread across geographically distributed data centers and edge nodes. In such environments, a single centralized scheduler may be neither feasible nor desirable. Research on multi-agent reinforcement learning in robotics and autonomous vehicles has shown that distributed agents can learn to coordinate their actions through local interactions and shared rewards (Yu et al., 2021; Liang et al., 2023). The queuing-aware deep Q-learning framework can be extended to such multi-agent cloud architectures, where each data center or cluster operates as an agent that learns to balance its local queues while contributing to global system objectives. The theoretical foundation provided by Kanikanti et al. (2025) thus opens the door to a new generation of distributed, self-organizing cloud schedulers.

However, this optimistic vision must be tempered by a critical examination of limitations and counterarguments. One of the most frequently cited challenges in deep reinforcement learning is the issue of sample efficiency and convergence. Training deep Q-learning agents often requires large amounts of interaction data, which in real-world cloud systems may be costly or risky to obtain. While simulation environments and digital twins can mitigate this problem to some extent, there remains a gap between controlled experimental settings and the messy realities of production clouds (Luo and Ball, 2025). Critics may therefore argue that the practical deployment of queuing-aware deep Q-learning schedulers faces significant barriers in terms of training time, stability, and interpretability.

Moreover, the reliance on neural networks introduces concerns about transparency and trust. In safety-critical or economically sensitive cloud applications, stakeholders may be reluctant to entrust scheduling decisions to opaque models whose internal logic is difficult to interpret. This concern has parallels in robotics and autonomous vehicles, where deep reinforcement learning controllers have demonstrated impressive performance but also unpredictable behavior in edge cases (Carlucho et al., 2018; Sufan and Troni, 2025). Addressing this limitation will require the development of explainable reinforcement learning techniques and robust validation frameworks, an area of active research across multiple domains.

Another potential limitation lies in the assumption of

stationarity or slow non-stationarity. Although Kanikanti et al. (2025) incorporated adaptive mechanisms to handle changing workloads, there may be scenarios in which cloud environments undergo abrupt regime shifts, such as during large-scale cyberattacks, hardware failures, or sudden spikes in user demand. In such cases, a learned policy may become temporarily suboptimal or even harmful. Hybrid approaches that combine reinforcement learning with rule-based safeguards or model-predictive control may therefore be necessary to ensure resilience and safety (Mostafavi and Hakami, 2020).

Despite these challenges, the broader scholarly trajectory strongly supports the continued integration of reinforcement learning and queuing theory in cloud computing. The convergence of evidence from diverse domains, including microgrids, autonomous vehicles, and multi-robot systems, suggests that learning-based control is uniquely suited to managing complex, stochastic, and resource-constrained environments (Dong et al., 2024; Rybak et al., 2020). The work of Kanikanti et al. (2025) stands as a particularly compelling example of how these ideas can be operationalized in the context of cloud task scheduling.

Looking to the future, several promising research directions emerge from this discussion. One avenue involves the incorporation of physiological or human-in-the-loop data, as explored in human-robot teaming studies, to adapt cloud services to user behavior and well-being (Korivand et al., 2024). Another involves the integration of digital twin technologies, which can provide high-fidelity simulations for training and validating reinforcement learning schedulers before deployment in live systems (Luo and Ball, 2025). A third involves the development of hybrid algorithms that combine deep Q-learning with evolutionary optimization or imitation learning to accelerate convergence and improve robustness (Guo et al., 2020; Chen et al., 2024).

In theoretical terms, the most profound implication of this research is the reframing of cloud computing as a learning system rather than a static infrastructure. Just as organisms and societies adapt through feedback and experience, cloud data centers can evolve their scheduling policies through reinforcement learning, guided by queuing-theoretic insights into system dynamics. This paradigm shift challenges traditional notions of systems engineering and invites a more organic, adaptive, and self-optimizing vision of digital infrastructure (Kanikanti et al., 2025; Sutton and Barto, 1998).

## CONCLUSION

This study has advanced a comprehensive and theoretically grounded exploration of deep reinforcement learning-driven dynamic task scheduling in cloud computing, with a particular focus on the integration of

optimal queuing theory as exemplified by Kanikanti et al. (2025). Through an extensive synthesis of research from cloud computing, energy management, robotics, and control theory, the article has demonstrated that queuing-aware deep Q-learning represents not merely an incremental improvement over traditional scheduling methods but a fundamental reorientation of how cloud systems can be designed and governed.

The analysis has shown that by embedding detailed queue dynamics and long-term reward structures into learning-based schedulers, cloud infrastructures can achieve higher levels of stability, efficiency, and adaptability than are possible with static or heuristic approaches. These benefits are not confined to computing alone but resonate with broader trends in cyber-physical systems, where reinforcement learning has proven effective in managing energy flows, coordinating autonomous agents, and optimizing complex production processes. While challenges remain in terms of training efficiency, interpretability, and robustness, the theoretical and empirical trajectory of the field strongly suggests that learning-driven, queuing-aware scheduling will play a central role in the future of cloud computing.

In closing, the convergence of deep reinforcement learning and optimal queuing theory offers a powerful and unifying framework for addressing the multifaceted challenges of modern cloud ecosystems. By embracing this paradigm, researchers and practitioners alike can move toward cloud infrastructures that are not only more performant and sustainable but also more intelligent, resilient, and responsive to the ever-changing demands of the digital world (Kanikanti et al., 2025; Ramesh et al., 2025).

## REFERENCES

1. Ariche, S.; Boulghasoul, Z.; Ouardi, A.E.; Elbacha, A.; Tاجر, A.; Espie, S. Enhancing Energy Management in Battery Electric Vehicles: A Novel Approach Based on Fuzzy Q-Learning Controller. *Eng. Sci. Technol. Int. J.* 2025, 67, 102070.
2. Yu, T.; Huang, J.; Chang, Q. Optimizing task scheduling in human-robot collaboration with deep multiagent reinforcement learning. *Journal of Manufacturing Systems* 2021, 60, 487–499.
3. Liang, Z.B.; Li, Q.; Fu, G.D. Multi-UAV Collaborative Search and Attack Mission Decision-Making in Unknown Environments. *Sensors* 2023, 23, 7398.
4. 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, 2025, 217–222.
5. Ramesh, S.; Sukanth, B.N.; Sathyavarapu, S.J.; Sharma, V.; Kumar, A.A.N.; Khanna, M. Comparative Analysis of Q-Learning, SARSA, and Deep Q-Network for Microgrid Energy Management. *Sci. Rep.* 2025, 15, 694.
  6. Ding, D.; Fan, X.; Zhao, Y.; Kang, K.; Yin, Q.; Zeng, J. Q-learning based dynamic task scheduling for energy-efficient cloud computing. *Future Generation Computer Systems* 2020.
  7. Puterman, M.L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*; John Wiley and Sons: Hoboken, NJ, USA, 1994.
  8. Sutton, R.; Barto, A. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 1998.
  9. Guo, W.; Tian, W.; Ye, Y.; Xu, L.; Wu, K. Cloud Resource Scheduling with Deep Reinforcement Learning and Imitation Learning. *IEEE Internet of Things Journal* 2020.
  10. Carlucho, I.; De Paula, M.; Wang, S.; Petillot, Y.; Acosta, G.G. Adaptive Low-Level Control of Autonomous Underwater Vehicles Using Deep Reinforcement Learning. *Robot. Auton. Syst.* 2018, 107, 71–86.
  11. Luo, Y.; Ball, P. Adaptive Production Strategy in Vertical Farm Digital Twins with Q-Learning Algorithms. *Sci. Rep.* 2025, 15, 15129.
  12. Xu, Y.J.; Li, H. Secondary Voltage Control Strategy for DC Microgrid Based on Reinforcement Learning. *Mech. Electr. Eng. Technol.* 2025, 54, 173–178.
  13. Mohan Sharma; Garg, R. An artificial neural network based approach for energy efficient task scheduling in cloud data centers. Elsevier 2020.
  14. Rybak, L.A.; Behera, L.; Averbukh, M.A.; Sapryka, A.V. Development of an Algorithm for Managing a Multi-Robot System for Cargo Transportation Based on Reinforcement Learning in a Virtual Environment. *IOP Conf. Ser. Mater. Sci. Eng.* 2020, 945, 012083.
  15. Korivand, S.; Galvani, G.; Ajoudani, A.; Gong, J.; Jalili, N. Optimizing Human–Robot Teaming Performance through Q-Learning-Based Task Load Adjustment and Physiological Data Analysis. *Sensors* 2024, 24, 2817.
  16. Sufan, V.; Troni, G. Swim4Real: Deep Reinforcement Learning-Based Energy-Efficient and Agile 6-DOF Control for Underwater Vehicles. *IEEE Robot. Autom. Lett.* 2025, 10, 7326–7333.
  17. 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. *Journal of Supercomputing* 2020.
  18. Chen, D.; Wang, H.; Hu, D.; Xian, Q.; Wu, B. Q-Learning Improved Golden Jackal Optimization Algorithm and Its Application to Reliability Optimization of Hydraulic System. *Sci. Rep.* 2024, 14, 24587.
  19. Dong, X.; Zhang, H.; Xie, X.; Ming, Z. Data-Driven Distributed H Infinity Current Sharing Consensus Optimal Control of DC Microgrids via Reinforcement Learning. *IEEE Trans. Circuits Syst. Regul. Pap.* 2024, 71, 2824–2834.
  20. Li, X.X. Research on Energy Management Strategies for Fuel Cell Hybrid Electric Ships. Wuhan University of Technology, Wuhan, China, 2023.
  21. Wang, J.J.; Zhou, H.M.; Guo, J.; Si, H.W.; Xu, C.; Zhang, M.H.; Zhang, Y.Q.; Zhou, G.X. A Q-Learning-Based Deep Deterministic Policy Gradient Algorithm for the Re-Entrant Hybrid Flow Shop Joint Scheduling Problem with Dual-Gripper. *Eng. Lett.* 2025, 33, 1632–1647.
  22. Wang, X.; Zhu, Q.X.; Zhu, Y.H.; Miao, L.Y. Path Planning for Mobile Robots Based on Improved Q-Learning Algorithm. *Comput. Simul.* 2025, 42, 371–377.
  23. Mostafavi, S.; Hakami, V. A Stochastic Approximation Approach for Foresighted Task Scheduling in Cloud Computing. *Wireless Personal Communications* 2020.
  24. Qi, Q.; Zhang, L.; Wang, J.; Sun, H.; Zhuang, Z.; Liao, J.; Yu, F.R. Scalable Parallel Task Scheduling for Autonomous Driving Using Multi-Task Deep Reinforcement Learning. *IEEE Transactions on Vehicular Technology* 2020.
  25. Zhu, H.; Li, M.; Tang, Y.; Sun, Y. A Deep Reinforcement-Learning-Based Optimization Approach for Real-Time Scheduling in Cloud Manufacturing. *IEEE Access* 2020.