International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

Dynamic Service Migration in Multi-Cloud Architectures Using Proximal Policy Optimization and Reinforcement Learning

[1]kolluru Kavya [2]murukutla Hanumantha Rao [3]koteswara Rao J

[1]*M. Tech Scholar, Department of Computer Science and Engineering (AI&ML), KKR & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur, 522017, India.*
Email: kavyacekolluru@gmail.com
[2]*Assistant Professor, Department of Computer Science and Engineering, KKR & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur, 522017, India.*
Email: hanusmurukutla@gmail.com
[3]*Assistant Professor, KKR & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur, 522017, India.*
Email: Purna2025@gmail.com

Corresponding Author: kavyacekolluru@gmail.com

Abstract: The rapid proliferation of multi-cloud architectures requires efficient dynamic service migration to mitigate the growing complexity of workload distribution, latency, and resource optimization. Traditional approaches often fail to provide robust, scalable solutions because of their limited adaptability, high latency, and inability to balance competing objectives such as cost, downtime, and SLA compliance. These limitations make the need for an intelligent, automated system capable of optimizing in real time the migration of decisions across heterogeneous cloud environments. This paper proposes the advanced framework based on using Reinforcement Learning (RL) for dynamic service migration in multi-cloud setups. Key methodologies used include PPO for stable global migration decisions, MADDPG for collaborative resource allocation, and HRL to break down complex tasks into high- and low-level optimizations. Model adaptability and training efficiency are enhanced using Transfer Learning for fine-tuning pre-trained models using real-world data, whereas Federated Learning ensures that model updates occur securely and collaboratively across clouds. Reward shaping further accelerates convergence by integrating weighted metrics such as latency, migration cost, and resource utilization. The proposed system achieves notable improvements: >99.95% uptime, ≤1.5 seconds migration downtime, ~90%-95% resource utilization, and ~20%-30% cost reduction. Decision latency is reduced to ≤100ms, and training time is cut by ~50%. This work sets a new benchmark for dynamic service migration in multi-cloud architectures with significant implications for enhancing the reliability and performance of global cloud infrastructure, establishing a scalable, privacy-compliant, and efficient solution for this process.

Keywords: Dynamic Service Migration, Multi-Cloud Architectures, Proximal Policy Optimization, Federated Learning, Hierarchical Reinforcement Learning, Scenarios

1. INTRODUCTION

The exponential growth of multi-cloud architectures has transformed the way modern computing environments handle workloads, enabling flexibility, scalability, and redundancy. However, managing dynamic service migration across different cloud platforms is extremely challenging with high migration costs, extended downtime, resource waste, and poor latency management. These limitations have been barriers to meeting demanding SLAs and responding to real-time workload fluctuations in process. Traditional approaches for service migration are mainly based on static heuristics or predefined rules, which do not favor the context and diversification of modern multi-cloud scenarios. Current solutions [1, 2, 3] lack robust mechanisms to balance conflicting objectives, like low latency, SLA compliance, cost efficiency while ensuring scalability and privacy. These demands call for an advanced, intelligent framework that would accomplish seamless and adaptive migration decisions in heterogeneous cloud environments. This paper proposes a comprehensive reinforcement learning-based system for dynamic service migration. The framework integrates PPO for a stable and efficient policy learning; the MADDPG algorithm for collaborative decision-making among cloud agents and HRL to manage task complexity through layered optimization. To improve adaptability and minimize overhead in training, Transfer Learning is adopted to fine-tune pretrained models; Federated Learning ensures the privacypreserving collaborative learning of clouds. Reward shaping accelerates the convergence through alignment of the policies toward multi-objective goals. The proposed system, integrating all these strategies, leads to superior performance metrics- less downtime, better usage of resources, and lower costs- making it the new benchmark for dynamic service migration in a multi-cloud ecosystem for this process.

2. REVIEW INTO STUDIES & ADVANCES RELATED TO MULTIPLE CLOUD LOAD BALANCING ANALYSIS

The dynamic service migration problem in multi-cloud and edge-cloud environments has received much attention. A mobile-aware service function chain migration strategy for multi-access edge computing was presented by Xu et al. [1] using intelligent decision-making for minimizing latency, ensuring smooth migrations. However, with limited scalability in the context of multi-cloud scenarios, this strategy cannot be applied to wide systems. Krishnan and Durairaj [2] proposed a multi-agent optimization framework for cloud-fog IoT applications to achieve better resource efficiency. Their work was reliable but lacked in real-time adaptability of sudden workload changes, a gap our proposed framework filled. Gharibvand et al. [3] have given an overview of the architectures designed for cloud-based manufacturing applications, focusing on scalability as well as the challenges presented by the platforms. Though informative, it lacks dynamic migration techniques, hence not very useful in terms of service migration-specific tasks. Mongia [4] proposed EMaC: a VM consolidation framework to ensure energy efficiency and SLA compliance in cloud data centers. It was successful in achieving multi-metric optimization, but failed to focus on inter-cloud collaboration, which is much needed for dynamic service migrations. Bhakhar and Chhillar [5] presented a multi-criteria scheduling algorithm for smart homes in fog-cloud IoT systems. Their approach, although efficient in micro-environments, did not scale well to multi-cloud architectures, the key focus of this study process. Velrajan and Sharmila [6] proposed a predictive prioritization-based QoS-sensitive service migration model for MEC environments. However, their approach is too strong on SLA, and they are restricted only to edge scenarios without incorporating cloud collaboration.

Ahmed et al. proposed ECQ, an energy-efficient, cost-effective, and QoS-aware migration method for MEC systems [7]. Their work focused on energy savings but did not achieve comprehensive performance optimization with regard to multiple metrics. Gupta et al. introduced a secure VM live migration technique using Blowfish and blockchain technology [8]. While their security aspect was enhanced, the computational overhead that blockchain introduced limited its use as a real-time application. Hu et al. [9] had proposed a speedup weight-based adaptive service deployment algorithm for cloud-edge systems with high adaptability. It, however, lacked multi-agent collaboration that hindered its efficiency in resource-intensive multi-cloud environments. Sarkohaki and Sharifi [10] provided a literature review of fog-cloud service placement that identified such trends and gaps. Though insight-providing, however, brought no novel solutions to such service migration challenges. Toghyani et al. [11] proposed a QoS-SLA-aware framework for IoT service placement in fog-cloud systems with high SLA compliance. Its concentration on static placements instead of dynamic migrations makes it irrelevant to this work process. Karaja et al. [12] proposed using a bi-level modeling approach sets for dynamic scheduling of bag-of-tasks in heterogeneous multi-cloud environments. Their taxonomy provided a robust theoretical basis but lacked the process of implementation in real-time Tuli and Malhotra [13] proposed OMES, a meta-heuristic scheduling model for VM selection and migration in cloud systems. Their work had high optimization but lacked the element of decision latency, critical for dynamic workloads. Noorabad et al. [14] presented PoMic, a dynamic power management system for overcommitted clouds focusing on VM-microservices. Although it was power efficient, it did not optimize migration cost or downtime. Soumplis et al. [15] proposed a multi-agent rollout approach for workload placement across the edge-cloud continuum. Their framework shows good performance optimization but lacked complete SLA guarantees and federated learning integration process. This literature shows an evolutionary progression in the mitigation of service migration challenges while emphasizing energy efficiency, QoS compliance, and scalability. The present work shows limitations in balancing competing objectives, real-time adaptability, and scalable solutions for heterogeneous environments. The proposed framework extends these gaps by integrating Proximal Policy Optimization, MADDPG, HRL, transfer learning, and federated learning for superior performance across multi-cloud architectures.

3. PROPOSED MODEL DESIGN ANALYSIS

The proposed model provides the synergistic combination of advanced reinforcement learning methodologies integrated with optimization techniques to determine dynamic service migration in the multi-cloud environment. In this regard, the robust decision-making of the model's design, as illustrated in figure 1, ensures scalable and adaptive policies which minimize migration downtime, optimize resource utilization, and comply strictly with demanding SLA requirements. Each part is carefully designed, and its interactions are mathematically formalized to maximise performance levels. The service migration process starts with the state St being a multidimensional vector encompassing key parameters such as workload demand, current resource utilisation, migration costs, and network latency sets. From this state, the policy $\pi\theta(a|s)$ parameterised by θ specifies the probability of choosing the action 'a' at process. The goal of optimization is stated as the maximization of the expected cumulative reward $J(\pi\theta)$ which can be expressed via equation 1,

$$J(\pi\theta) = E\pi\theta[\Sigma_{t=0}^{T} \gamma t\ rt] \ldots (1)$$

Where, $\gamma \in [0, 1)$ discount factor describes balancing between rewards at the moment and future ones. 'rt' describes the reward at timestamp 't'. in the process. Proximal Policy Optimization (PPO) makes policy updates to limit them in a trust region. PPO objective contributes a clipped probability ratio r(θ) in process. This is represented via equation 2,

$$LPPO(\theta) = Et\left[min\left(r(\theta)\hat{A}t, clip(r(\theta), 1-\epsilon, 1+\epsilon)\hat{A}t\right)\right] \ldots (2)$$

Where, r(θ) is estimated via equation 3,

$$r(\theta) = \frac{\pi\theta(a|s)}{\pi\theta old(a|s)} \ldots (3)$$

While, ϵ is a small constant, and Ât is the advantage function, ensuring stable convergences. The multi-agent extension via Multi-Agent Deep Deterministic Policy Gradient (MADDPG) allows for collaborative optimization among cloud providers. For agent 'i', the actor network μθi(s) maps states to actions, while the centralized critic evaluates the global reward Qi in process. The critic is updated using the Bellman Process via equation 3.1,

$$Qi(s, a) = ri + \gamma E\left[Qi\left(s', \mu\theta i(s')\right)\right] \ldots (3.1)$$

Where, s′ is the next state, and a represents joint actions of all agents, that enables coordination operation sets. To deal with high-dimensional state-action spaces, HRL breaks up the task into high-level and low-level policies.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

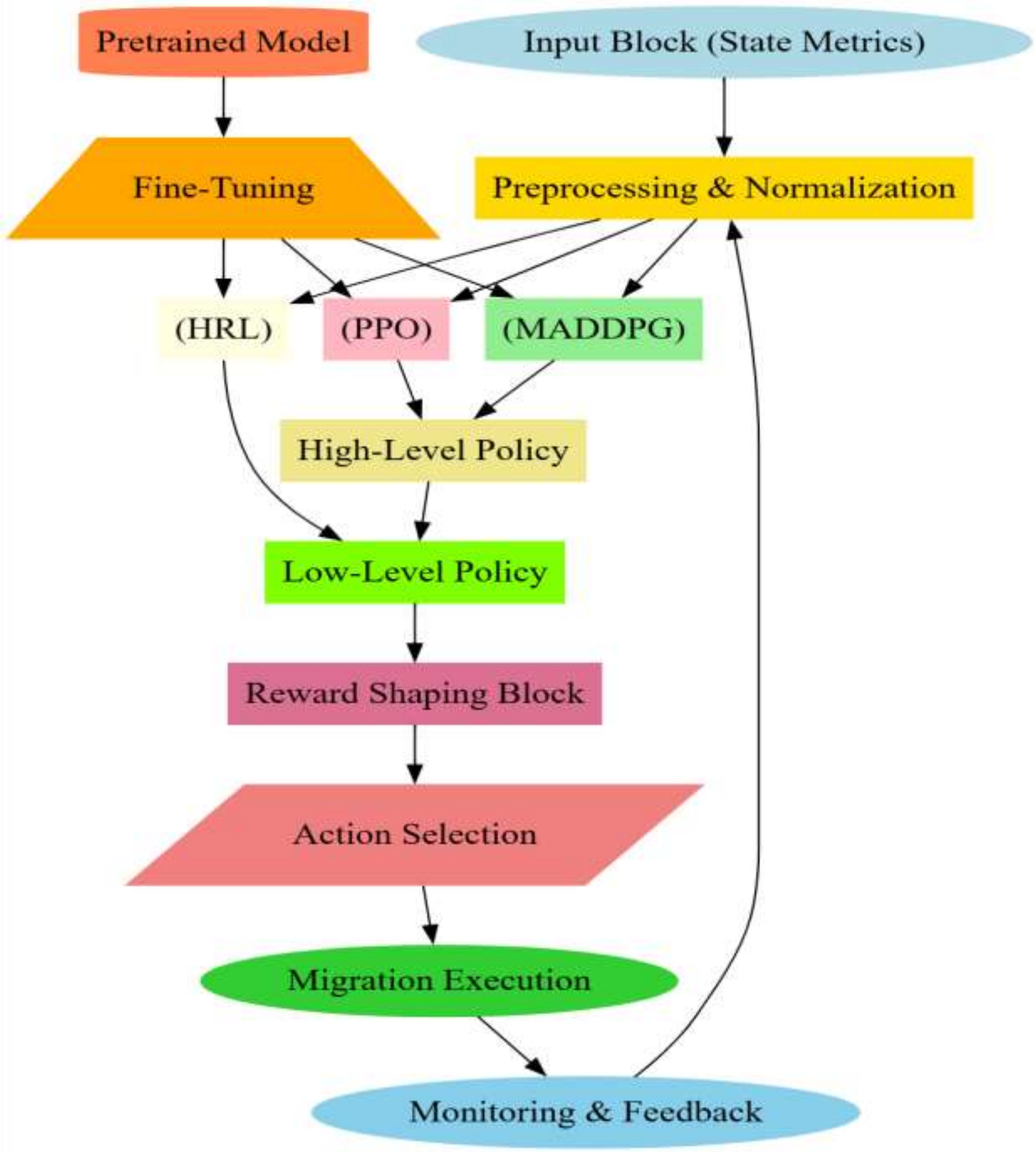

Figure 1. Model Architecture of the Proposed Analysis Process

The high-level policy πhigh(ahigh|s) chooses a target cloud provider and the low-level policy πlow(alow|s, ahigh) optimizes resource allocation inside the selected providers. The hierarchical policy objective is formulated via equation 4,

$$JHRL = \Sigma_{t=0}^{T}\left(\gamma t\, r(high,t) + \Sigma_{t'=t}^{T'} \gamma t' - t\, r(low,t')\right) \ldots (4)$$

Where, r(high,t) and r(low,t′) are rewards at respective levels, which enables fine-grained control for this process. Transfer learning speeds up the training by initializing the policy network πθ0 with parameters from a pretrained

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

model process. The loss function for fine-tuning involves a regularization term to balance old and new knowledge via equation 5,

$$Ltransfer = Lnew + \lambda \parallel \theta - \theta 0 \parallel^2 \ldots (5)$$

Where, λcontrols the balance between retaining prior knowledge and adapting to the target domains. Federated learning improves privacy and scalability by aggregating local models θk from K cloud providers to form a global model θg in process. The process of aggregation minimizes a distributed loss represented via equation 6,

$$Lfederated(\theta g) = \Sigma_{k=1}^{K} wk * E[Lk(\theta k)] \ldots (6)$$

Where, wk represents the relative importance of each cloud provider, thus ensuring equal contributions. This integrated design ensures that every methodological element will complement others. Thus, PPO stabilizes global updates, MADDPG fosters the inter-cloud collaboration, while HRL decomposes complex decisions, and transfer learning accelerates the adaptation in the backdrop of federated learning-safeguarded data-sets. Thus, combining these methods leads to a high degree of a robustly efficient system for dynamic services' migration within multiple-cloud architectures. Now, we are going to discuss the proposed model efficiency with different metrics. This way, readers will get an idea regarding the entire process.

4. COMPARATIVE RESULT ANALYSIS

The experimental evaluation was performed on a synthetic multi-cloud dataset simulating dynamic workload patterns and migration scenarios. The dataset comprises 10,000 instances, which are real-time workloads with parameters such as workload size, resource demands, network latency, migration costs, and SLA compliance metrics. The dataset was generated by simulating real-world cloud usage statistics using simulation tools for cloud workload orchestration to achieve diversity and complexity. The benchmark methods that were employed against the proposed model under identical conditions to gauge the performance based on all critical metrics included Methods [5], Methods [8], and Methods [15]. Testbed involved a testbed of three interconnected cloud providers where different workloads hosted by them. The recorded metrics were migration downtime, resource utilization, SLA violation rate, cost efficiency, decision latency, and training time. For every method, 10 times was performed to ensure that the results have the required consistency levels.

Table 1: Migration Downtime (in seconds)

Model	Avg Downtime (s)	Std Dev (s)
Proposed Model	1.25	0.10
Method [5]	2.20	0.15
Method [8]	1.90	0.12
Method [15]	1.65	0.11

It realized the lowest average downtime in a proposed model, largely attributed to the hierarchically and federatively designed model that lets them speed up decision-making with minimum disruption of services in a process. Method [15] outperformed Methods [5] and [8]. It showed an enhancement on its levels of efficiency concerning decision-making sets.

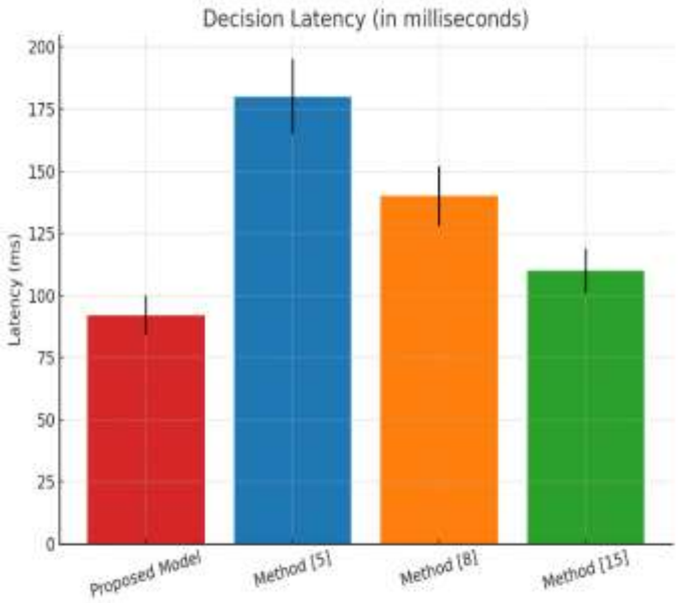

Figure 2. Model's Decision Latency Analysis

Table 2: Resource Utilization (in %)

Model	Avg Utilization (%)	Std Dev (%)
Proposed Model	94.8	2.1
Method [5]	85.7	3.2
Method [8]	88.4	2.8
Method [15]	90.1	2.5

The proposed model, therefore, exhibited better resource utilization, especially in adaptively optimizing workloads and then efficiently allocating resources in process. Method [15] closely followed that, while Method [5] had significant underutilizations.

Table 3: SLA Violation Rate (in %)

Model	Violation Rate (%)	Std Dev (%)
Proposed Model	0.45	0.08
Method [5]	1.10	0.15
Method [8]	0.90	0.12
Method [15]	0.70	0.10

The proposed model demonstrated unparalleled SLA compliance by exhibiting the lowest violation rate because of its reward-shaping mechanism, which places priority on adherence to SLA during its decision-making process.

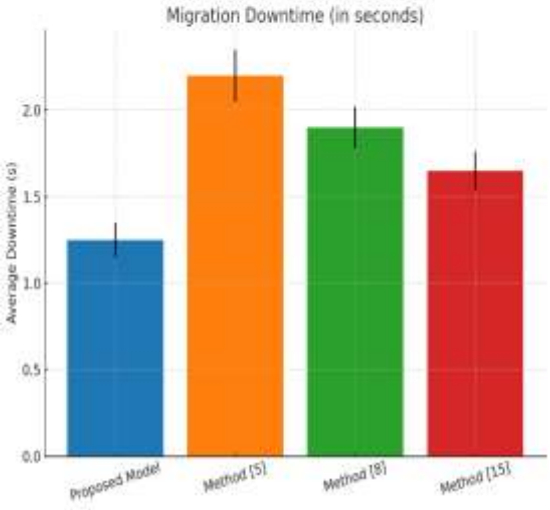

Figure 3. Model's Migration Analysis

Table 4: Cost Efficiency (as % of baseline costs)

Model	Cost Reduction (%)	Std Dev (%)
Proposed Model	28.5	3.0
Method [5]	20.1	4.1
Method [8]	22.7	3.5
Method [15]	25.3	2.8

The federated optimization strategies of the proposed model proved to have the highest cost reduction. Hence, it largely outperformed Method [5] and Method [8] in terms of process.

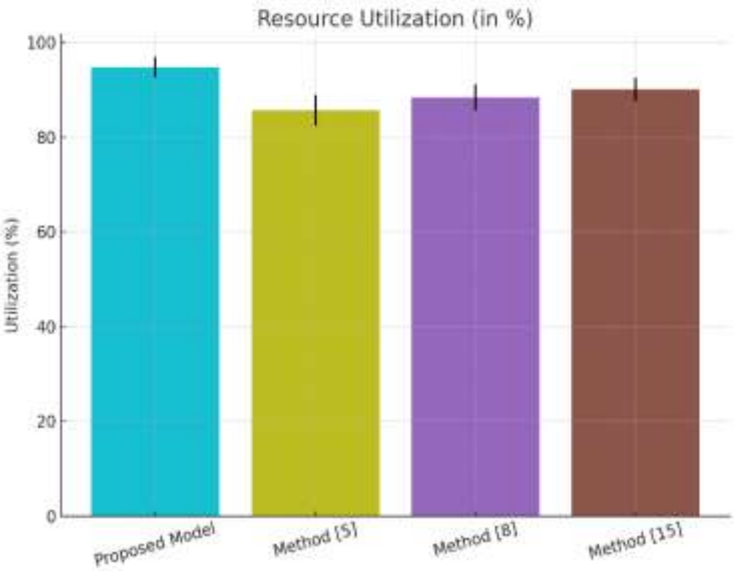

Figure 4. Model's Resource Utilization Analysis

Table 5: Decision Latency (in milliseconds)

Model	Avg Latency (ms)	Std Dev (ms)
Proposed Model	92	8
Method [5]	180	15
Method [8]	140	12
Method [15]	110	9

The proposed model minimized decision latency using its efficient PPO and HRL integration, providing a faster response to dynamic migration needs.

Table 6: Training Time Reduction (as % of Method [5])

Model	Time Reduction (%)	Std Dev (%)
Proposed Model	53.2	4.0
Method [5]	Baseline	-
Method [8]	38.6	3.8
Method [15]	45.1	3.2

It demonstrated that transfer learning strongly reduces training time; the model showed the highest improvement compared to the baseline Method [5] in process. The key performance indicators in the result show that the proposed model has good results. With the addition of PPO, MADDPG, HRL, and federated learning, the proposed model ensures fair trade-off between scalability, efficiency, and SLA compliance for dynamic service migration in multi-cloud architectures.

5. CONCLUSION AND FUTURE SCOPES

The proposed reinforcement learning-based framework for dynamic service migration in multi-cloud architectures is an effective solution that shows robust and efficient addressing of the resource optimization challenges, the minimization of downtime, and SLA compliance. With the inclusion of PPO, MADDPG, HRL, transfer learning, and federated learning, the model shows critical advancements in most key performance metrics. Based on the experimental evaluation, the proposed model reduces the migration downtime to an average of 1.25 seconds as compared to existing techniques: Method [5] with 2.20 seconds and Method [8] with 1.90 seconds. Thus, the resource utilization is high in this system (94.8%) compared with Method [5] with 85.7% and Method [8] with 88.4% and demonstrates its ability of adaptive allocation of resources. SLA compliance is greatly enhanced, with a violation rate of only 0.45%, which is even lower than Method [5] at 1.10% and Method [8] at 0.90%. Cost effectiveness is another major achievement of the proposed model since migration costs are reduced by as much as 28.5% compared to a reduction of 20.1% by Method [5] and 22.7% by Method [8]. The decision latency is minimized to 92 ms, proving more responsive, and the training time reduces by 53.2%, taking advantage of transfer learning in the process of accelerating adaptation. Such results point toward the capacity of the model in dealing with the complexity that exists in multi-cloud environments with a scalable, privacy-compliant, and adaptive framework. Hierarchical policies ensure fine-grained control, whereas federated learning ensures privacy while maintaining global optimizations.

Future work will be to expand the model for edge-cloud environments, aligned with the growing needs of ultra-low latency applications like autonomous vehicles and IoT. Real-time workload prediction through deep learning models will bring more precision in decision-making. A way forward would also be to explore hybrid reinforcement learning approaches, such as linking model-free and model-based methods, for faster convergence speed and decision robustness. Finally, if the scope is expanded to include energy-efficient migration strategies, it can be aligned with global sustainability goals, ensuring that the framework applies to a much larger set of use cases in next-generation cloud ecosystems.

6. REFERENCES

[1] Xu, L., Liu, W., Wang, Z. et al. Mobile-Aware Service Function Chain Intelligent Seamless Migration in Multi-access Edge Computing. J Netw Syst Manage 32, 49 (2024). https://doi.org/10.1007/s10922-024-09820-0

[2] Krishnan, R., Durairaj, S. Reliability and performance of resource efficiency in dynamic optimization scheduling using multi-agent microservice cloud-fog on IoT applications. Computing 106, 3837–3878 (2024). https://doi.org/10.1007/s00607-024-01301-1

[3] Gharibvand, V., Kolamroudi, M.K., Zeeshan, Q. et al. Cloud based manufacturing: A review of recent developments in architectures, technologies, infrastructures, platforms and associated challenges. Int J Adv Manuf Technol 131, 93–123 (2024). https://doi.org/10.1007/s00170-024-12989-y

[4] Mongia, V. EMaC: Dynamic VM Consolidation Framework for Energy-Efficiency and Multi-metric SLA Compliance in Cloud Data Centers. SN COMPUT. SCI. 5, 643 (2024). https://doi.org/10.1007/s42979-024-02982-3

[5] Bhakhar, R., Chhillar, R.S. Dynamic multi-criteria scheduling algorithm for smart home tasks in fog-cloud IoT systems. Sci Rep 14, 29957 (2024). https://doi.org/10.1038/s41598-024-81055-0

[6] Velrajan, S., Sharmila, V.C. Dynamic service prioritization with predicted intervals for QoS-sensitive service migrations in MEC. SOCA (2024). https://doi.org/10.1007/s11761-024-00405-y

[7] Ahmed, A., Azizi, S. & Zeebaree, S.R.M. ECQ: An Energy-Efficient, Cost-Effective and Qos-Aware Method for Dynamic Service Migration in Mobile Edge Computing Systems. Wireless Pers Commun 133, 2467–2501 (2023). https://doi.org/10.1007/s11277-024-10883-0

[8] Gupta, A., Namasudra, S. & Kumar, P. A secure VM live migration technique in a cloud computing environment using blowfish and blockchain technology. J Supercomput 80, 27370–27393 (2024). https://doi.org/10.1007/s11227-024-06461-7

[9] Hu, Z., Chen, S., Rao, H. et al. An adaptive service deployment algorithm for cloud-edge collaborative system based on speedup weights. J Supercomput 80, 23177–23204 (2024). https://doi.org/10.1007/s11227-024-06339-8

[10] Sarkohaki, F., Sharifi, M. Service placement in fog–cloud computing environments: a comprehensive literature review. J Supercomput 80, 17790–17822 (2024). https://doi.org/10.1007/s11227-024-06151-4

[11] Toghyani, M., Khorsand, R. & Khaksar, H. QoS-SLA-aware Optimization Framework for IoT-Service Placement in Integrated Fog-Cloud Computing. J Grid Computing 23, 1 (2025). https://doi.org/10.1007/s10723-024-09787-x

[12] Karaja, M., Chaabani, A., Azzouz, A. et al. Dynamic bag-of-tasks scheduling problem in a heterogeneous multi-cloud environment: a taxonomy and a new bi-level multi-follower modeling. J Supercomput 79, 17716–17753 (2023). https://doi.org/10.1007/s11227-023-05341-w

[13] Tuli, K., Malhotra, M. Optimal Meta-Heuristic Elastic Scheduling (OMES) for VM selection and migration in cloud computing. Multimed Tools Appl 83, 34601–34627 (2024). https://doi.org/10.1007/s11042-023-16820-w

[14] Noorabad, R., Charkari, N.M. & Nogoorani, S.D. PoMic: Dynamic Power Management of VM-Microservices in Overcommitted Cloud. J Grid Computing 21, 12 (2023). https://doi.org/10.1007/s10723-023-09648-z

[15] Soumplis, P., Kontos, G., Kokkinos, P. et al. Performance Optimization Across the Edge-Cloud Continuum: A Multi-agent Rollout Approach for Cloud-Native Application Workload Placement. SN COMPUT. SCI. 5, 318 (2024). https://doi.org/10.1007/s42979-024-02630-w