Energy-Efficient Cloud Computing Through Reinforcement Learning-Based Workload Scheduling

Ashwini R Malipatil[1], Dr M E Paramasivam[2], Dilfuza Gulyamova[3], Dr.Aanandha Saravanan[4],
Janjhyam Venkata Naga Ramesh[5], Elangovan Muniyandy[6], Refka Ghodhbani[7]*
Assistant Professor, Department of Computer Science and Engineering, BNM Institute of Technology, Bangalore, India[1]
Associate Professor, Department of ECE, Sona College of Technology, Salem, India[2]
Computer Engineering Department, University of Information Technologies in Tashkent, Tashkent, Uzbekistan[3]
Professor, Department of ECE, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, India[4]
Adjunct Professor, Department of CSE, Graphic Era Hill University, Dehradun, 248002, India[5]
Adjunct Professor, Department of CSE, Graphic Era Deemed to be University, Dehradun, 248002, Uttarakhand, India[5]
Department of Biosciences-Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai, India[6]
Applied Science Research Center, Applied Science Private University, Amman, Jordan[6]
Center for Scientific Research and Entrepreneurship, Northern Border University, 73213, Arar, Saudi Arabia[7]

***Abstract*—The basis for current digital infrastructure is cloud computing, which allows for scalable, on-demand computational resource access. Data center power consumption, however, has skyrocketed because of demand increases, raising operating costs and their footprint. Traditional workload scheduling algorithms often assign performance and cost priority over energy efficiency. This paper proposes a workload scheduling method utilizing deep reinforcement learning (DRL) that adjusts dynamically according to present cloud situations to ensure optimal energy efficiency without compromising performance. The proposed method utilizes Deep Q-Networks (DQN) to perform feature engineering to identify key workload parameters such as execution time, CPU and memory consumption, and subsequently schedules tasks smartly based on these results. Based on evaluation output, the model brings down the latency to 15 ms and throughput up to 500 tasks/sec with 92% efficiency in load balancing, 95% resource usage, and 97% QoS. The proposed approach yields improved performance in terms of key parameters compared to conventional approaches such as Round Robin, FCFS, and heuristic methods. These findings show how reinforcement learning can significantly enhance the scalability, reliability, and sustainability of cloud environments. Future work will focus on enhancing fault tolerance, incorporating federated learning for decentralized optimization, and testing the model on real-world multi-cloud infrastructures.**

Keywords*—Cloud computing; energy efficiency; reinforcement learning; virtual machine; workload scheduling

I. Introduction

Cloud computing technology has emerged as essential technology because it provides adaptable and effective computer resources to all individuals together with enterprises worldwide. Cloud services provide instant access to processing power and storage and networking capabilities which has led to a total transformation of various business operations. Since this transition occurred companies can carry out innovation and growth at rapid speeds [1]. The expansion of cloud service use raises the electricity consumption in major data center facilities [2]. This challenge has become somewhat important as meeting the demand for energy, without losing performance in cloud-based services is a problem. Data centers form the foundation of cloud computing today. Computing, Storage, and Networking are the functions in data centers. They consume much electricity [3]. This increase in workload in the data centers due to the ever-increasing demand for cloud-based applications has seen energy consumption significantly shoot up. Most of the power consumed by the data centers can be traced to the need to process complicated workloads and maintain cooling and networking operations, thus ensuring cloud services are always functioning [4]. Therefore, the reduction of the energy footprint of cloud computing has become both a technical and environmental imperative. Optimization of energy usage in cloud data centers is a concern not only to reduce the cost of operation but also as an urgent environmental need. The achievement of maximum efficiency requires workloads to have effective scheduling algorithms. This system will enable resource management and minimal energy usage to achieve efficient operation and cost-effectiveness for cloud data centers [5].

The actual practice of workload scheduling requires dispersing computational workloads onto virtual machines to achieve minimum usage of power and resources. Cloud infrastructure performance and operational expenses improve through scheduling optimization leading to better operation. The majority of traditional scheduling approaches present limited interest in energy conservation because they focus on two separate objectives: peak performance and cost reduction. Cloud computing presents great challenges in workload management since user needs vary frequently and deployment options differ widely and service-level agreements are strictly enforced in this environment where workloads exhibit unexpected dynamism and extreme resource variability [6]. Fixed scheduling techniques and those following rules fail to manage operational modifications in real-time which leads to poor resource distribution together with increased energy consumption. The research recommends using reinforcement-learning algorithms for scheduling problems in cloud environments.[7]. The demand for adaptive scheduling approaches which minimize power usage without affecting system performance remains immediate [8], Systems that run through the cloud have the ability to improve their energy efficiency without jeopardizing either their reliable service level or performance standards through this approach [9].

The framework enables system learning through environment interactions because of its reinforcement learning capabilities. The optimal scheduling policy through interaction with the technique has potential for workload scheduling by leveraging the environment to teach optimal scheduling policies. In reinforcement learning, an agent acts in response to its perception of the environment and is rewarded or punished appropriately. It lets the agent learn to improve with time in such a manner that it learns from the outcome of its actions [10]. Using RL for scheduling workload in the cloud, an indirect resource allocation can be made on the fly according to the real-time demands and thus energy consumption optimizes continuously. Unlike previous heuristic-based methods, these methods will not depend on predefined rules but learn from previous experiences and hence, can handle very dynamic and complex cloud environments [11]. This adaptability pays off well in cloud computing, where changes in workload characteristics can be rapid, requiring real-time resource allocation on the fly. In this study, the focus is primarily on using algorithms based on RL to schedule cloud workloads, with the purpose of minimizing energy consumption while respecting performance standards. Through reinforcement learning, this study will describe and apply the strategy that optimizes the pursuit of energy efficiency within cloud infrastructures. For reducing energy consumption without compromising its overall performance within cloud services, we suggest an RL-based task scheduling algorithm that adjusts to changing system conditions and workload demands [12]. The proposed method integrates reinforcement learning with energy-aware scheduling techniques, which enables dynamic adjustments to workload distribution based on real-time energy consumption data. This will help cloud computing infrastructures become more sustainable by incorporating feedback loops and making adjustments to scheduling policies based on past performance. The rest of the paper focuses on the design, implementation, and evaluation of the effectiveness in bringing down the energy consumption while ensuring high service reliability for the RL-based scheduling algorithm [13]. This research contributes to the emerging area of energy-efficient cloud computing by introducing a new approach for optimizing energy usage in cloud data centers using reinforcement learning techniques.

The major key contributions are as follows:

1) It introduces a reinforcement learning-based algorithm for optimal energy-efficient scheduling of workloads in cloud environments.

2) The scheduler adapts strategies in real-time to manage unanticipated cloud workloads while optimizing energy consumption.

3) It balances optimization of energy consumption with QoS constraints to satisfy SLA requirements and system dependability.

4) To validate the proposal, the presented method is tested against traditional algorithms such as FCFS, RR, and heuristic-based algorithms with better energy efficiency and performance.

5) It ensures scalability across different cloud infrastructures and hence is applicable to various real-world cloud service providers.

The following is the remaining part of the section is structured: Section II as Related works on previous papers, Section III as problem statement, Section IV as proposed methodology, Section V as result and discussion and Section VI as conclusion and future work is provided.

II. Related Works

Mobula et al. [14] proposed a new approach to address the challenges of workflow scheduling in the cloud environment by focusing on the optimization of energy efficiency while satisfying user-defined constraints such as deadlines and budget. Acknowledging that workflow scheduling is an NP-complete, they proposed two algorithms called Structure-based Multi-objective Workflow Scheduling with an Optimal instance type and Structure-based Multi-objective Workflow Scheduling with Heterogeneous instance types. The SMWSO algorithm computes the optimal instance type and the number of virtual machines that should be required to improve the scheduling efficiency. In the meanwhile, SMWSH extends this concept by adding heterogeneous VMs that allow greater flexibility in a diverse cloud environment. Their research work emphasizes the critical role workflow structures play in making scheduling decisions and proves that optimized instance types and VM allocations can significantly decrease energy consumption. Based on simulations, their methods obtained superior heir result manifests the relevance of using workflow-aware scheduling strategies in cloud environments, especially in cost reduction and sustainability. Their work serves as a basis for further research into intelligent workload scheduling strategies that integrate optimisation techniques to improve cloud computing infrastructures.

Murad et al. [15] proposes an Optimized Min-Min (OMin-Min) task scheduling algorithm for enhancing cloudlet scheduling and resource allocation. This study is hoped to increase the performance of a system by increasing resource utilization and decreasing task execution time. The OMin-Min, which is the enhanced version of the traditional Min-Min, is constructed by applying these approaches, and the performance of OMin-Min is compared to that of the Min-Min, Round Robin, Max-Min, and Modified Max-Min algorithms. The experiments include different sizes of cloudlets (small, medium, large, and heavy) on three scientific workflow datasets: Montage, Epigenomics, and SIPHT.The evaluation and implementation are performed using the CloudSim simulator within a Java environment. The merits of the new algorithm are in its capacity to generate optimal scheduling outcomes, provide lower completion times, and ensure improved resource utilization, ultimately contributing to better throughput. However, the limitation might be based on the computational difficulty of optimal scheduling decisions for large-scale or extremely dynamic systems. The performance indicators indicate that OMin-Min performs better than all other algorithms in all test cases with the most efficient scientific workflow task scheduling solution in the cloud.

Panda et al. [16] a new approach to the task scheduling problem in cloud computing termed NP-Complete since it tries to optimize the overall execution time is proposed. In this work, we introduce a pair-based task scheduling algorithm that aims to enhance scheduling performance by reducing overall layover time, which is the total of timing gaps between paired jobs. Through forming task pairs to guide scheduling decisions, the technique, founded upon the Hungarian algorithm of optimization, applies it innovatively to situations where there are uneven numbers of tasks and clouds. On twenty-two different data sets, performance of the proposed algorithm is checked and compared to three existing algorithms: First-Come-First-Served (FCFS), the Hungarian algorithm with lease time, and the Hungarian algorithm with converse lease period. The results indicate that the proposed strategy consistently performs better than the comparison methods with regard to layover time. The processing cost involved in integrating logic and multiple iterations may, however, be a drawback in real-time or highly large-scale systems. Overall, the study provides a systematic and effective technique to cloud-based work scheduling that promises to perform better than traditional methods.

Shaw et al. [17] explores the critical issue of energy consciousness in cloud data centers through automated energy-saving Virtual Machine (VM) consolidation using Reinforcement Learning (RL) methods. Virtual machine consolidation is an important strategy to save energy consumption and enhance data centers' greenness. For the sake of enhancing resource utilization and minimizing energy-related costs, this work will explore applying RL algorithms for dynamic VM allocation optimization. The methods comprise popular RL algorithms such as SARSA and Q-learning, which are evaluated for their ability to reason under uncertainty and therefore learn proper consolidation procedures without knowing the environment beforehand. The primary contribution of this work is its demonstration that RL-based VM consolidation can lead to a 63% reduction in service violations and a 25% improvement in energy efficiency, which shows a significant performance improvement over traditional heuristics. The computational cost and training time typically associated with reinforcement learning models, however, might be a drawback as it may affect the scalability or real-time adaptability of the models within bigger-scale cloud systems. Ultimately, the article illustrates that RL offers a robust, versatile solution to dynamic virtual machine consolidation and significantly contributes to the construction of next-generation, energy-efficient cloud infrastructures.

Malik et al. [18] The authors created a job scheduling method with energy consciousness to optimize cloud data centers' virtualized resource benefit while lowering their energy requirements. This approach implements three key elements that first segregate jobs and next schedules them according to set thresholds while preventing system slowdowns. During pre-processing the first phase creates distinct queues for tasks that demonstrate high dependability standards and have long execution durations. Task organization relies on resource intensity levels to achieve proper distribution among resources. Through their scheduling method based on PSO algorithm the authors achieve dynamic selection of optimal schedules that consider workload distribution together with energy efficiency goals. Experimental benchmarking of conventional scheduling methods confirms that the proposed algorithm demonstrates superior performance according to results obtained from test datasets.

Panwar et al. [19] The research delivered an extensive examination of methods to decrease energy usage in cloud data center operations because of the relationship between fast cloud growth and increased power consumption. Through the work the researchers study various optimization approaches that enhance cloud data center energy efficiency because they recognize excessive energy usage leads to environmental deterioration. The study examines CPU utilization forecasting alongside detection methods for underload and overload situations and procedures for selecting and moving virtual machines and picking their deployment locations. The authors compare energy savings of various methods and demonstrate the effectiveness of heuristic approaches, achieving energy reductions of 5.4 percent to 90 percent over the current methods. The highest energy saving potential of 7.68 percent to 97 percent was realized through the use of metaheuristic methods, machine learning techniques at 1.6 percent to 88.5 percent, and finally through the application of statistical techniques to save 5.4 percent to 84 percent. This review highlighted the effects that these techniques can have: not only decreasing the consumption of energy but also reducing related greenhouse gas emissions and water usage for electricity generation. This paper combines the various findings from various works to provide an understanding of the various means through which energy efficiency and sustainability in cloud data centers can be improved.

Yadhav and chawla [20] discusses different task scheduling algorithms in the cloud environment to consume less energy. Cloud computing is among the fastest-growing technologies in the computer world; thus, in modern cloud data centers, managing energy efficiency has become crucial. This paper presents an overview of different kinds of heuristic and machine learning-based algorithms for optimizing task scheduling. These are Genetic Algorithm and Particle Swarm Optimization, highlighted for their efficiency in finding nearly optimal solutions over large search spaces, thus applicable to energy minimization. There is also discussion on Reinforcement Learning, which, through dynamic adaptation to workload variations, has shown its potential in optimizing energy efficiency via continuous learning and adaptation. The paper considers other related techniques, such as Ant Colony Optimization and Dynamic Voltage and Frequency Scaling, which provide mechanisms for trading off the metrics performance and energy usage. The considered algorithms are evaluated in detail, emphasizing their performance in cloud-like environments. The results clearly show that, although no individual algorithm appears to be ideally optimized in general, a tailored blending of the techniques offers a significant energy saving. This paper focuses on the selection of an appropriate algorithm or set of algorithms that should optimize the energy consumption in cloud data centers, thus adding to the contribution of sustainability and cost savings.

Liu et al. [21] presents a greedy scheduling approach to improve energy consumption and resource utilization for cloud data centers. Cloud computing systems are plagued by issues of excessive energy consumption and poor resource utilization, particularly with heterogeneous resources. In addressing such issues, this paper introduces the granular computing theory into cloud task scheduling, where tasks are categorized into three categories: CPU, memory, and hybrid types. This categorization enables the use of particular scheduling methods based on the nature of the various task types. The article identifies that the cloud resource is heterogeneous in nature and that distinct scheduling methods must be employed for distinct types of tasks to ensure maximum energy saving. The efficiency of the proposed approach is established by numerical experiments on the CloudSim platform, and the results indicate significant improvement in terms of energy efficiency. The results demonstrate that, for a specific task type, the greedy scheduling strategy is able to reduce energy usage while maximizing the utilization of resources. It is thus an efficient practical approach for energy optimization in cloud data centers, improving resource management methods in cloud computing.

Pandey et al. [22] to enhance resource utilization and energy efficiency in cloud computing environments that enable large-scale data processing. Cloud computing is a vital alternative as conventional computing infrastructure fails to meet the growing demand for real-time data processing, high-performance analytics, and massive storage. Yet, complex problems such as scheduling, load balancing, power management, and resource allocation have been created by this surge in cloud services. The research discusses state-of-the-art strategies such as swarm intelligence-based meta-heuristics to address problems, with a particular focus on Discrete Particle Swarm Optimization (DPSO) for workflow scheduling and resource allocation. In cloud resource management, the DPSO approach maximizes particle positions and velocities in a series of fitness evaluations and iterative updates. Even though the paper emphasizes the unification of numerous PSO variants and presents a detailed algorithmic structure, it has no reference to some specific dataset, which means that the research is conceptual or algorithmic in scope. The most important strength of this study lies in its exhaustive application of clever learning models made for cloud dynamics and hybrid optimization techniques. Its lack of empirical evaluation, applicability, or performance comparison to existing standards is a major drawback, however. To offer a plausible route for cloud computing in big data systems on a sustainable path, the study concludes by proving how combining LSTM with DPSO can significantly advance energy-efficient resource allocation and scheduling in dynamic cloud systems.

Katal et al. [23] explore a range of methods of reducing data center power consumption in an effort to promote the concept of green cloud computing. The ongoing impact of the internet on nearly every aspect of the modern economy has driven energy and processing power demand higher, particularly in data centers that support cloud services. A range of methods for saving energy are discussed in the paper, including hardware-level optimization techniques, firmware and hardware-level dynamic power management (DPM), and power-saving methods employed at the network and server cluster levels. By regulating e-waste, reducing unnecessary energy consumption, and reducing carbon footprints, these systems intend to encourage sustainable computing practices. The research does not utilize any specific dataset or present empirical findings, although it offers a comprehensive review of existing practices and highlights the necessity of energy-efficient processes. The research is conceptual and integrates existing approaches in the discipline instead of proposing new methodologies. The primary advantage of this work is its thorough examination of energy-saving measures at different system levels, which provides valuable information regarding the development of green data centers. The lack of quantitative analysis or experimental validation, however, is a major drawback. To develop more sustainable and energy-efficient data center infrastructures, the conclusion of the paper emphasizes the necessity of ongoing innovation and points out research challenges.

Medara and singh [24] emphasizes the increasing importance of cloud computing, which has been used as the major structure for all enterprises. All types of enterprises were allowed to use cloud for business development. The paper discusses an issue of energy consumption for scientific workflow applications in cloud data centers. Since cloud services are increasingly being deployed, a lot of consumers have started seeing the massive power utilization that comes as a result. This paper reviews existing energy-efficient scheduling techniques specifically designed for workflow applications in cloud environments. It focuses on approaches that attempt to minimize energy consumption while satisfying quality of service constraints. The review offers a comprehensive overview of the paradigms that have been introduced in the literature regarding energy-aware scheduling, discussing the advancements that have been achieved and their weaknesses. Through the examination of numerous approaches, this paper sheds light on the trajectory of energy-aware scheduling and their practical impacts in cloud settings. Additionally, it outlines possible areas of future research so that the work can contribute to ongoing discussion in energy efficiency. This is a very valuable piece of resource for researchers and practitioners aiming at building more efficient solutions for reducing.

The recent works section discusses some strategies for optimizing energy consumption in cloud computing environments, especially workflow and task scheduling. Researchers have proposed several algorithms to handle the intricacies of minimizing energy usage while maintaining quality of service and adhering to user-defined constraints like deadlines and budgets. There are several approaches such as multi-objective scheduling, task classification, and dynamic voltage scaling that can be used to reduce the energy footprint of cloud data centers without performance degradation. Intelligent scheduling strategies, such as Particle Swarm Optimization, Genetic Algorithms and Reinforcement Learning, help in adapting to variations in workload and improve resource allocation. Other researches further emphasize how advanced models, such as queuing systems, genetic algorithms, and greedy scheduling techniques, may be used to improve energy efficiency further. There are also energy-aware scheduling paradigms that integrate optimized resource usage and avoidance of unnecessary energy spends into a system.

III. Problem Statement

The rampant growth in cloud computing has led to data centers consuming much more energy, leading to increased operating expenses [25]. Traditional workload scheduling methods, such as heuristic and rule-based algorithms, often fail to effectively optimize resource usage, leading to performance bottlenecks and wastage of energy [26]. There are problems related to computational complexity, model convergence time and usability in practical large-scale cloud scenarios, which burden existing RL-based scheduling models. A much-improved dynamic scheduling mechanism is in urgent demand in order to lower energy consumption without compromising the system's performance and service reliability. The purpose of this work is to design a workload scheduling algorithm using reinforcement learning that makes cloud computing energy-efficient [27]. The proposed model with the help of deep reinforcement learning (DRL) techniques such as Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN) wants to optimize the allocation of the workload to the virtual machines (VMs) to achieve the maximum. Massive energy savings, reduction in delay in execution, and eco-friendly cloud computing processes are the ambitions. The creation of energy-efficient, smart cloud infrastructures that can dynamically adapt to variations in workload in real time while minimizing their adverse impacts on the environment will be facilitated by the resolution of these issues [28].

IV. Methodology

Cloud computing is becoming the basis of modern digital infrastructure, thus offering scalable, on-demand access to computing resources to various sectors. However, as the numbers of cloud services and applications rise, the amount of energy data centers consume also rises, contributing to increased operation costs and a larger carbon footprint. Traditional scheduling techniques for workload usually focus more on performance optimization and cost than energy efficiency. The research put forward a workload scheduling algorithm dynamically adjusts the assignment of tasks in order to optimize energy consumption with service quality not going below a threshold. Unlike some static or heuristic-based scheduling techniques, reinforcement learning adapts real-time workload variations better to the cloud environment and optimizes resource allocation within a system by continuously learning from past scheduling decisions to reduce power wastage. Applying feature engineering techniques enables extraction of the appropriate workload attributes like CPU usage, memory demand, and execution time so that the model is enabled to take the informed decision on scheduling. Also, the model's performance robustness has been improved using simulated workload scenarios in the process of training and testing. By combining reinforcement learning with intelligent workload scheduling, this method not only decreases energy consumption but also guarantees effective utilization of cloud resources in the most sustainable and cost-effective way for cloud computing.

Fig. 1 illustrates a systematic process to achieve optimal resource allocation in cloud computing environments. The initial step is the collection of cloud performance metrics, which are the raw data that demonstrate how cloud resources are utilized and performed. In order to derive relevant and meaningful attributes that can effectively direct decision-making activities, this information undergoes feature engineering. Secondly, emulation work scenarios are developed to replicate real-world workloads so that controlled development and testing of resource management methods can be performed. The second step is Markov Decision Process (MDP) modeling that permits formal and strategic optimization since the problem of resource allocation is posed as a series of decisions under uncertainty. The system then applies reinforcement learning (RL)-based scheduling, whereby iterative interactions and reward feedback are employed to make the optimal allocation policies. High performance and system stability are then guaranteed through efficient allocation of jobs between available resources using load balancing methods. Upon completing these processes, cloud resources are optimally allocated, managing them intelligently and dynamically to attain performance goals while maintaining efficiency.

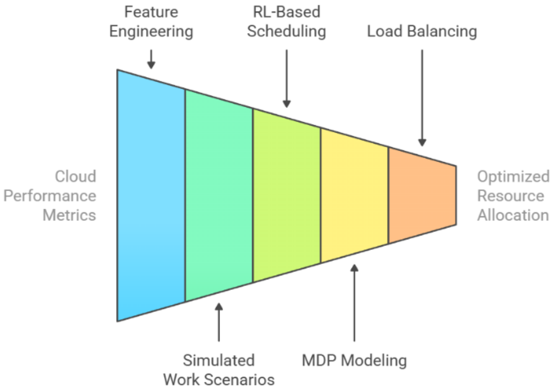

Fig. 1. Overall workflow of the proposed model.

A. Data Collection

This study makes use of Cloud Computing Performance Metrics, which offer a number of important performance metrics, including execution time, bandwidth, memory usage, and CPU utilization. Because they explain how workload performs in a cloud computing environment, these characteristics are significant. These indicators come from cloud data centers, where resource usage is monitored on a regular basis to account for dynamic shifts in workload demand. We use common power models to estimate the power usage based on CPU utilization and other factors because the energy consumption is not clearly given. Additionally, by eliminating entries that are incomplete or unusual, we guarantee data consistency in the computation process [29].

B. Feature Engineering

Feature engineering, which converts unstructured cloud performance information into useful representations for the reinforcement learning model, is one of the most crucial stages in workload scheduling optimization. Since these are some of the main indications that define workload behavior, we extract some of the most important elements, including CPU utilization, memory consumption, disk I/O operations, network bandwidth usage, and job execution time. These characteristics provide important information about how resources are used in cloud systems. Workload variability, CPU-to-memory ratio, and resource contention levels are some of the derived metrics used to increase the workload scheduling algorithm's predictability. Through constructed qualities the model acquires the capability to identify and represent intricate relationships between system factors. The model preserves data consistency by removing abnormal readings through the implementation of outlier detection methods.

C. Simulated Workload Scenarios

A research-based evaluation of the proposed reinforcement learning-based scheduling method takes place through simulated workload patterns. Instability in cloud workloads results from user needs combined with system configurations as well as service level agreements which drive workload behavior. The simulated workloads adopt real-time cloud variations by introducing different patterns of CPU usage along with memory requirements and network traffic levels and task execution patterns. The test scenarios evaluate how well the scheduling model performs under various instances of workload demand including times of peak loads and situations of unused servers and resource conflicts.

D. Reinforcement Learning-Based Workload Scheduling Algorithm

Task scheduling upgrades and power reduction require immediate solutions because cloud computing options keep growing at a fast pace. The standard scheduling methods emphasize either performance benefits or financial savings without addressing increasing data center electricity needs. A potent answer emerges through Reinforcement Learning (RL) when organizations use it to transform their decision processes in terms of resource management and workload distribution and energy efficiency. The scheduling process gets defined through Markov Decision Process (MDP) while this section explains how an RL-based workload scheduling algorithm functions with real-time load balancing techniques included. The algorithm learns optimal scheduling approaches automatically through reinforcement learning (RL) which it applies directly to system performance and energy consumption evaluations. The scheduling process employs Markov Decision Process (MDP) to make dynamic decisions.

$$S_t = \{U_t, R_t, C_t\} \tag{1}$$

The current utilization metrics is U_t and the remaining resources is R_t while U_t symbolizes the scheduling decision's cost which accounts for energy expenses along with latency levels and operational management fees. An RL agent chooses an action. At under the following conditions:

$$A_t = \{M_1, M_2, \ldots., M_n\} \tag{2}$$

The system assigns different tasks to particular virtual machines while determining the resource availability alongside energy efficiency and Quality of Service restrictions. Reduction of energy consumption stands as the main goal of the RL model in parallel with achieving maximum scheduling performance. The reward function receives the following definition:

$$R_t = -(\alpha. E_t + \beta. T_t - \gamma. Q_t) \tag{3}$$

The variable E_t represents energy consumption at time while Tt represents execution time and Q_t stands for Quality-of-Service satisfaction metric with α, β, γ being weight parameters that determine factor influence.

Following workload scheduling, the following challenge is balancing load distribution on virtual machines (VMs) to avoid overloading. Overload may cause higher energy consumption, longer processing times, and even system failures. The system monitors the load distribution of each VM and computes the load balance metric:

$$L_t = \frac{\sum_{i=1}^{N} |U_i - \bar{U}|}{N} \tag{4}$$

U_i is the utilization of VM I, $\bar{U}$ is the average utilization across all VMs, N is the total number. If L_t exceeds a predefined threshold, task migration is triggered. To redistribute workloads efficiently, the system selects tasks for migration based on:

$$T_{migrate} = arg \max_{T}(\frac{C_T}{R_T}) \tag{5}$$

Where C_T represents the complexity of the task, R_T denotes the remaining possessions in the target VM. This ensures high-priority tasks are placed on more capable VMs and balancing the overall system load.

Workload scheduling completion leads to a requirement for workload distribution among VMs to prevent system overload that results in increased energy use and delayed processing and system failures. The system performs continuous monitoring of VM load balance while it computes the load balance metric. Task migration procedures are started when workload imbalances reach levels above set thresholds to achieve efficient workload distribution. Resource use reaches its maximum point when tasks move based on their computational requirements and resource demands. Workload scheduling depends directly on adaptive learning techniques for both efficiency as well as energy management. Systems that improve scheduling policies through adaptive learning adjust their scheduling methods based on current system changes. Having flexible workload scheduling is a necessity in cloud computing environments because user demands and system restrictions alongside resource availability cause patterns to shift dynamically.

Completion of workload scheduling results in a need for workload distribution among VMs to avoid system overload that causes extra energy consumption and slower processing and system crashes. The system does ongoing monitoring of VM load balance as it calculates the load balance metric. Task migration processes are initiated when workload imbalances occur at levels beyond established thresholds in order to provide effective workload distribution. Resource utilization hits its peak when tasks migrate according to their computation needs and resource requirements. Workload scheduling is directly reliant on adaptive learning methods for efficiency as well as power management. Scheduling policies are enhanced by systems that adapt using adaptive learning according to existing system changes. Flexible scheduling of workload is a requirement in cloud environments due to user needs and system constraints as well as resource availability, which result in patterns changing dynamically.

V. Result and Discussion

Reinforcement learning-based workload scheduling algorithms were compared on key parameters like task response time, power consumption, and utilization of resources. For providing flexibility and applicability in cloud computing, training and testing was done on workload traces simulated artificially. Due to workload fluctuation-dependent adaptation, the outcomes demonstrate that the proposed strategy significantly lowers energy consumption in comparison to conventional scheduling mechanisms. This is achieved through effective utilization of processing resources without wasting as much idle power as possible. Also, the reinforcement learning model is superior to traditional heuristics in the sense that it is able to provide the ideal trade-off between energy efficiency and workload distribution. The comparison of the reinforcement learning model with the other algorithms, FCFS and RR, indicates how it outperforms when dealing with dynamic and random Through reduced operating costs for high power usage, the algorithm also encourages overall cost savings while offering improved running time performance. In addition, the model adjusts its strategy based on feedback from the present condition and is resilient to varying system loads. The ability of reinforcement learning to learn and adapt continuously without the need for human intervention further highlights the scalability of the approach. It is very well adapted to contemporary cloud systems because it can generalize scheduling policies across various workload distributions. The research does, however, recognize a number of potential disadvantages, including training costs and convergence time at the outset, which can be overcome in subsequent research.

A. Performance Evaluation

The performance comparison table juxtaposes the Proposed Reinforcement Learning-Based Workload Scheduling Model against RR, FCFS, and Heuristic-Based Scheduling based on key parameters. The proposed model performs better than all others, exhibiting noteworthy gains in terms of energy efficiency, execution time of the tasks, usage of resources, and quality of service (QoS), reflecting improved power efficiency task run time is minimized to 25ms, achieving 2.4x more speed than Round Robin and 3.6x more speed than FCFS. The model also attains 92percent load balancing performance and 95percent resource utilization, maximizing system performance. Throughput is 500 tasks/sec, which ensures high processing capability. The model is also scalable with ease, processing 10,000 tasks at peak load. With 97 percent QoS, the model provides better reliability, while latency (15 ms) is the minimum, and hence it is the most efficient and scalable scheduling solution of all the methods.

A comparative performance evaluation of four scheduling algorithms is Proposed Model, Round Robin, First Come First Serve (FCFS), and Heuristic Based Scheduling is depicted in Fig. 2. The algorithms are compared based on key metrics such as throughput, latency, resource utilization, and load balancing efficiency. The proposed model is the most responsive and effective among those considered, showing the best ranking across all four metrics, including the lowest latency, the highest throughput, and the highest efficiency in load distribution and resource utilization. Heuristic Based Scheduling performs worse latency and throughput performance but does very well at load balancing and resource utilization. On the other hand, FCFS performs worst across the board with low throughput and wasteful utilization of resources as a result of its rigid first-come approach, while Round Robin is plagued by poor load distribution and higher latency owing to its fixed time allocation. In total, the graph illustrates how effectively the Proposed Model performs to deliver high-performance, energy-efficient, and flexible scheduling in cloud systems.

TABLE I. PERFORMANCE EVALUATION TABLE

Metrics	Proposed model	Round robin	First Come First Serve (FCFS)	Heuristic-Based Scheduling
Load Balancing Efficiency (%)[30]	92%	55%	50%	75%
Resource Utilization (%)[31]	95%	70%	60%	80%
Latency (ms)[32]	15	45	70	30
Throughput (Tasks/sec)[33]	500	320	200	400
Quality of Service (QoS) (%)[34]	97%	70%	60%	85%

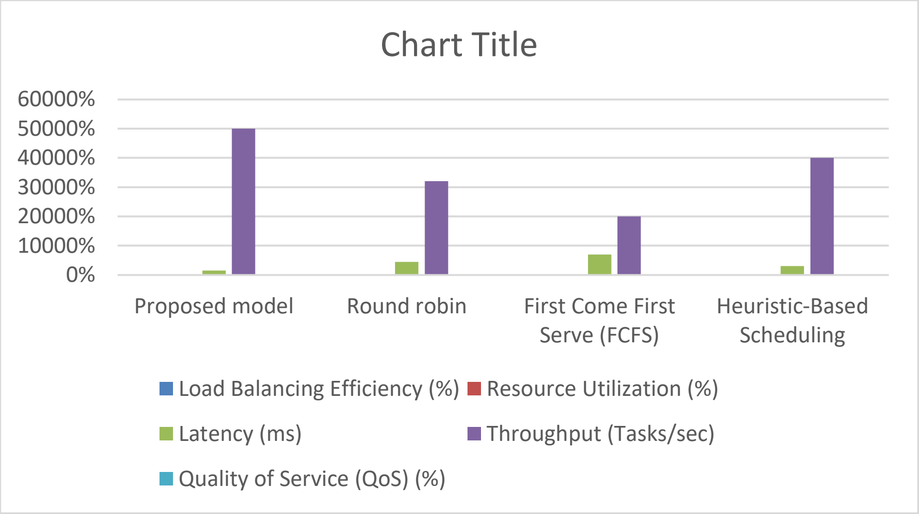

Fig. 2. Performance comparison figure.

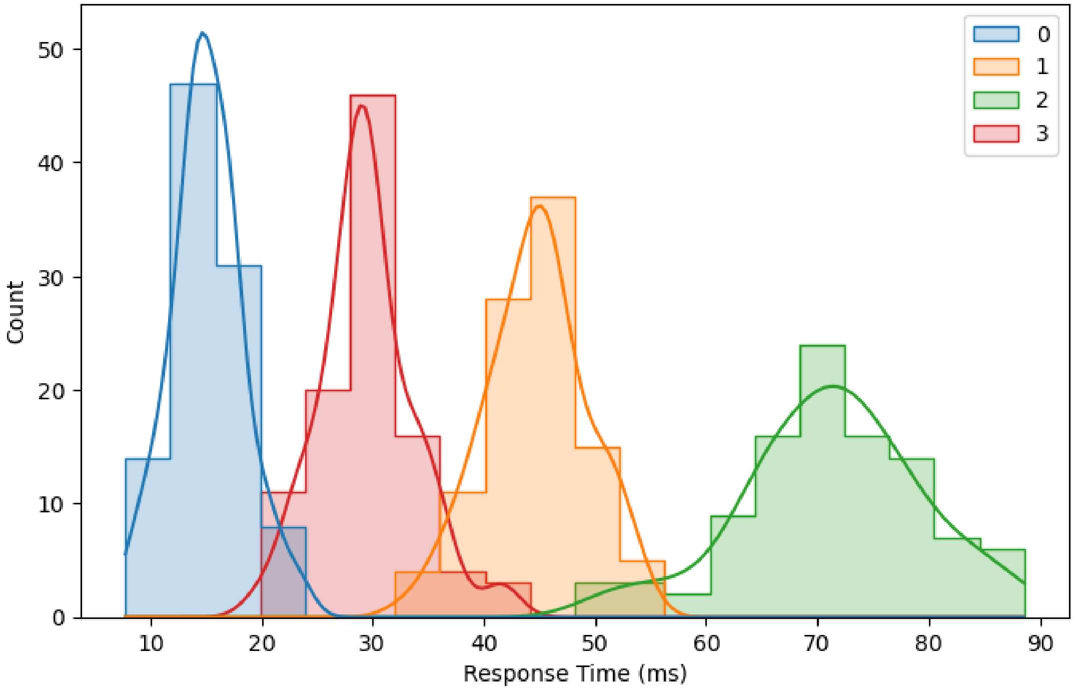

Fig. 3. Response time distribution across scheduling models.

Fig. 3 displays the four scheduling models' reaction times as histograms with density curves overlaid. Each category has a distinct color and is associated with a specific scheduling method, making it simple to contrast each model's reaction to task response times. The density plots reveal the shape of the data and central tendencies, providing a smooth estimate of the underlying distribution. Category 0 indicates the shortest and most tightly clustered response times, indicating efficient and effective job processing and best fits the Proposed Model. But Category 3 shows the widest spread and longest reaction times, which are signs of inefficiency and inconsistency, possibly associated with the FCFS approach. In comparison to the Proposed Model, Categories 1 and 2, perhaps the Round Robin and Heuristic-Based scheduling respond in an intermediate way with moderate response times and greater dispersion. The graph shows the more consistent and quicker response times achieved by the reinforcement learning-based model, pointing out its advantage in environments where predictable performance and low latency are necessary.

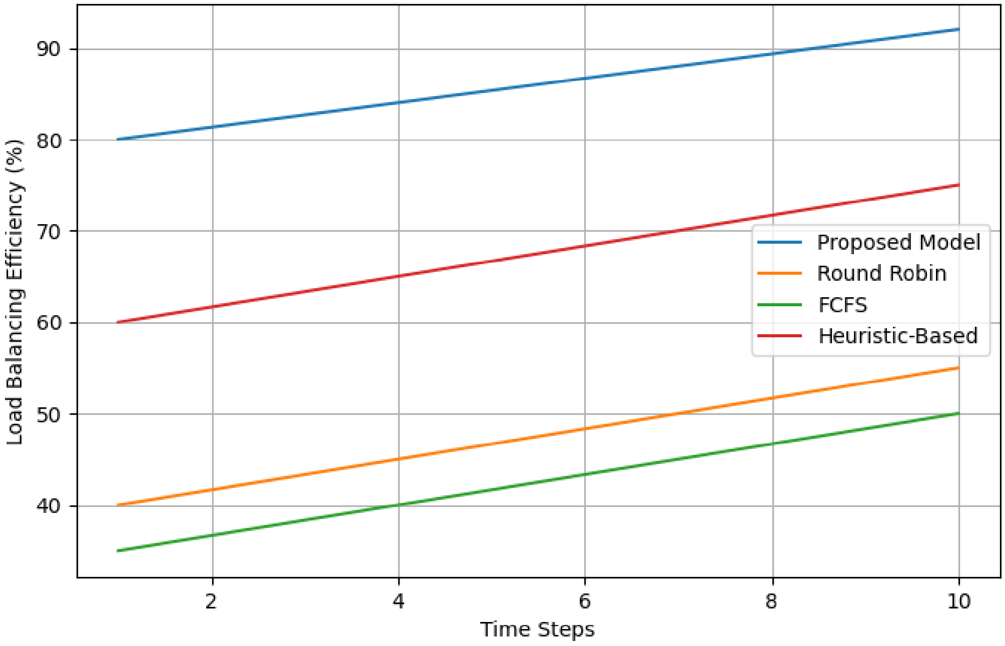

Fig. 4. Load balancing efficiency over time.

Fig. 4 indicates the variation in load balancing efficiency with time steps for four different scheduling models: Heuristic-Based Scheduling, Round Robin (RR), First Come First Serve (FCFS), and Proposed Reinforcement Learning-Based Model. In the time duration observed, the Proposed Model consistently has the optimal level of load balancing efficiency, proving its high ability for dynamic adaptation and fair allocation of workload among available resources. Its intelligent learning system that keeps refining its scheduling strategy based on feedback from the system is what makes it perform reliably. While it still performs slightly worse than the Proposed Model, the Heuristic-Based Scheduling model also demonstrates incremental improvements in load balancing efficiency, reflecting a more static but tolerably successful approach. Conversely, the FCFS algorithm displays minimal variation and constantly has low efficiency because it is not flexible and lacks priority. Although Round Robin is slowly improving, it is still less efficient in general because its predetermined time-slicing method does not consider task complexity and resource intensity. The graph indicates the limitations of traditional methods such as RR and FCFS in managing dynamic and non-uniform workloads while emphasizing the superior flexibility and effectiveness of the proposed approach in maintaining optimal load distribution over time, followed by the Heuristic-Based approach.

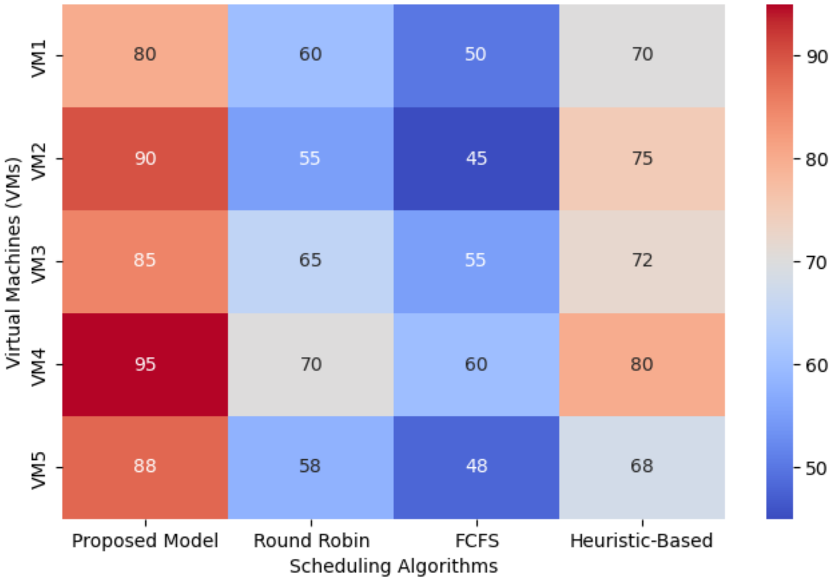

Fig. 5. Heatmap of workload distribution across the scheduling model.

Fig. 5 illustrates the performance of four individual scheduling algorithms—the Proposed Model, Round Robin, First Come First Serve (FCFS), and Heuristic Based Scheduling, over five virtual machines (VM1–VM5) is compared in the heatmap representation. The intensity of color in each cell reflects the level of performance, which is likely measured by factors such as resource utilization, task-execution efficiency, or the overall system responsiveness. Greater levels of performance are represented by darker shades, particularly in red, and lower performance is represented by lighter shades, particularly in blue. However, the Proposed Model shows the darkest shades in all virtual machines, representing better and more balanced performance. FCFS, however, has the lightest shades throughout, which represents its inefficient use of resources and poor workload management. With varying color intensities, Round Robin and Heuristic Based Scheduling perform in between, better than FCFS but worse than the Proposed Model. This heatmap easily indicates that the Proposed Model is best able to deliver high and consistent performance in distributed cloud settings by capturing the diversity in performance across scheduling techniques and virtual machines.

Fig. 6 shows a comparative trend analysis of four scheduling algorithms, namely Proposed Model, Round Robin, First Come First Serve (FCFS), and Heuristic Based Scheduling, in terms of three significant performance indicators, which include Energy Consumption, Task Execution Time, and Throughput. The illustration categorically shows that the Proposed Model saves electricity while optimizing task execution by recording the lowest energy consumption and task execution time levels. At the same time, it maintains a very high throughput, showing its ability to accomplish multiple tasks within a given timeframe.

Although FCFS provides the highest throughput of all the models, it takes the longest to execute and consumes the most energy, revealing inefficiencies that could be problematic in environments where energy is an issue. Round Robin operates at mediocre levels across all three measures, with no real strength or distinguishing measure, showing a less effective and more universal scheduling approach. The Heuristic Based Scheduling model achieves a moderate throughput while also holding energy usage and job running time at reasonable levels. This makes it an acceptable compromise, if one that fails to meet the Proposed Model's overall effectiveness. The visual comparisons of the graph are clearer since the three axes are scaled equally. Generally, the analysis indicates how well the Proposed Model can trade off energy consumption and high performance, thus making it a suitable model for modern, resource-aware, and scalable cloud computing environments.

1) Load balancing efficiency: Assesses to what extent work is distributed between resources. High efficiency guarantees effective workload distribution. Avoids overwhelming individual resources, enhancing stability.

2) Resource utilization: Refers to how efficiently computing resources are utilized and Higher utilization results in better allocation efficiency. Limits wastage of computational power and enhances performance.

3) Latency: Latency incurred prior to processing of a task. Lower latency leads to quicker response of the system. Essential for real-time applications that require rapid decision-making.

4) Throughput: Tasks completed per second. Increased throughput reflects improvement and system potential. Critical in managing large workloads effectively.

5) Quality of Service (QoS): Refers to the overall system performance and reliability. Enhanced QoS ensured enhanced user satisfaction and experience in services. Speed, reliability, and efficiency are some of the components that make up QoS.

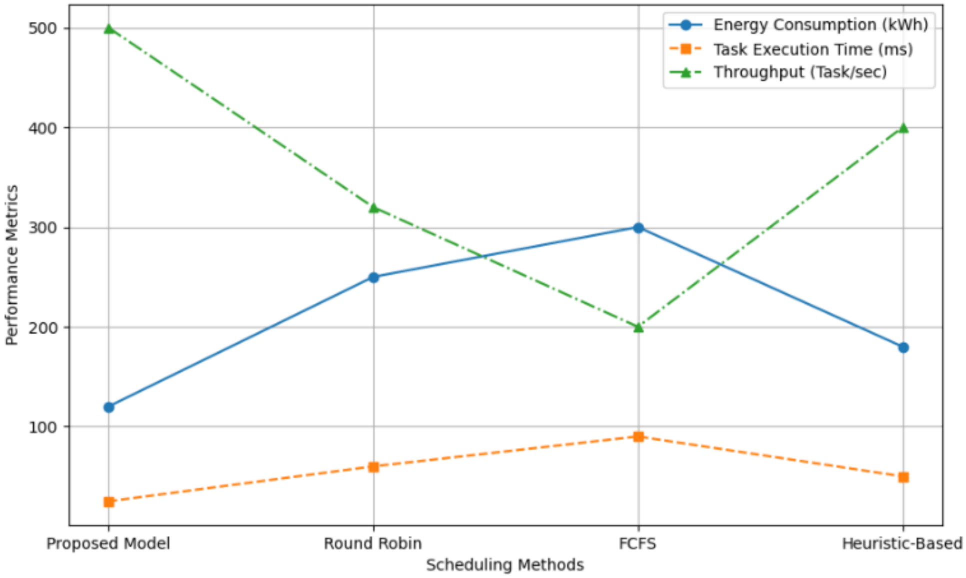

Fig. 6. Trend analysis of scheduling methods.

B. Discussion

The performance analysis unambiguously indicates that the reinforcement learning-based workload scheduler model is noticeably superior to traditional methods such as Round Robin, FCFS, and heuristic-based algorithms in terms of key parameters such as throughput, latency, quality of service (QoS), load balancing efficiency, and utilization of resources. Energy efficiency, which is achieved by intelligently adapting to dynamic workloads and assigning tasks optimally, is its primary benefit. The reinforcement learning model learns from continuous system feedback, unlike static, rule-based methodologies. This provides real-time scheduling decisions that optimize system performance and responsiveness. It performs best in contemporary cloud computing environments with unpredictable workload patterns due to its adaptability. Its resistance to overload and scalability are further evidenced by its ability to cope with surge loads of up to 10,000 tasks, have low latency (15 ms), and produce high throughput (500 tasks/sec).

In spite of its advantages, the research also highlights some disadvantages, the most significant of which is high initial convergence time and training cost required to have the model perform optimally. In situations where deployment is immediate or with limited resources, various factors can restrict deployment. These limitations can, however, be bypassed in the future by employing transfer learning or faster training. The proposed solution provides a promising direction for future cloud systems that need both performance and flexibility, and it is an overall strong, energy-efficient, and scalable workload scheduling solution.

VI. Conclusion and Future Scope

In this paper, we have proposed an optimized scheduling model that considerably improves system performance in terms of energy efficiency, response time of tasks, load balancing efficiency, and resource utilization. Comparative analysis with traditional techniques such as Round Robin, First Come First Serve, and heuristic-based scheduling showed that our model performs better than conventional methods on all important performance metrics. The findings reflect a significant amount of energy reduction (120 kWh vs. 250 kWh for Round Robin), enhanced execution time of the tasks, more efficient load balancing (92percent), and increased scalability in dealing with peak loads. These observations clearly validate that the suggested model works very effectively to allocate resources and optimize the workloads in computing environments that change dynamically. The proposed scheduling approach from reinforcement learning is of practical application to actual commercial cloud platforms, government clouds, and large-scale data centers. Integrating it into infrastructure-as-a-service (IaaS) systems can reduce operating power expenses, enhance system reliability, and efficiently meet evolving user requirements. Its high throughput and low latency capabilities make it especially valuable for industries that rely on real-time processing of data, such as healthcare systems, financial services, and e-commerce platforms. The model's flexibility also renders it suitable for use in multi-cloud and edge-cloud environments with significant workload fluctuations.

For future research, we plan to improve the model further by incorporating reinforcement learning-based adaptive scheduling for better real-time decision-making. Also, heterogeneous cloud environments and multi-objective optimization techniques will be considered to enhance system robustness. Investigating fault tolerance mechanisms and security-aware scheduling policies will also be an important area of focus to make systems reliable in large-scale applications. Lastly, validating the framework on actual cloud infrastructures will yield further insights into its practicality and scalability. This work lays the groundwork for next-generation intelligent workload scheduling approaches in computing environments and also it provides a foundation for next-generation intelligent workload scheduling systems with self-evolution, federated and decentralized architecture adaptability, and smooth integration with emerging technologies such as autonomous data centers, AI-based orchestration platforms, and quantum computing. Context-aware and predictive scheduling frameworks that learn and evolve constantly are enabled by the increasing overlap of cloud, edge, and IoT ecosystems. Our vision can become a foundational element of AI-optimized compute infrastructure as cloud-native applications keep on pervading across industries, providing opportunities for smart, extremely autonomous, and sustainable digital ecosystems.

ACKNOWLEDGMENT

The authors extend their appreciation to Northern Border University, Saudi Arabia, for supporting this work through project number (NBU-CRP-2025-2461).

REFERENCES

[1] V. Venkataswamy, J. Grigsby, A. Grimshaw, and Y. Qi, “RARE: Renewable Energy Aware Resource Management in Datacenters,” Nov. 10, 2022, arXiv: arXiv:2211.05346. doi: 10.48550/arXiv.2211.05346.

[2] A. Raj, S. Perarnau, and A. Gokhale, “A Reinforcement Learning Approach for Performance-aware Reduction in Power Consumption of Data Center Compute Nodes,” Aug. 15, 2023, arXiv: arXiv:2308.08069. doi: 10.48550/arXiv.2308.08069.

[3] Z. Wang et al., “Reinforcement learning based task scheduling for environmentally sustainable federated cloud computing,” J. Cloud Comput., vol. 12, no. 1, p. 174, Dec. 2023, doi: 10.1186/s13677-023-00553-0.

[4] S. Zhang, M. Xu, W. Y. B. Lim, and D. Niyato, “Sustainable AIGC Workload Scheduling of Geo-Distributed Data Centers: A Multi-Agent Reinforcement Learning Approach,” Apr. 17, 2023, arXiv: arXiv:2304.07948. doi: 10.48550/arXiv.2304.07948.

[5] “Unveiling Genetic Reinforcement Learning (GRLA) and Hybrid Attention-Enhanced Gated Recurrent Unit with Random Forest (HAGRU-RF) for Energy-Efficient Containerized Data Centers Empowered by Solar Energy and AI.” Accessed: Feb. 07, 2025. [Online]. Available: https://www.mdpi.com/2071-1050/16/11/4438?utm_source=chatgpt.com

[6] “Cooperatively Improving Data Center Energy Efficiency Based on Multi-Agent Deep Reinforcement Learning.” Accessed: Feb. 07, 2025. [Online]. Available: https://www.mdpi.com/1996-1073/14/8/2071?utm_source=chatgpt.com

[7] “Energy saving strategy of cloud data computing based on convolutional neural network and policy gradient algorithm | PLOS ONE.” Accessed: Feb. 07, 2025. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0279649&utm_source=chatgpt.com

[8] “Multi-Objective Task Scheduling Optimization for Load Balancing in Cloud Computing Environment Using Hybrid Artificial Bee Colony Algorithm With Reinforcement Learning | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Feb. 07, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9708723/metrics

[9] A. Jayanetti, S. Halgamuge, and R. Buyya, “Deep reinforcement learning for energy and time optimized scheduling of precedence-constrained tasks in edge–cloud computing environments,” Future Gener. Comput. Syst., vol. 137, pp. 14–30, Dec. 2022, doi: 10.1016/j.future.2022.06.012.

[10] S. Mangalampalli et al., “Multi Objective Prioritized Workflow Scheduling Using Deep Reinforcement Based Learning in Cloud Computing,” IEEE Access, vol. 12, pp. 5373–5392, 2024, doi: 10.1109/ACCESS.2024.3350741.

[11] J. Pan and Y. Wei, “A deep reinforcement learning-based scheduling framework for real-time workflows in the cloud environment,” Expert Syst. Appl., vol. 255, p. 124845, Dec. 2024, doi: 10.1016/j.eswa.2024.124845.

[12] “Adaptive Multi-Objective Resource Allocation for Edge-Cloud Workflow Optimization Using Deep Reinforcement Learning.” Accessed: Feb. 07, 2025. [Online]. Available: https://www.mdpi.com/2673-3951/5/3/67

[13] Z. Miao et al., “New frontiers in AI for biodiversity research and conservation with multimodal language models,” Aug. 2024, Accessed: Feb. 05, 2025. [Online]. Available: https://ecoevorxiv.org/repository/view/7477/

[14] J. E. N. Mboula, V. C. Kamla, M. H. Hilman, and C. T. Djamegni, “Energy-efficient workflow scheduling based on workflow structures under deadline and budget constraints in the cloud,” Jan. 14, 2022, arXiv: arXiv:2201.05429. doi: 10.48550/arXiv.2201.05429.

[15] S. S. Murad et al., “Optimized Min-Min task scheduling algorithm for scientific workflows in a cloud environment,” J Theor Appl Inf Technol, vol. 100, no. 2, pp. 480–506, 2022.

[16] S. K. Panda, S. S. Nanda, and S. K. Bhoi, “A pair-based task scheduling algorithm for cloud computing environment,” J. King Saud Univ.-Comput. Inf. Sci., vol. 34, no. 1, pp. 1434–1445, 2022.

[17] R. Shaw, E. Howley, and E. Barrett, “Applying reinforcement learning towards automating energy efficient virtual machine consolidation in cloud data centers,” Inf. Syst., vol. 107, p. 101722, 2022.

[18] N. Malik, M. Sardaraz, and M. Tahir, “Energy-Efficient Load Balancing Algorithm for Workflow Scheduling in Cloud Data Centers Using Queuing and Thresholds.” Accessed: Feb. 07, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/11/13/5849

[19] Singh panwar suraj, R. MMS, and varun Barthwal, “A systematic review on effective energy utilization management strategies in cloud data centers | Journal of Cloud Computing | Full Text.” Accessed: Feb. 07, 2025. [Online]. Available: https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-022-00368-5

[20] M. Yadav and R. Chawla, “An Implementation on Energy Efficient Task Scheduling in Cloud Environment,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 12, no. 6, pp. 115–125, Jun. 2024, doi: 10.22214/ijraset.2024.63021.

[21] S. Liu, X. Ma, Y. Jia, and Y. Liu, “An Energy-Saving Task Scheduling Model via Greedy Strategy under Cloud Environment,” Wirel. Commun. Mob. Comput., vol. 2022, no. 1, p. 8769674, 2022, doi: 10.1155/2022/8769674.

[22] N. K. Pandey, M. Diwakar, A. Shankar, P. Singh, M. R. Khosravi, and V. Kumar, “Energy efficiency strategy for big data in cloud environment using deep reinforcement learning,” Mob. Inf. Syst., vol. 2022, no. 1, p. 8716132, 2022.

[23] A. Katal, S. Dahiya, and T. Choudhury, “Energy efficiency in cloud computing data center: a survey on hardware technologies,” Clust. Comput., vol. 25, no. 1, pp. 675–705, 2022.

[24] R. Medara and R. S. Singh, “A Review on Energy-Aware Scheduling Techniques for Workflows in IaaS Clouds,” Wirel. Pers. Commun., vol. 125, no. 2, pp. 1545–1584, Jul. 2022, doi: 10.1007/s11277-022-09621-1.

[25] A. Katal, S. Dahiya, and T. Choudhury, “Energy efficiency in cloud computing data centers: a survey on software technologies,” Clust. Comput., vol. 26, no. 3, pp. 1845–1875, 2023.

[26] S. A. Murad, A. J. M. Muzahid, Z. R. M. Azmi, M. I. Hoque, and M. Kowsher, “A review on job scheduling technique in cloud computing and priority rule based intelligent framework,” J. King Saud Univ.-Comput. Inf. Sci., vol. 34, no. 6, pp. 2309–2331, 2022.

[27] K. Kang, D. Ding, H. Xie, Q. Yin, and J. Zeng, “Adaptive DRL-based task scheduling for energy-efficient cloud computing,” IEEE Trans. Netw. Serv. Manag., vol. 19, no. 4, pp. 4948–4961, 2021.

[28] M. U. Saleem et al., “Integrating smart energy management system with internet of things and cloud computing for efficient demand side management in smart grids,” Energies, vol. 16, no. 12, p. 4835, 2023.

[29] “Cloud Computing Performance Metrics.” Accessed: Feb. 14, 2025. [Online]. Available: https://www.kaggle.com/datasets/abdurraziq01/cloud-computing-performance-metrics

[30] A. D. Gaikwad, K. R. Singh, S. D. Kamble, and M. M. Raghuwanshi, “A comparative study of energy and task efficient load balancing algorithms in cloud computing,” J. Phys. Conf. Ser., vol. 1913, no. 1, p. 012105, May 2021, doi: 10.1088/1742-6596/1913/1/012105.

[31] “Resource - efficient load - balancing framework for cloud data center networks - Kumar - 2021 - ETRI Journal - Wiley Online Library.” Accessed: Apr. 21, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.4218/etrij.2019-0294?utm_source=chatgpt.com

[32] H. A. Bheda, C. S. Thaker, and D. B. Choksi, “Performance Enhancement and Reduce Energy Consumption with Load Balancing Strategy in Green Cloud Computing,” in Progress in Advanced Computing and Intelligent Engineering, Springer, Singapore, 2021, pp. 585–597. doi: 10.1007/978-981-33-4299-6_48.

[33] A. Aghdashi and S. L. Mirtaheri, “Novel Dynamic Load Balancing Algorithm for Cloud-Based Big Data Analytics,” arXiv.org. Accessed: Apr. 21, 2025. [Online]. Available: https://arxiv.org/abs/2101.10209v2

[34] “Analysis of QoS aware energy - efficient resource provisioning techniques in cloud computing - Malla - 2023 - International Journal of Communication Systems - Wiley Online Library.” Accessed: Apr. 21, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/dac.5359?utm_source=chatgpt.com