International Journal of Computational Intelligence in Engineering (IJCIE)
Vol. 01, No. 03, May 2026, pp. 22-30
ISSN: XXXX-XXXX, DOI: 10.5281/zenodo.20395041

A Novel Virtual Machine Migration Model for Optimizing Cloud Resource Utilization

Anitha R

Professor & Head, Dept. of CSE, The National Institute of Engineering, Mysuru, Karnataka, India. Email: anithar@nie.ac.in

Article Info

Article history:

Received Apr 15, 2026
Revised May 20,2026
Accepted May 23, 2026

Keywords:

Load Balancing
Cloud Computing
Energy-Efficient Computing
Virtual Machine Migration
Resource Utilization
SLA-Aware Optimization

ABSTRACT

Cloud Computing has emerged as an important paradigm for delivery of scalable computing and storage resources using virtualization techniques. As cloud workloads grow rapidly, however, virtual machine (VM) overloading, energy usage, SLA violations and low system performance are common results. To address these issues, this work proposes an adaptive VM migration design for efficient resource utilization and VM load-balancing in a cloud environment. First, the proposed approach proposes to allocate VMs based on resources, taking into account the availability of processor resources, memory, and network conditions. A threshold-based migration strategy is then used to identify the overloaded VMs and migrate them automatically to appropriate physical machines which have lesser usage. Furthermore, to reduce the migration overhead and enhance the computation efficiency, the cost estimation based on latency, communication delay and hop distance are introduced for migration. An optimization model is also formulated to optimize the use of the energy and minimize the migration cost. Up to 200 virtual machines, with various host utilization thresholds and 40 simulation runs, are used in experimental analysis. Results show that the proposed method is able to decrease the number of active hosts, energy consumption and the frequency of VM migration when compared to other approaches like FFO-EVM and ACO. From the results obtained, it is concluded that the proposed adaptive migration framework is effective in enhancing the performance of the cloud, utilization of resources and efficiency of task completion with low computational overhead.

Corresponding Author:

Anitha R
Professor & Head, Dept. of CSE, The National Institute of Engineering, Mysuru, India.
Email: anithar@nie.ac.in

1. INTRODUCTION

In recent two decades, the significant expansion of Internet and web-based communication systems has created huge quantities of structured and unstructured data, which need to be efficiently processed, stored and computed [1]. To overcome these challenges, cloud computing has become a scalable and cost-effective computing paradigm that delivers virtualized computing resources via service modes like Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [2]. Cloud computing allows for scalable computation, efficient resource sharing, and dynamic task execution, all facilitated by the use of virtualization technologies. The primary goal of virtualization is to make better use of physical resources and computational efficiency by allowing multiple operating systems and applications to operate on a single set of physical resources. Another benefit of live virtual machine (VM) migration is that overloaded machines can move workloads to underutilized machines, which enhances the overall cloud performance and task completion [3].

Despite all these benefits, there are some challenges in cloud data centers, like workload volatility, resource imbalance, communication overhead, energy usage and SLAs violation. Traditional physical server architectures may have disadvantages, such as low resource efficiency, high maintenance cost and energy waste. Such problems can be addressed by applying virtualization and load balancing techniques that dynamically allocate loads to available servers to achieve scalability, responsiveness and resource efficiency [4-5]. But, variability of workloads and dynamic traffic or network patterns still impact performance and stability of cloud systems.

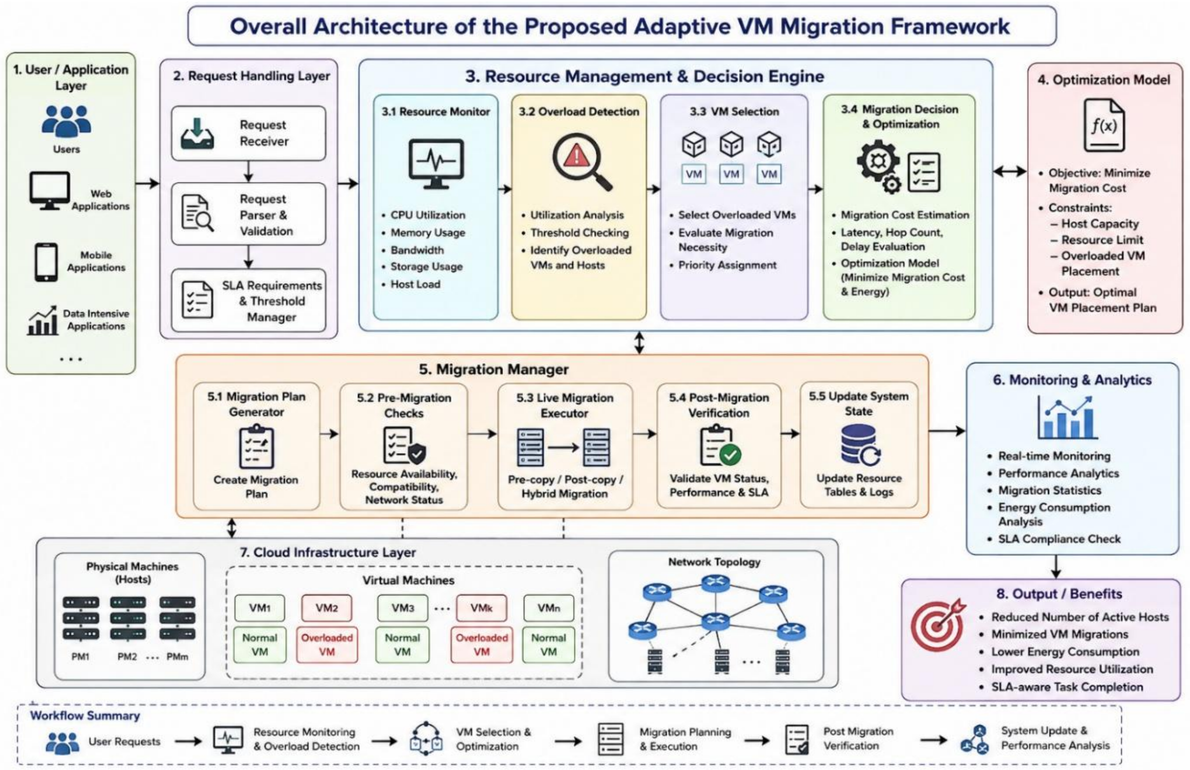

Fig. 1. Architecture diagram of proposed system

Virtualization is an important technique to improve cloud performance by consolidating servers and managing VMs [6]. As cloud services become more popular, the need for efficient VM placement and migration to this cloud environment is becoming critical for ensuring the system's performance and lowering operational expenses. Traditional scheduling methods provide better resource allocation however they do not provide good management of dynamic traffic demands and migration overhead [7]. VM migration schemes [8] and adaptive local resizing approaches [9] have been used to tackle these issues. There are existing migration techniques such as pre-copy, post-copy, and hybrid migration techniques which help to minimize migration time and downtime, but there is still a room for improvement in the minimization of resources usage and communication overheads [10]. Tiwari et al. [11] also pointed out that load balancing combined with VM migration is crucial for ensuring performance of cloud systems.

To address the aforementioned issues, in this paper, the authors propose an adaptive VM migration framework to optimize resource usage, minimize migration overhead, and maximize energy efficiency in the cloud. The proposed solution involves VM allocation and migration based on resource awareness and thresholds, respectively, which will lead to efficient VM allocation and task completion as well as better cloud resource management. The main contributions of this work are as follows:

- Introduction to cloud computing and virtualization techniques,
- Discussion of recent VM migration approaches and trends, and
- Creation of an adaptive migration framework for performance enhancement in virtualization and migration, and
- Comparative experimental analysis to assess the effectiveness of the proposed model.

The rest of the paper is organized as follows: Section II discusses recent studies related to cloud computing and VM migration, Section III describes the methodology proposed in this work, Section IV discusses the experimental results and comparisons, and Section V concludes the work.

2. LITERATURE SURVEY

A brief discussion about recent techniques for load balancing, virtualization and VM migration is presented in this section. The area of Cloud computing has been seen as an attractive area of research for offering efficient computing and storage space to clients. To undertake these kinds of efforts, cloud service providers must build up several data centers. Such data centers are in charge of carrying out the assigned tasks. During computation process, huge amount of energy is required which cases more power consumption [12]. Likewise, in order to improve the performance of the cloud systems and satisfy the clients, the assigned task should be completed within a stipulated time period. Because of the growing demand for Cloud computing based system, lots of tasks are assigned which results in overhead and leads to degrade performance. To overcome this issue, load balancing is considered as most promising scheme [13]. Energy consumption, however, is still a difficult research field in this area. Recently, several techniques have been presented to overcome the issue of power consumption which include migration schemes and without migration schemes. During this process, live VM migration is taken into consideration as a major technology for lowering the energy consumption [14]. VM Migration process is as per, If VM is overloaded or down, VM is transferred to another active node. The process has the significant impact on energy saving in cloud computing systems. Generally, VM migration schemes are not as simple because it includes other stages such as VM placement [15], VM consolidation [16] etc.

In this study, Khosravi et al. [17] explained the energy consumption problems associated with cloud data center and also calculated the amount of energy that can be saved with the use of renewable energy in data centers. Generally, renewable energy has irregular nature of production which may cause issues for power supply and generation. So, virtual machine migration close to the datacenter may help to achieve the desired performance. Authors in this work introduce an offline algorithm, in which the level of the energy sources is known to be renewable.

Duggan et al. [18] concentrated on the live VM migration for cloud computing system and also discussed that the consumption of more network resources may result into the more time requirement for VM migration. Currently, various industries use manual process for VM migration which is addressed in this work and network aware live VM migration scheme is presented which considers the current demand of the network resources and performs the suitable action. In order to learn the requirement parameters, artificial intelligence based reinforcement learning scheme is introduced which learns the network demand patterns and optimal scheduling time for any given network traffic model. Another work of interest is that of Wu et al. [19] who considered the live VM migration and talked about its significance in cloud computing systems. Conventional techniques of migration are based on the pre-copy and post-copy methods of migration which require more time hence time-series based pre-copy and post-copy based methods presented here. General model of pre-copy techniques can provide the better migration by reducing the time taken for migration but it may lead to longer downtime which will cause more time for task completion. The entire migration process is divided into three phases: (a) Push phase (b) stop and copy and (c) pull phase and downtime and migration time are used as measures of migration performance. A memory prediction scheme is also presented in this approach which can select the pages to be migrated during pre-copy or stop and copy phase.

Private clouds are also extensively utilized by different industries lately. Private clouds have very limited resources hence a proper load lancing, minimal energy consumption, resource utilization is required to obtain the optimal performance from the computing architecture. Sohrabi et al [20] proposed a novel approach of VM migration in private cloud computing systems based on this assumption. In this work, a novel combined approach is presented in which selection of VMs and VM mapping are discussed together, with the aim of obtaining an energy enhancement. The problem is unpredictable in terms of the number of tasks to be processed; so, optimizing the performance is very important, crucial, tsk. Hence, a new approach with adaptive computation process is introduced where Bayesian interference is used for likelihood estimation for VM migration.

As discussed before, for VM migration consolidation and placement play important roles. Various methodologies have been presented for VM placement and consolidation to enhance the system performance. Recent growth in efficient computation of cloud computing, cloud providers such as Microsoft Azure and Amazon have come up with IaaS based new cloud model where MapReduce based schemes are utilized to perform the task for huge workloads. It is considered that the IaaS have multiple pricing option, we discuss about batch-oriented consolidation and placement of virtual machines. In this work, DVFS based TRP-FS scheme is introduced for task management and VM consolidation along with this it also guarantees to follow the SLA criteria [21].

Zola et al [22] also focused on the cloud performance optimization using datacenter virtualization and VM migration process. in order to reduce the energy consumption, cloud operators consolidate various virtual machine to a smaller number of physical servers which may lead to the SLA violations during peak load hence a proper method is required for VM migration. Hence, a novel approach presented using optimization computation process in this process, VM consolidation is modeled in the robust Mixed Integer Linear Program, which take into account the resource-based constraint of cloud. Abdullah et al. [23] proposed a heuristic method for minimizing energy consumption and meeting the SLA requirements. To allocate the VMs, fast best-fit decreasing (FBFD) algorithm is presented and dynamic utilization rate (DUR) algorithm for space utilization and VM migration is also introduced.

3. PROPOSED MODEL

This part introduces the proposed adaptive virtual machine (VM) migration framework in cloud computing environments. The main aim of the proposed model is to reduce VM overloading, Resource Overhead for migration and improve resource utilization in cloud computation. A cloud environment has a collection of virtual machines and physical machines represented as::

$$VM = \{VM_1, VM_2, \dots, VM_{n\}}$$

$$PM = \{PM_1, PM_2, \dots, PM_{m\}}$$

As workloads grow, some VMs may start to be overloaded. The overloaded VM set is described as::

$$O \subset VM$$

The cloud data center is represented as a dependency graph for communicating between VMs:

$$G = (V, E)$$

whereas V represents Virtual machines and E represents the communication links between the Virtual machines. The traffic demand between VMs is represented by:

$$W(VM_i, VM_j) \propto (VM_i, VM_j)$$

Every VM is using computational resources like CPU, memory and bandwidth, and every physical machine has a limit of resources. Migration time, communication delay and energy resource are taken into consideration during migration. The migration cost is calculated as:

$$C(VM_i, PM_k) = Dist(PM_k, PM_l) \times W(VM_i, VM_j)$$

Where $Dist(PM_k, PM_l)$ is the network distance measured by latency, delay and hops between physical machines.

To determine VM placement, a binary allocation parameter is defined as:

$$X_{ik} = \begin{cases} 1, & if\ VM_i\ is\ allocated\ to\ PM_k \\ 0, & otherwise \end{cases}$$

The optimization goal is to minimize the overhead of migration:

$$min \sum C(VM_i, PM_k) \times X_{ik}$$

In order to prevent overloading on the server, the load assigned to each physical machine should comply with:

$$\sum X_{ik} \times Load_i \leq Capacity_k$$

where $load_i$ is the workload of the VM and $Capacity_k$ is the available capacity of the physical machine. VMs are moved to other machines with sufficient resources in case of overload.

The proposed migration approach is based on the assumption that cloud applications have data intensive tasks. The data input is divided into several blocks and spread out through the cloud's storage nodes. The data distribution matrix is shown to be:

$$D_{n\times m} = d_{i,j}$$

where n is storage nodes, and m is data blocks. The overall size of data is calculated as follows:

$$DataSize = \sum_{i=1}^{n} d_{i,m}$$

The communication speed between VM and storage node is given as:

$$S = \frac{s}{\Delta t}$$

s is the size of the packet and Δt is the time it takes to transmit the packet.

The proposed system evaluates the available computer resources (processor capacity and memory) for each new task. The available resources are calculated as:

$$CR_{available} = \{PM_{available}, M_{available}\}$$

$$PM_{available} = PM - \sum PM_{VM}$$

$$M_{available} = M - \sum M_{VM}$$

The VM allocation process is a process that chooses the machine with the least amount of data transfer time that meets the required resources. Resource information is dynamically updated once allocation is done.

Cloud environments are very dynamic, and during peak load times the environment can cause network latency, potentially violating the Service Level Agreements (SLA). To tackle this problem, a threshold-based migration mechanism is presented. If $execution\_time$ and/or $transfer\_time$ is greater than the predefined SLA threshold T, then the alarm is activated. It automatically triggers migration with this option:

$$\sum T(x,y) > T_{SLA}$$

The system then looks for another physical machine with minimal transfer time and enough resources to release the VM and frees it up. In overall, the proposed adaptive VM migration framework makes the cloud more efficient by minimizing the amount of work that needs to be done during migration, balancing workloads, reducing the amount of energy consumed, and meeting SLA requirements.

4. EXPERIMENTAL STUDY

Experimental Evaluation of the Proposed Adaptive Virtual Machine (VM) Migration Framework for Cloud Computing Environments is presented in this section. Different numbers of virtual machines, node utilization thresholds, and multiple runs of execution are used to evaluate the performance of the proposed approach. The main goal of the experiments is to assess the advantages of the proposed model with respect to host utilization, VM migrations frequency, and energy consumption.

Experiments were carried out with up to 200 virtual machines under various workloads for performance evaluation. To obtain reliable and consistent results, 40 simulation runs have been performed. The proposed VM migration framework was contrasted with the present VM migration strategies like FFO-EVM and ACO-based migration strategies [24]. This study takes into account the following evaluation criteria:

- Number of times the number of active hosts is needed for the computation.
- The total number of VM migrations.
- Average energy consumption
- Resource utilization efficiency
- Task completion performance based on SLA.

4.1 Number of Hosts Performance

The comparative performance analysis (based on number of active hosts used throughout the computation process) is shown in Figure 2. This experiment was carried out by adding up virtual machines and measuring the utilization of the corresponding host.

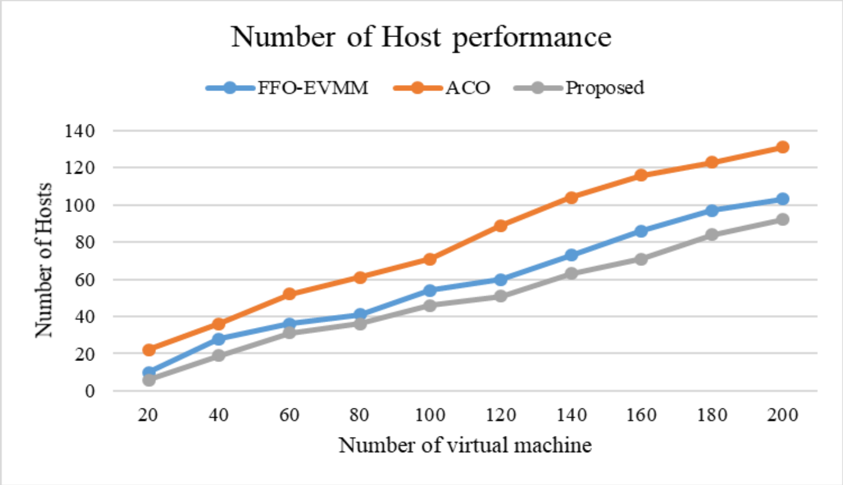

Fig 2. Number of hosts performance

The results show that the proposed adaptive VM migration approach reduces the number of active hosts as compared to the existing approaches. The improvement is made by the proposed framework which does VM allocation based on resources and efficiently consolidates workloads on a proper physical machine. The proposed approach addresses the challenge of minimizing resources consumed in the activation of the hosts, thus optimizing resources usage and minimizing infrastructure overhead. Moreover, effective host consolidation also helps in minimizing power usage and improving resource utilization in the cloud.

4.2 VM Migration Performance

Figure 3 shows the comparative analysis of the total number of VM migrations performed during execution process. VM migration is deemed as one of the significant parameters since too much migration leads to excessive communication overhead, migration delay, and system instability.

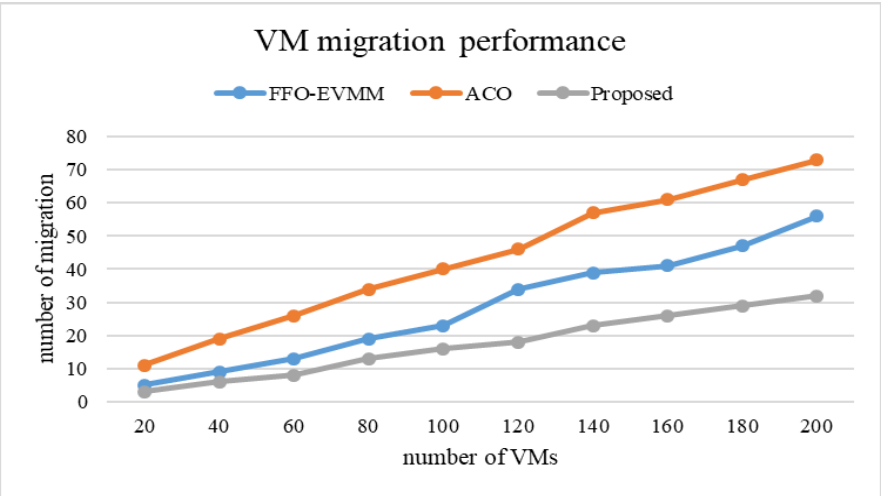

Fig.3. VM migration performance

The experimental results show that the proposed framework can significantly reduce VM migrations compared to the existing ones. In the threshold-based migration strategy, the migration only takes place if the workload is above the defined SLA threshold. This causes unnecessary migrations to be avoided, and reduces migration overhead and improves system stability. Furthermore, fewer migrations can help in achieving efficient task completion within the required execution time, as well as reduce network congestion.

4.3 Energy Consumption Performance

The comparative performance evaluation is presented in terms of average energy consumption in Figure 4 for various node utilization percentages. In cloud data centers, energy efficiency is one of the primary concerns since they need to process a significant amount of power to run large-scale computations and manage VMs.

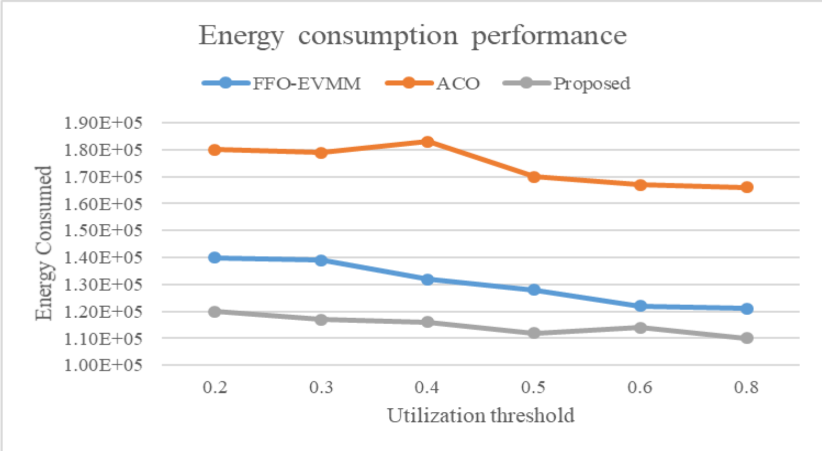

Fig. 4. Energy consumption performance

The proposed approach is shown to be energy-efficient as compared to the state-of-the-art technologies like FFO-EVM and ACO. This is accomplished by optimizing VM placement, efficient host consolidation, and fewer migrations. The proposed framework significantly reduces computational overhead and energy consumption in cloud environments by reducing the number of hosts that need to be active and migrating, and unnecessary migrations.

The experimental analysis demonstrates that the proposed adaptive VM migration framework offers substantial benefits in resource usage, migration efficiency and energy saving. Based on the results obtained, it is clear that the proposed approach is better from the performance and scalability aspects than compared to the existing VM migration techniques.

5. CONCLUSION

This paper proposed an adaptive and efficient framework for the virtual machine migration in cloud computing environment. Virtual machines are allocated to several physical machines in cloud data centers to run different user apps and calculations. But, VM allocation problems and the increased workloads can cause VM overload, elevated energy usage, migration cost, and poor system performance. A resource-aware and threshold-based VM migration framework was proposed to solve these problems. First, the model is proposed to implement intelligent VM allocation, taking into account resource availability, communication delay and workload conditions. After that a threshold-based migration mechanism is used to detect overloaded virtual machines and to migrate these machines to lower utilized physical machines. The framework also introduces migration cost optimization to reduce migration latency and communication overheads and energy consumption. A comprehensive experimental study was carried out with various VM configurations and utilization levels. Comparative performance evaluation showed that the proposed approach needed a smaller number of active hosts, lesser number of VM migrations and lesser energy consumption than the existing VM migration techniques. The results obtained proves the proposed framework can utilize the cloud resources, stabilize cloud system, decrease the computational burden, and meet the SLA requirements well. For future research, the proposed framework can be enhanced by incorporating AI and machine learning algorithms to optimize VM allocation and intelligent workload management in large-scale cloud setups.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY
Data can be provided on genuine request.

REFERENCES

[1] K. Rajammal and M. Chinnadurai, “Dynamic load balancing in cloud computing using predictive graph networks and adaptive neural scheduling,” *Scientific Reports*, vol. 15, no. 1, pp. 1–18, 2025, doi: 10.1038/s41598-025-97494-2.
[2] D. Bodra and S. Khairnar, “Machine learning-based cloud resource allocation algorithms: A comprehensive comparative review,” *Frontiers in Computer Science*, vol. 7, pp. 1–22, 2025, doi: 10.3389/fcomp.2025.1678976.
[3] A. Kushchazli, K. Leonteva, I. Kochetkova, and A. Khakimov, “Evaluating QoS in dynamic virtual machine migration: A multi-class queuing model for edge-cloud systems,” *Journal of Sensor and Actuator Networks*, vol. 14, no. 3, pp. 47–61, 2025, doi: 10.3390/jsan14030047.
[4] K. Kotteswari, R. K. Dhanaraj, B. Balusamy, A. Nayyar, and A. K. Sharma, “EELB: An energy-efficient load balancing model for cloud environment using Markov decision process,” *Computing*, vol. 107, no. 81, pp. 1–20, 2025, doi: 10.1007/s00607-025-01439-6.
[5] Z. Jianping, “Optimal resource allocation and load balancing in SDN-based cloud computing: A comprehensive review and heuristic optimization approach,” *SSRN Electronic Journal*, pp. 1–8, 2025, doi: 10.2139/ssrn.5336093.
[6] A. I. Maiyza, H. A. Hassan, W. M. Sheta, K. Banawan, and N. O. Korany, “VTGAN based proactive VM consolidation in cloud data centers using value and trend approaches,” *Scientific Reports*, vol. 15, no. 1, pp. 20133–20148, 2025, doi: 10.1038/s41598-025-04757-z.
[7] S. Sakib, A. Katangur, and R. Dubey, “A dynamic approach to load balancing in cloud infrastructure: Enhancing energy efficiency and resource utilization,” *arXiv preprint arXiv:2508.05821*, pp. 1–12, 2025, doi: 10.48550/arXiv.2508.05821.
[8] C. Liu, L. Ma, M. Zhang, and H. Long, “Optimizing cloud resource management with an IoT-enabled optimized virtual machine migration scheme for improved efficiency,” *Journal of Network and Computer Applications*, vol. 237, p. 104137, 2025, doi: 10.1016/j.jnca.2025.104137.
[9] G. Sripathi and D. A. Khan, “Proactive resource allocation framework with SVM regression for cloud environments,” *SN Computer Science*, vol. 7, no. 1, p. 90, 2026, doi: 10.1007/s42979-025-04541-w.
[10] N. K. Sharma, S. Bojjagani, R. Uyyala, A. K. Maurya, and S. Kumari, “Deep learning BiLSTM and Branch-and-Bound based multi-objective virtual machine allocation and migration with profit, energy, and SLA constraints,” *Sustainable Computing: Informatics and Systems*, vol. 46, p. 101128, 2025, doi: 10.1016/j.suscom.2025.101128.
[11] Y. Y. Raghav and V. Vyas, “Cloud load balancing: Performance comparison between Bird Swarm and Ant Colony Optimization algorithms,” *SN Computer Science*, vol. 7, no. 1, p. 107, 2026, doi: 10.1007/s42979-025-04712-9.
[12] S. H. H. Madni, M. S. A. Latiff, and Y. Coulibaly, “Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities,” *Journal of Network and Computer Applications*, vol. 68, pp. 173–200, 2016, doi: 10.1016/j.jnca.2016.04.016.
[13] M. Xu, W. Tian, and R. Buyya, “A survey on load balancing algorithms for virtual machines placement in cloud computing,” *Concurrency and Computation: Practice and Experience*, vol. 29, no. 12, pp. 1–24, 2017.
[14] N. Doumbouya, M. Hifi, and T. Saadi, “Parallel multi-objective optimization for predictive virtual machine placement and scheduling in cloud computing,” *Future Generation Computer Systems*, vol. 165, pp. 210–228, 2026.
[15] Y. Han *et al.*, “Using virtual machine allocation policies to defend against co-resident attacks in cloud computing,” *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 1, pp. 95–108, Jan.–Feb. 2017.
[16] S. Xu *et al.*, “Energy-efficient dynamic consolidation of virtual machines in big data centers,” in *Proc. Int. Conf. Green, Pervasive, and Cloud Computing*, Cham, Switzerland: Springer, 2017, pp. 211–223.
[17] A. Khosravi, A. N. Toosi, and R. Buyya, “Online virtual machine migration for renewable energy usage maximization in geographically distributed cloud data centers,” *Concurrency and Computation: Practice and Experience*, vol. 29, no. 24, pp. 1–19, 2017.

[18] M. Duggan *et al.*, “A network aware approach for the scheduling of virtual machine migration during peak loads,” *Cluster Computing*, vol. 20, no. 3, pp. 2089–2101, 2017.
[19] T.-Y. Wu, N. Guizani, and J.-S. Huang, “Live migration improvements by related dirty memory prediction in cloud computing,” *Journal of Network and Computer Applications*, vol. 90, pp. 83–89, 2017.
[20] S. Sohrabi *et al.*, “Energy-efficient adaptive virtual machine migration mechanism for private clouds,” *Concurrency and Computation: Practice and Experience*, vol. 29, no. 18, pp. 1–17, 2017.
[21] F. Teng *et al.*, “Energy efficiency of VM consolidation in IaaS clouds,” *The Journal of Supercomputing*, vol. 73, no. 2, pp. 782–809, 2017.
[22] E. Zola and A. J. Kassler, “Optimising for energy or robustness? Trade-offs for VM consolidation in virtualized datacenters under uncertainty,” *Optimization Letters*, vol. 11, no. 8, pp. 1571–1592, 2017.
[23] M. Abdullah *et al.*, “A heuristic-based approach for dynamic VMs consolidation in cloud data centers,” *Arabian Journal for Science and Engineering*, vol. 42, no. 8, pp. 3525–3539, 2017.
[24] N. J. Kansal and I. Chana, “Energy-aware virtual machine migration for cloud computing—A firefly optimization approach,” *Journal of Grid Computing*, vol. 14, no. 2, pp. 327–345, 2016.