IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received 8 September 2023, accepted 26 September 2023, date of publication 29 September 2023, date of current version 5 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3320729

RESEARCH ARTICLE

A Novel Energy Efficient Multi-Dimensional Virtual Machines Allocation and Migration at the Cloud Data Center

NEERAJ KUMAR SHARMA[1], SRIRAMULU BOJJAGANI[1], (Senior Member, IEEE), Y. C. A. PADMANABHA REDDY[2], MANOJKUMAR VIVEKANANDAN[1], JAGADEESAN SRINIVASAN[3], AND ANUP KUMAR MAURYA[4]

[1]Department of Computer Science and Engineering, School of Engineering and Applied Sciences (SEAS), SRM University AP, Amaravati, Andhra Pradesh 522240, India
[2]Department of Computer Science and Engineering, B V Raju Institute of Technology, Narsapur, Telangana 502313, India
[3]School of Computer Science Engineering and Information Systems, Department of Software and Systems Engineering, Vellore Institute of Technology, Vellore, Tamilnadu 632014, India
[4]Goa Institute of Management, Goa 403505, India

Corresponding author: Neeraj Kumar Sharma (neeraj16ks@gmail.com)

ABSTRACT Due to the rapid utilization of cloud services, the energy consumption of cloud data centres is increasing dramatically. These cloud services are provided by Virtual Machines (VMs) through the cloud data center. Therefore, energy-aware VMs allocation and migration are essential tasks in the cloud environment. This paper proposes a Branch-and-Price based energy-efficient VMs allocation algorithm and a Multi-Dimensional Virtual Machine Migration (MDVMM) algorithm at the cloud data center. The Branch-and-Price based VMs allocation algorithm reduces energy consumption and wastage of resources by selecting the optimal number of energy-efficient PMs at the cloud data center. The proposed MDVMM algorithm saves energy consumption and avoids the Service Level Agreement (SLA) violation by performing an optimal number of VMs migrations. The experimental results demonstrate that our proposed Branch-and-Price based VMs allocation with VMs migration algorithms saves more than 31% energy consumption and improves 21.7% average resource utilization over existing state-of-the-art techniques with a 95% confidence interval. The performance of the proposed approaches outperforms in terms of SLA violation, VMs migration, and Energy SLA Violation (ESV) combined metrics over existing state-of-the-art VMs allocation and migration algorithms.

INDEX TERMS Cloud computing, data center, virtual machine, physical machine, energy-aware, branch-and-price, SLA, ESV.

I. INTRODUCTION

To provide different types of cloud services, such as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS), cloud data center's consist of thousands of interconnected Physical Machines (PMs). Therefore, nowadays, geographically distributed data centers consume more energy. A cloud services provider provides different services to customers on a pay-as-you-go model. The pay-as-you-go model says that customers have to pay the same amount for resources/services utilized by the customer [1]. A study on energy consumption of worldwide geographically distributed data centers highlighted that from 2010 to 2020, US-based data centers' energy consumption and worldwide data centers' energy consumption increased by 36% and 56% respectively [2], [3]. Another study demonstrated that worldwide data center's energy consumption is expected to increase in the coming years. Hence, the worldwide cloud data center's energy consumption bill will rise to 78 billion dollars by 2025

The associate editor coordinating the review of this manuscript and approving it for publication was Somchart Fugkeaw.

[4]. Thus, the increased energy consumption will cause several consequences, such as high operational cost, high heat density, reduced reliability, degraded service performances, and environmental pollution [5], [6].

Therefore, the cloud data centre must reduce energy consumption to resolve the highlighted issues. Several approaches, such as energy-efficient VMs allocation and migration, energy-efficient task scheduling, green energy utilization, etc., can reduce the energy consumption of the cloud data center. In this paper, we are only dealing with energy-efficient VMs allocation and migration policies. The multi-dimensional VMs allocation problem is called an variable size bin packing problem [7], where VMs are the items and PMs are the bins. Since the bin packing problem is NP-hard/NP-complete, the complexity of the multi-dimensional VMs allocation problem is also defined in non-polynomial time [8]. Therefore, an approximate or fast exact algorithm is required to solve the VM allocation problem. The key limitation of an approximate algorithm is that it does not guarantee an optimal solution to the given problem. Thus, this paper focuses on the exact algorithm and proposes a Branch-and-Price-based VM allocation algorithm. On the other hand, in the VM migration problem, the migration of VMs from underutilized/over-utilized PM to energy-efficient PM is required. The migration of VMs from one PM to another PM is performed in such a manner to reduce energy consumption and SLA violation at the cloud data center. Hence, both VM allocation and migration problems are critical and challenging. There are many approaches for VM allocation and migration, such as approximation algorithms [9], bio-inspired algorithms [10], [11], and greedy approaches [12]. However, the critical limitations of the existing algorithms are based on a heuristic system and do not ensure the optimal solution to the multi-objective VM allocation problem.

In short, the core idea of our proposed work is described as follows:

- A Linear Programming (LP) based mathematical model of the proposed multi-objective VMs allocation problem is designed.
- Branch-and-Price based VMs allocation algorithm is designed at the cloud data center. In the proposed Branch-and-Price based VMs allocation algorithm, a MP in VMs allocation is generated. This MP is converted into a RMP by developing a primary solution to the problem.
- An Linear Programming Relaxation (LPR) is solved for the RMP, and a dual of the restricted MP is generated.

After generating the dual of the RMP, a sub-problem is solved by passing duals. Column generation and power consumption evaluation are performed to calculate the optimal solution to VMs allocation problem. A VMs migration algorithm is designed and developed in the second phase at the cloud data center. The proposed MDVMM algorithm is based on the multiple resources utilization and least square method. The proposed VMs allocation and migration algorithms are applied to the data center. To check the performance of the proposed algorithms in terms of different performance metrics such as energy consumption, resource utilization and SLA violation are calculated at the cloud data center. In short the key contributions of the proposed work in this paper are as follows:

1) To design a mathematical model of a multi-objective VMs allocation problem using Linear Programming.
2) To design and develop a Branch-and-Price based VMs allocation algorithm for reducing power consumption and resource wastage.
3) To design and develop a VMs migration policy for reducing energy consumption and SLA violation.
4) To evaluate the performance of our proposed VMs allocation and migration algorithms with other state-of-the-art algorithms in different performance metrics.

The remainder of this paper is organized as follows. Section II deals with the background and related work. Section III describes the proposed work. Section IV describes the experimental setup and performance evaluation of the proposed VMs allocation and migration algorithms. Finally, the concluding remarks are given in Section V.

II. BACKGROUND AND RELATED WORK

The existing work on energy-efficient resource allocation is described in terms of energy-efficient VMs allocation, task scheduling, and VMs migration policies at the cloud data center. To reduce the energy consumption at the cloud data center, the existing approaches are described as follows:

Masoudi et al. [20] proposed a two-phase energy-efficient load balancing through VM migration using Particle Swarm Optimization (PSO). In the first phase, the author deactivated many PMs, reducing energy consumption at the cloud data center. In the second phase, the author performed load balancing using the PSO. Their proposed approach also considered the Dynamic Voltage Frequency Scaling (DVFS) technique. In experimental work, they saved approximately 10% energy consumption of the cloud data center by using their proposed method over the PSO algorithm.

Mandal et al. [21] proposed an energy-aware VM migration technique at the cloud data center. In their proposed approach, they selected the VM for the live migration in such a manner that reduced the energy consumption and avoided the SLA violation. The proposed VM selection algorithm is known as MECpVmS. The author conducted their experimental work on a cloudsim simulator, and they utilized a real time dataset. Their proposed approach reduced a significant amount of energy with low SLA violation at the cloud data center.

Li et al. [22] proposed a pareto optimal multi-objective VMs consolidation algorithm (MOEA/D) in the cloud environment. They mainly focused on reducing energy consumption and improving resource utilization. In their proposed work, the authors applied a dynamic workload, saving energy consumption and resource wastage at the cloud data center. Ullah and Chakir [23] proposed a task

TABLE 1. Summary of related work.

Author	Methodology	Experimental setup and other details	Limitations
Rodrigo A. C. et al. [13]	Proposed a topology aware VMs placement algorithm (TAVMP) at the cloud data center. The proposed algorithm reduced the energy consumption of networking switches and servers at the cloud data center.	**Data center environment**: Network aware cloud data center **Constraints**: Bandwidth **Usages**: Network devices power consumption **Simulator**: Cloudsim	Author in their proposed approach did not consider the QoS parameter during VMs placement. Therefore, high SLA violation at the cloud data center.
D. Borgetto et al. [14]	Designed and developed a multi-objective (QoS, and energy-aware) modular data center architecture for the resources management.	**Data center environment**: Multicloud data centers **Constraints**: Workload at the data center **Usages**: Optimal resources utilization **Simulator**:NEST Simulator	In the proposed multi-objective modular cloud data center architecture, the author did not propose any real-time VMs allocation and migration algorithms for resource allocation and migration at the cloud data center.
V. Ebrahimirad et al. [15]	Proposed two different resource scheduling algorithms such as Virtualized Homogeneous Earliest Start Time (VHEST) and Energy-Aware Scheduling Algorithm (EASA) in multi-cloud data centers environment. The VHEST algorithm minimized the makespan of services provided by the cloud data center. On the other hand, EASA algorithm minimized the makespan of cloud services and maximized the utilization of PMs at the cloud data center.	**Data center environment**: Single cloud data center **Constrains**: Priority based VMs allocation **Usages**: Energy aware scheduling **Simulator**: Designed Virtualized Data Center Simulator (VDCS)	To match the user's requirement the proposed algorithms dealt with resource sharing among multi cloud data centers environment. Author did not consider the resource management within a cloud data center.
N. J. Kansal et al. [16]	Designed an energy-aware VMs migration algorithm at the cloud data center. The proposed algorithm is based on the Firefly algorithm. To reduce the energy consumption they migrated the VMs from highly utilized PMs to least utilized PMs by applying their proposed migration algorithm at the cloud data center.	**Data center environment**: Single cloud data center **Constrains**: Distance of VMs **Usages**: Energy-aware VMs migration **Simulator**: Cloudsim	In their proposed approach they did not consider the SLA violation at the cloud data center. Further, the proposed work does not deal with VMs allocation problems.
S. Ilager et al. [17]	To reduce the local hotspots, the author proposed an Energy and Thermal Aware Scheduling (ETAS) algorithm at the cloud data center. They dynamically consolidated the VMs to PMs for saving energy and thermal temperature at the cloud data center.	**Data center environment**: Single cloud data center with cooling devices **Constraints**: Thermal temperature and energy efficiency **Usages**: Mulitobjective VMs consolidation **Simulator**: Cloudsim	The proposed work dealt with energy and thermal aware VMs allocation algorithm, but commercial cloud data center requires VMs migration over some time. In their proposed work author did not consider the VMs migration policy at the cloud data center.
H. Li et al. [18]	Proposed a Customer Satisfaction Level (CSL) Driven Energy Efficient Resource Scheduling (CDEERS) algorithm at the cloud data center. They reduced the energy consumption and avoided the SLA violation at the cloud data center.	**Data center environment**: Single cloud data center **Constrains**: Customer satisfaction and energy efficiency **Usages**: Multiobjective VMs allocation **Simulator**: Cloudsim	Author in their proposed work did not consider the VMs migration problem at the cloud data center.
A. A. Khan [19]	Proposed an energy efficient SLA aware VMs migration algorithm at the containerized cloud data center environment. They reduced the energy consumption and SLA violation at the cloud data center.	**Data center environment**: Single cloud data center **Constrains**: Energy and performance aware VMs allocation **Usages**: Multiobjective VMs allocation **Simulator**: Cloudsim	The proposed work not deals with VMs allocation problem at the cloud data center.

distribution approach using a load balancing technique. The author applied a modified BAT algorithm for load balancing in their proposed method. They performed two changes in the BAT algorithm:

(1) modifying the fitness function during load balancing and (2) modifying the search process in the BAT dimension section. The proposed algorithm improved accuracy and efficiency regarding task distribution at the cloud data center.

Memari et al. [24] proposed a meta-heuristic improved TABU search algorithm for VMs allocation. They improved the performance of the TABU search algorithm by applying ANN and fruit fly optimization (FOA) algorithms. In their experiment, they optimized different factors such as execution time, latency, and memory allocation compared to other Bio-inspired algorithms. Further, They also performed latency-aware task scheduling with their proposed approach.

Li et al. [25] proposed an energy-efficient task scheduling optimizer for the cloud data center. The author considered the data centre's energy efficiency in their proposed approach. Further, to reduce the energy consumption at the cloud data center Polverini et al. [26] designed and developed an online multi-objective batch job scheduling algorithm (GreFar) in a geographically distributed data center environment. In their proposed approach, authors reduced multiple cloud data centres' queuing delay, energy cost, and thermal temperature. Further, by utilizing the offered GreFar jobs scheduling algorithm, they processed the jobs when the queue length was sufficiently large or electricity prices were adequately low.

In their proposed work, Huang et al. [27] collected VMs from different users and allocated the VMs on PMs based on their throughput requirement and current CPU utilization. They also proposed a VMs migration policy by setting the utilization threshold of a PM. In their proposed approach, the author reduced energy consumption and avoided the SLA violation at the cloud data center.

Lv et al. [28] proposed a multi-cloud broker model called the CoMCLOUD. In their proposed approach for VMs allocation, they applied Ant Colony Optimization (ACO) at the cloud data center. Further, they also designed and developed a multi-broker model to handle the cost and

Quality of Service (QoS) trade-off. They also proposed a minimum disruption based VMs migration policy. The proposed VMs allocation and migration policies reduced the cost and SLA violation and the minimum disruption time of a VM at the cloud data center. The proposed approach is helpful for live VM migration. Cerroni and Esposito [29] proposed a Linux-based live VMs migration policy; in their proposed approach, they reduced the interruption time of the VMs during the live migration. They also presented a geometric programming model for VMs migration; the key objectives of the proposed geometric programming model are to optimize the bit rate allocation and reduce the live migration time of the VMs. Anwar et al. [30] designed a VM migration approach for solving the VM migration timing problem. Their proposed method used game theory concepts to solve the VM migration timing problem. They formulate the VM migration problem for capturing the data leakage model. The proposed model is helpful to pay off when the cloud utilizes intrusion detection systems that detect side-channel attacks.

Nikzad et al. [31] proposed a cloud resource management technique by applying multi-objective VMs allocation in a dynamic cloud environment. The author considered eight criteria to reduce energy consumption and SLA violation in their proposed approach. Further, they solved the same multi-objective problem by applying heuristics and meta-heuristic algorithms. In their proposed work, they saved energy consumption up to 12.5% compared to other algorithms. They also reduced the SLA violation and number of VMs migrations at the cloud data center.

Li et al. [32] proposed a Modified Particle Swarm Optimization (MPSO) based energy efficient VMs migration and consolidation algorithm at the cloud data center. Their proposed work migrated the VMs based on double threshold values and consolidated the maximum number of VMs to fewer PMs at the cloud data center. Their scheme reduced energy consumption and avoided the SLA violation at the cloud data center. Li et al. [33] designed and developed an Energy Efficient and QoS-aware (EEQoS) VMs consolidation algorithm at the cloud data center. In their proposed work, first, they defined the objective function in terms of energy and QoS for the cloud data center. Further, they utilized the PSO algorithm for VMs consolidation to fewer PMs. Their experiment reduced the response time, SLA violation, and energy consumption by 27.2%, 31.4%, and 3.8%, respectively. Further, Table 1 summarizes the related work w.r.t. VM allocation and migration at the cloud data center.

III. PROPOSED WORK

Two different algorithms, such as Branch-and-Price based VMs allocation and multi-resource utilization based VMs migration algorithms, are proposed in this paper. Since the VMs allocation problem is a combinatorial multi-dimensional optimization problem, it looks like a multi-dimensional bin packing problem. Hence, we applied a Branch-and-Price method to solve the multi-dimensional combinatorial optimization problem. Further, the proposed VMs migration problem is a pareto optimal multi-objective problem in which we not only reduce the energy consumption but also need to avoid the SLA violation at the cloud data center. Fig. 1 shows the block diagram of the proposed work concerning multi-objective VMs allocation and migration problems at the cloud data center. The block diagram includes a broker, central unit, and node controller. The user requests the resources in terms of VMs from the broker. A broker collects all the user's requested VMs and forwards them to the cloud services provider. The cloud services provider provides the cloud services by the data center. After getting the VMs allocation requests from the broker, the proposed algorithm executes itself and allocates VMs to PMs at the cloud data center. Hence, the central maintenance unit is responsible for assigning and migrating the VMs on the PMs at the cloud data center. The proposed VMs allocation and migration algorithms are deployed in the centre maintenance unit. One node controller unit is deployed on each PM at the cloud data center. The node controller unit is responsible for keeping the resource utilization record of a PM, and it is also responsible for allocating the resources among the VMs. Hence, a mathematical model of the VM allocation problem is required before designing the VM allocation and migration algorithms.

A. VMS ALLOCATION PROBLEM FORMULATION

The LP based VMs allocation problem formulation at the cloud data center is described as follows. Let us consider a data center consisting of the *'m'* number of PMs and *'n'* number of VMs requested by the users at time instance *'t'*. The different resources of a PM (pm_j) ($j \in \{1, 2, \ldots m\}$), such as MIPS, RAM, Storage, and Processing elements, are described by pm_j^{mips}, pm_j^{ram}, $pm_j^{storage}$, pm_j^{pe} respectively. Similarly, the dimensions (mips, ram, storage, and processing element) of a VM (vm_i) where $i \in \{1, 2, \ldots n\}$) are defined as vm_i^{mips}, vm_i^{ram}, $vm_i^{storage}$, vm_i^{pe} respectively. We aim to allocate the *'n'* number of VMs to a minimum number of energy-efficient PMs at the cloud data center. Hence, an objective function of a VMs allocation problem is described as follows:

$$min \quad f = \sum_{j=1}^{m} P_j x_j \tag{1}$$

where P_j represents the power consumption of a PM (pm_j); x_j represents a binary variable, its value is 1 when at least one VM is allocated to pm_j; otherwise, it is 0. The power consumption of a pm_j in Eq. 1 is described as follows:

$$P_j = (P_j^{max} - P_j^{min})u + P_j^{min} \tag{2}$$

where (u_t) represents the CPU utilization of a PM at time instance 't'; The maximum and minimum power consumption of the PM is described by (P_j^{max}) and (P_j^{min}) respectively.

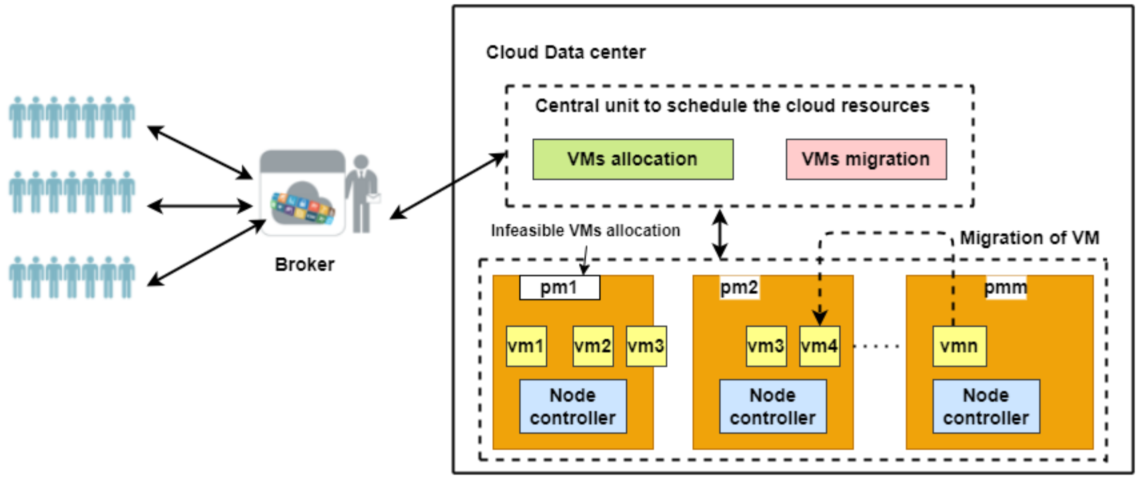

FIGURE 1. **Block diagram of proposed work.**

subject to the following constraints

$$\sum_{j=1}^{m} x_{ij} = 1, \quad \forall i \in \{1, 2, \ldots .n\} \tag{3}$$

$$\sum_{i=1}^{n} vm_i^{mips} x_{ij} \leq pm_j^{mips}, \quad \forall j \in \{1, 2, \ldots .m\} \tag{4}$$

$$\sum_{i=1}^{n} vm_i^{ram} x_{ij} \leq pm_j^{ram}, \quad \forall j \in \{1, 2, \ldots .m\} \tag{5}$$

$$\sum_{i=1}^{n} vm_i^{storage} x_{ij} \leq pm_j^{storage}, \quad \forall j \in \{1, 2, \ldots .m\} \tag{6}$$

$$vm_i^{pe} x_{ij} \leq pm_j^{pe}, \quad \forall j \in \{1, 2, \ldots .m\} \tag{7}$$

Eq. 1 describes the objective function of the VMs allocation problem in terms of minimizing the power consumption of the cloud data center by utilizing a minimum number of energy-efficient PMs; Eq. 2 describes the power consumption of a PM; Eq. 3 ensures that each VM is assigned to only one PM; Eqs. 4 to 7 describe the resources capacity constraints such as mips, ram, storage, and pe, respectively (The sum of a particular resource of all the VMs is lesser than the capacity of a PM on which these VMs are allocated); x_{ij} describes a binary variable; its values is one if vm_i is assigned to pm_j otherwise it values is '0'.

Let us consider a possible set of solutions for VMs allocation in the case of pm_j is represented by $k_j = \{x_1^j, x_2^j, \ldots x_k^j\}$. A feasible solution such as x_k^j satisfies all the constraints defined in Eq. 4 to Eq. 7. Out of 'k' feasible solutions of pm_j, we can select only one possible solution. In Fig. 1 allocation of $vm1$, $vm2$, and $vm3$ on $pM1$ represents the infeasible VMs allocation because this allocation not satisfy the constraint defined in Eq. 4 to Eq. 7. The selection of a feasible solution for pm_j is described by a binary variable y_k^j.

y_k^j=1 *if feasible allocation* x_k^j *is selected;otherwise* y_k^j=0.

Hence, based on feasible solutions, the Generalized Assignment Problem (GAP) is formulated as follows:

$$min \quad f' = \sum_{j=1}^{m}\sum_{k=1}^{k_i}(\sum_{i=1}^{n} P_j x_{ij}) y_k^j \tag{8}$$

Subject to the following constraints:

$$\sum_{j=1}^{m}\sum_{k=1}^{k_j} x_k^j y_k^j = 1, \quad \forall i \in \{1, 2, \ldots n\} \tag{9}$$

$$\sum_{k=1}^{k_j} y_k^j \leq 1, \quad \forall j \in \{1, 2, \ldots m\} \tag{10}$$

$$y_k^j \in \{0, 1\}, \quad j \in \{1, 2, ..m\} \quad k \in k_j \tag{11}$$

where Eq. 8 represents the new objective function of a GAP; Eq. 9 ensures that each VM is assigned to only one PM in GAP; Eq. 10 provides that a maximum one feasible solution is selected for a pm_j*; Eq.11 describes that a boolean variable* (y_k^j) *can take a binary value.*

B. PROPOSED BRANCH-AND-PRICE ALGORITHM

The Branch-and-Price is a combinatorial optimization technique for solving Integer Linear Programming (ILP) and Mixed Integer Linear Programming (MILP) problems with many variables. This technique is similar to a Branch-and-Cut technique. In the case of the Branch-and-Cut technique, we mainly focus on row generation, whereas, in the case of Branch-and-Price, we focus on column generation. The Branch-and-Cut is defined as Branch-and-Bound with cutting planes such as (Branch-and-Cut = Branch-and-Bound + Cutting Planes); it is called row generation. Therefore, the Branch-and-Price algorithm combines Branch-and-Bound with column generation such as (Branch-and-Price=Branch-and-Bound+column generation). A search tree is constructed to calculate the solution to the VMs allocation problem in the

Branch-and-Price technique. A column may be generated at each search tree node to improve the LP relaxation. In the case of the Branch-and-Price technique, a set of columns is left out as an LP relaxation for solving a large ILP problem. Thus, using the Branch-and-Price technique, we can handle many variables in the ILP problem.

Since the optimal solution to an original VMs allocation problem consists of many zero columns, a sub-problem is solved to calculate the optimal solution of the given situation. A branch is generated when the solution to the problem does not satisfy the defined integrity constraints. A flowchart of the proposed Branch-and-Price algorithm is described in Fig. 2. First, we formulate the original VMs allocation problem as an LP problem. Solving it takes work since the original LP problem consists of many variables and constants. Hence, there is a need to convert an actual LP problem into two problems: a GAP problem and a subproblem. The GAP problem is known as the MP. The main advantage of reformulating the original problem into a MP is that we can apply Dantzig-Wolfe decomposition to the MP. Since the MP still consists of many variables, the MP is converted to the Restricted Master Problem (RMP).

The RMP consists of a lesser number of variables. Therefore, an LPR is solved for the same RMP. The solution of the LPR gives the dual of the RMP. This dual is passed to solve the sub-problem. After solving the subproblem, there is a need to verify the different columns generated in the solution. If the solution is integral, accept the solution that consumes less power; otherwise, we need to create a new branch and start the process from the LPR, as mentioned in Fig. 3. Further, we can generate the additional columns for the RMP in terms of VMs allocation problem by solving the following two subproblems:

$$max(\quad 1 \leq j \leq m\{z(KP_j) - v_j\}) \tag{12}$$

where v_i describes the dual associated with convexity constraints of pm_j; $z(KP_i)$ represents the optimal solution to the RMP knapsack problem.

$$\sum_{i=1}^{n} vm_{ij}^{mips} x_k^j \leq pm_j^{mips}, \quad \forall j \in \{1, 2, \ldots m\} \tag{13}$$

$$\sum_{i=1}^{n} vm_{ij}^{ram} x_k^j \leq pm_j^{ram}, \quad \forall j \in \{1, 2, \ldots m\} \tag{14}$$

$$\sum_{i=1}^{n} vm_{ij}^{storage} x_k^j \leq pm_j^{storage}, \quad \forall j \in \{1, 2, \ldots m\} \tag{15}$$

$$\sum_{i=1}^{n} vm_{ij}^{pe} x_k^j \leq pm_j^{pe}, \quad \forall j \in \{1, 2, \ldots m\} \tag{16}$$

From Eq. 13 to Eq. 16, describe the resource constraints such as MIPS, RAM, Storage, and Pe for the RMP problem.

Hence, our objective is to search for the optimal allocation of VMs to PMs at the cloud data centre, done by the first sub-problem.

$$max(1 \leq j \leq mz(KP_j) - v_j) \tag{17}$$

The sub-problem generates the column with a highly reduced cost. The RMP keeps growing during the column generation process. It is optional to solve the sub-problem optimally and create the column with reduced power consumption. If the objective function value is less than or equal to '0,' then the current optimal solution of the RMP is also optimal for the MP.

1) BRANCHING STRATEGIES

If a solution to the MP is fractional, then the solution for a common problem is also fractional. Each branch in the search tree of standard formulation of the problem corresponds to an equivalent branch in the Master's Problem. In the typical VMs allocation problem, by fixing x_{ij} to zero, we can prohibit vm_i from being assigned to pm_j. In the case of the MP, this condition can be achieved by the following description: if x_k^j=1, then y_k^j=0 for all $k \in K_i$. Conversely, in case we want vm_i to be assigned to pm_j, if x_k^j=0, then y_k^j=0 for all k$\in$ K_i and for some L$\neq$ i if x_k^j=1, then y_k^l=0 for 1$\leq$L$\neq$i$\leq$m and k $\in K_i$.

2) EXAMPLE

An example describes the overall description of the proposed Branch-and-Price based VMs allocation problem. Let us consider the Generalized Assignment Problem (GAP) consists of m=2 PMs and n=3 VMs, and power consumption associated with allocation of vm_i to pm_j is described by P_{ij}. Further, k_j represents the feasible allocation of VMs to PMs at the cloud data center. Thus, k_1=(1,0,0,), (0,1,0), (0,0,1), (1,0,1), and k_2=(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (0, 1, 1). Hence, the Master Problem (MP) is formulated as follows:

$$\begin{aligned} max \quad z = 5y_1^1 + 7y_2^1 + 3y_3^1 + 8y_4^1 + 2y_1^2 \\ + 10y_2^2 + 5y_2^3 + 12y_2^4 + 15y_2^5 \end{aligned} \tag{18}$$

Subject to the following constraints:

$$y_1^1 + 0 + 0 + y_4^1 + y_1^2 + 0 + 0 + y_2^4 + 0 = 1 \quad (u_1) \tag{19}$$

$$0 + y_1^2 + 0 + 0 + 0 + y_2^2 + 0 + y_2^4 + y_2^5 = 1 \quad (u_2) \tag{20}$$

$$0 + 0 + y_1^3 + y_4^1 + 0 + 0y_2^3 + 0 + y_2^5 = 1 \quad (u_3) \tag{21}$$

$$y_1^1 + y_1^2 + y_1^3 + y_4^1 \leq 1 \quad (v_1) \tag{22}$$

$$y_1^1 + y_2^2 + yy_2^3 + y + 2^4 + y_2^5 \leq 1 \quad (v_2) \tag{23}$$

where u_i, and v_j represent the duals associated with vm_i and pm_j respectively.

Let us randomly choose the columns y_1^1 and y_2^5 which are feasible solutions. Hence, the Restricted Master's Problem (RMP) is described as follows:

$$max \quad z = 5y_1^1 + 15y_2^5 \tag{24}$$

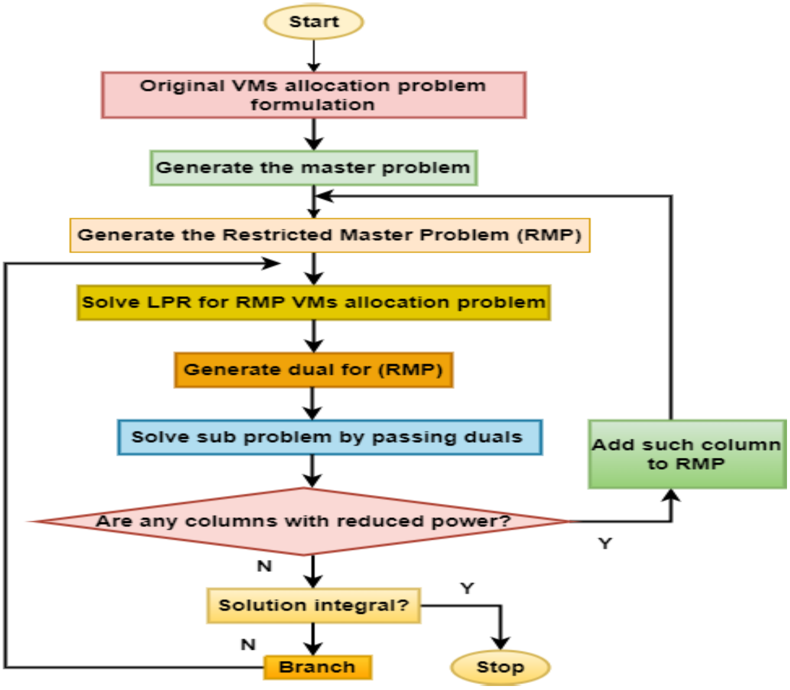

FIGURE 2. Flow chart of proposed Branch-and-Price algorithm.

Subject to Constraints

$$y_1^1 + 0 = 1 \quad (u_1) \tag{25}$$

$$0 + y_5^2 = 1 \quad (u_2) \tag{26}$$

$$0 + y_5^2 = 1 \tag{27}$$

$$y_1^1 + 0 = 1 \tag{28}$$

$$0 + y_5^2 = 1 \tag{29}$$

The last three constraints are redundant. Therefore their corresponding duals are zero.

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{30}$$

Therefore

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{31}$$

$$c_B B^{-1} = \begin{pmatrix} 5 & 15 & 5 \end{pmatrix} \tag{32}$$

Hence, u_1=5, u_2=15, u_3=5, v_1=0, v_2=0. Thus, the sub problem for pm_1 is described as follows:

$$max(5-5)x_1^1 + (7-15)x_1^2 + (3-5)x_3^1 \tag{33}$$

Subject to the following constraints

$$3x_1^1 + 4x_1^2 + 2x_3^1 \leq 5 \tag{34}$$

The optimal solutions are (1,0,0) and (0,0,0), each with z=0. Hence, z(KP_1)-v_1=0-0=0

Sub problem for pm_2

$$max(2-5)x_1^2 + (10-15)x_2^2 + (5-5)x_3^2 \tag{35}$$

Subject to the following constraints

$$5x_1^2 + 3_x2^2 + 4x_2^3 \leq 8 \tag{36}$$

The optimal solutions are (0,0,1) and (0,0,0), each with z=0. Hence, z(KP_2)-v_2=0-0=0. Thus the reduced cost for all the columns is 0. Therefore y_1^1=1, y_5^1=1, y_3^2=0 is the optimal solution. Hence, the optimal VM allocations are: (1,0,0)for pm_1, such that vm_1 is allocated to pm_1.

(1,1,0) for pm_2, such that vm_1, and vm_2 are allocated to pm_2. Thus, the optimal solution is $z_{optimal}$=5+10+5=20.

C. PROPOSED MDVMM TECHNIQUE

Our proposed online MDVMM technique reduces energy consumption and avoids SLA violations at the cloud data center. The proposed MDVMM is an online VMs migration technique therefore, to meet multiple objectives, such as reducing energy consumption and SLA violation, there is a need to select the underutilized and overutilized PMs at the cloud data center. Migrating VMs from underutilized PM to energy-efficient PM will reduce energy consumption, and migrating VMs from overutilized PM to energy-efficient PM will reduce the SLA violation.

The proposed MDVMM technique consists of four steps: (1) Detecting the overutilized and underutilized PMs in the cloud data center. (2) Selecting VMs from the overutilized PMs for the migration. (3) Selecting new PMs where VMs to be migrated. (4) Implementing the proposed migration technique to all the overutilized and underutilized PMs at the cloud data center.

To detect the overutilized and underutilized PMs at the cloud data centre, there is a need to set an upper and lower threshold utilization at the cloud data centre. A PM's upper and lower threshold utilization can be set by two approaches, i.e., static and dynamic. Since the cloud data centre works in a dynamic environment where the workload of a PM changes continuously (unpredictable), and our proposed technique is an online VM migration technique. Therefore, instead of taking a static threshold utilization of a PM, we need to calculate the dynamic threshold utilization of a PM at the cloud data center. Therefore, in our proposed MDVMM technique, we place a dynamic threshold of a PM. The calculation of a dynamic threshold of a PM is based on the statistical analysis of the historical utilization of a PM.

Further, the aggressive VMs migration (consolidating more VMs to fewer PMs) will increase the SLA violation, and fewer VMs consolidated to more PMs will increase resource wastage and energy consumption at the cloud data center. Therefore, we performed the trade-off between SLA violation and energy consumption during our proposed MDVMM algorithm.

To ensure SLA violation and energy efficiency in the proposed MDVMM technique, we set a PM's upper and lower threshold utilization defined in [34]. The upper and lower threshold utilization in [34] is based on the Inter Quartile Range (IQR) technique. The IQR technique is applied to the historical data (previous utilization of a PM in terms of CPU, RAM, and Storage), and a whisker plot of a PM is generated in the proposed approach. The developed whisker plot of a PM gives different information such as min value, IQR, max value, and outliers. Hence, in our proposed approach, we set a PM's lower and upper threshold utilization based on the IQR range.

Let us consider w_1, w_2, and w_3 are the weight of the CPU, RAM, and Storage of a PM in the cloud data center, i.e. ($w_1 + w_2 + w_3 = 1$). The CPU, RAM, and Storage utilization of a PM are described as pm^{cpu}, pm^{ram}, and $pm^{Storage}$, respectively. Thus, the load (L) of a PM in terms of Multi Dimensional Load (MDL) is defined as follows:

$$L = \frac{w_1}{1 - pm^{cpu}} \times \frac{w_2}{1 - pm^{ram}} \times \frac{w_3}{1 - pm^{Storage}} \tag{37}$$

where $0 \le pm_j^{mips}, pm_j^{ram}, pm_j^{storage} < 1$

The core idea of the proposed MDVMM technique is that in a given time series data, if there are 'w' MDL values (L) more than as compared to the upper threshold utilization value of 'z' most recently observed values of a PM, where ($w \le z$) then the PM is treated as overutilized PM, on the other hand, if there are 'w' MDL values less than as compared to the lower threshold utilization value of 'z' most recently observed values of a PM then the PM is treated as underutilized PM. In this approach, there are two possible conditions associated with 'w', such as for a given 'z', a small value of 'w' will cause aggressive detecting of the overloaded PMs at the cloud data center, which will lead to more number of VMs migration at the cloud data center. Thus, more SLA violations at the cloud data center. On the other hand, a large value of 'w' will take a lot of time to predict the overloaded PM at the cloud data center. In the case of (w=z=1), the highest aggressive approach is to find the overloaded and underloaded PMs at the cloud data center.

Hence, to improve the accuracy of the proposed MDVMM technique in addition to 'w' out of 'z' values, we need to check a PM's overload and underload conditions on the following predicted value, i.e. (z+1). The expected value of 'L' on the (z+1) time instance is based on the previous 'z' values; hence, if the PM is overloading on 'w' values, it should be overloaded on the (z+1) value. The proposed approach uses a time series prediction technique to predict a PM's load 'L' on (z+1) time instance.

Let us consider a sequence of observed values 'L' of a PM is $L_1, L_2, \ldots\ldots L_z$ to predict the L_{z+1}, we use the least square method to calculate the load L_{z+1} at time instance (z+1). The prediction of L_{z+1} is described as follows:

$$L_{z+1} = a + bX_{z+1} \tag{38}$$

where 'a' and 'b' represents the constants; 'X_{z+1}' represents the time. The normal equation for parameter 'a' is described as:

$$\sum_{i=1}^{z} L_i = za + b\sum_{i=1}^{z} X_i, \quad where \quad i \in \{1, 2, \ldots z\} \tag{39}$$

The normal equation for parameter 'b' is described as:

$$\sum_{i=1}^{z} X_i L_i = a\sum_{i=1}^{z} X_i + b\sum_{i=1}^{z} X_i^2, \quad i \in \{1, 2, \ldots z\} \tag{40}$$

If $\sum_{i=1}^{z} X_i = 0$ then a=$\frac{\sum_{i=1}^{z} L_i}{n}$, and b=$\frac{\sum_{i=1}^{z} X_i L_i}{\sum_{i=1}^{z} X_i^2}$.

Where, $X_i = \frac{x_i - x_{origin}}{t}$, $x_{origin} = (z+1)/2$ if 'z' is odd, otherwise if 'z' is even then $x_{origin} = z/2$; t=Time interval.

1) VMS MIGRATION PERFORMANCE METRIC

After selecting the underutilized and overutilized PMs at the cloud data center, VMs are selected for migration. In the case of VMs migration VMs are reallocated from their parent PM to another PM. The migration of a VM affects both the performance, such as source and destination PMs at the cloud data center. In a cloud data center, one online VM migration from source PM to destination PM takes approximately 10% overhead in terms of PM resources (CPU, RAM, Storage) defined in [34]. This overhead leads to SLA violations and performance degradation at the cloud data center. Based on the existing work [34], the VM migration time (T_i^m) and performance degradation (U_i^d) are described for an online

VM migration as follows:

$$T_i^m = \frac{M_i}{B_i} \tag{41}$$

$$U_i^d = 0.1 \int_{t_0}^{t_0+T_i^m} U_i(t)dt \tag{42}$$

where T_i^m represents the completion time of migration for vm_i; M_i and B_i describe the amount of memory and bandwidth are used by vm_i during migration; U_j^d describes the degradation of performance by vm_i; $U_j(t)$ represents the CPU utilization of vm_i

2) SLA VIOLATION METRIC

SLA violation is an important performance metric for the cloud service provider because high SLA violation leads to QoS degradation. Further, the cloud service provider may lose the customer due to high SLA violations. A cloud SLA violation is defined in different terms, such as increased response time, low bandwidth, high downtime, minimum throughput, etc. The definition of the SLA violation depends on the application type running on the cloud. Hence, the SLA violation metric at the IaaS level is defined in two different parameters:

(1) Percentage of time CPU utilization is at total capacity, i.e. 100% CPU utilization. In case of full CPU utilization, it will not provide the required performance to the VM [34]. A mathematical expression of SLA violation for the PMs in case of full CPU utilization (SLA(f))is defined as follows:

$$SLA(f) = \frac{1}{m} \sum_{j=1}^{m} \frac{T_j^f}{T_j^a} \tag{43}$$

where 'm' describes the number of PM; T_j^f represents the total time span pm_j utilized in full capacity; T_j^a represents the active time span of a pm_j.

(2) The SLA violation due to migration of a vm_i from the pm_j is described as follows:

$$SLA(m) = \frac{1}{n} \sum_{i=1}^{n} \frac{C_i^d}{C_i^t} \tag{44}$$

where 'n' represents the number of VMs to be migrated; C_i^d represents the performance degradation due to vm_i migration; C_i^t represents the amount of resource requested by vm_i.

A combined performance metrics of SLA violation (SLAV) in terms of full PM utilization and degradation of performance due to VM migration is defined as follows:

$$SLAV = SLA(f) \times SLA(m) \tag{45}$$

Further, a combined performance metric (ESV), which depends on energy consumption and SLA violation, is described as follows:

$$ESV = \int_0^t P_j(t)dt \times SLAV \tag{46}$$

3) REALLOCATION OF VMS

After collecting the VMs from underutilized and overutilized PMs, there is a need to migrate these VMs to energy-efficient PMs. In the proposed MDVMM technique, a selection of PMs for reallocating the VMs is performed based on their energy efficiency. A detailed description of the proposed MDVMM technique is described as follows: First, we list VMs to be migrated from the underutilized and overutilized PMs at the cloud data center. After that, we arrange the VMs in decreasing order based on their resource utilization and place the PMs in non-decreasing order based on their resource capacity. After setting the VMs and PMs in reducing and non-decreasing order, we migrate the VMs based on the First Fit technique. Algorithm 1 describes the overload and underload PMs detection, and Algorithm 2 describes the detailed description of the proposed VMs migration policy.

Algorithm 1 Overload and Underload PM Detection

```
Input: List of PMs [], List of VMs []
Output: Overload list (O[]), Underload list (U [])
O[]={}, U[]={}
foreach PM_j in PM[] do
    Compute L_j //Computes total load
    if (L_j ≥ Th_j^u) then
        Add PM_j in O[];
    else
        (L_j ≤ Th_j^l)
        Add PM_j in U[];
    end
end
return O and U
```

D. CONFIDENCE INTERVAL

To check the acceptability of the proposed VMs allocation and migration algorithms, we calculate the confidence interval of the proposed algorithms. In this process, we conducted our experiment multiple times and calculate the different parameters. The calculation of confidence interval is described as follows:

$$CI = \bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{47}$$

'σ'represents the standard deviation; 'n' represents sample size; '$\bar{x}$' represents sample mean;α represents significance value.

IV. EXPERIMENTAL SETUP AND RESULT ANALYSIS

We used a cloudsim simulator to check the performance of our proposed Branch-and-Price VMs allocation and MDVMM VMs migration algorithms. The cloudsim simulator provides different packages such as power and network packages and files such as VM specification, PM specification, etc. Therefore, a vital advantage of the cloudsim simulator is that, we can crate the cloud data

TABLE 2. PM's configuration.

PM Type	Processor	PE (Number of cores)	Memory	Storage
HPE ProLiant DL380 Gen10	Dual Intel Xeon scalable processors	28	3 TB	4 x 1.6 TB
Dell PowerEdge R740	2 x Intel Xeon Gold 6240R, 24 cores each	48	256 GB	4 x 1.6 TB
Cisco UCS C220 M5	Intel Xeon scalable	28	128 GB	77TB

Algorithm 2 VMs Migration

Input: List of Overload PMs (O []), List of Underload PMs (U[])
Output: Destination PM
foreach *PM_j in PM[]* **do**
  Compute $T_c = \left(PM_j^{CPU} \times PM_j^{RAM} \times PM_j^{Storage}\right)$
end
Arrange PMs in increasing order based on T_c
foreach *VM_i in VM[]* **do**
  Compute $T_r = \left(VM_i^{CPU} \times VM_i^{RAM} \times VM_i^{Storage}\right)$
end
Arrange VMs in decreasing order based on T_r
foreach *PM_j in $(O\,[] + (U\,[])$* **do**
  foreach *VM_i in PM_j:* **do**
    if $(VM_i \in O\,[])$ **then**
      Reallocate $VM_{i]}$ using FFD policy
      After reallocating VM_i calculate L_j
      if $\left(L_j < Th^U\right)$ **then**
        break;
      else
        if $(VM_i \in U\,[])$ **then**
          Reallocate $VM_{i]}$ using FFD policy
          After reallocating VM_i calculate L_j
          if $\left(L_j == 0\right)$ **then**
            break;
      end
    else
      break;
    end
  end
end
return list of VMs[] migrated;

center environment as per our requirement and deploy our proposed VMs allocation and migration policies. In cloudsim simulator, we created a heterogeneous cloud data center environment by taking three different types of PMs at the cloud data center [35]. Table 2 describes the resource capacity of all the PMs at the cloud data center. After creating the heterogeneous cloud data center environment in cloudsim simulator, we created five different types of VMs requested by the cloud customers at the cloud data center. To create a number of VMs, we used PlanteLab dataset [36], and Amazon Ec2 VM instances. An overall

TABLE 3. VM's instances.

VM Type	PE (Number of cores)	Memory	Storage
Small VM	2	4 GB	50 GB SSD
Medium VM	4	8 GB	100 GB SSD
Large VM	8	16 GB	200 GB SSD
Extra Large VM	16	32 GB	400 GB SSD
2x Large VM	32	64 GB	800 GB SSD

configuration of different types of VMs is described in Table 3. In the PlanetLab dataset, we randomly selected several VMs requested in 10 days from March 2011 to April 2011. Further, we divided the total number of VMs requested by the customer during a day into five categories: small, medium, large, extra large, and 2 large. Fig. 3 shows the day-wise VMs requested by the cloud customer a the cloud data center. Further, we performed VMs allocation and migration at the cloud data center by implementing our proposed VMs allocation and migration algorithms in cloudsim simulator's VMs allocation and migration classes. To check the superiority of our proposed Branch-and-Price VMs allocation algorithm, we compared our proposed algorithm with state-of-the-art designed four approximation algorithms such as First Fit Decreasing (FFD), Best Fit Decreasing (BFD), Modified Best Fit Decreasing (MBFD) [20], First Fit Decreasing Height (FFDH) [37], and two Bio-inspired algorithms such as Genetic Algorithm (GA), and Energy Efficient Task Scheduling Algorithm (ETSA) [38]. The performance of the proposed VMs allocation with existing algorithms is compared on the basis of different performance metrics such as resource utilization and power consumption. Fig. 4 shows the day-wise percentage of CPU utilization at the cloud data center.

The percentage of CPU utilization is high in the case of our proposed Branch-and-Price VMs allocation algorithm compared to other existing algorithms. The key reason behind the high CPU utilization in the case of Branch-and-Price is that Branch-and-Price selects the optimal number of PMs for the VMs allocation. Therefore, CPU wastage is less in the case of the Branch-and-Price algorithm. Fig. 5 and Fig. 6 show the percentage of RAM and percentage of storage utilization respectively. In the case of both the resources such as RAM and storage the proposed Branch-and-Price VMs allocation algorithm performs better than existing approximation and Bio-inspired algorithms. The key reason behind the high resource utilization in the case of

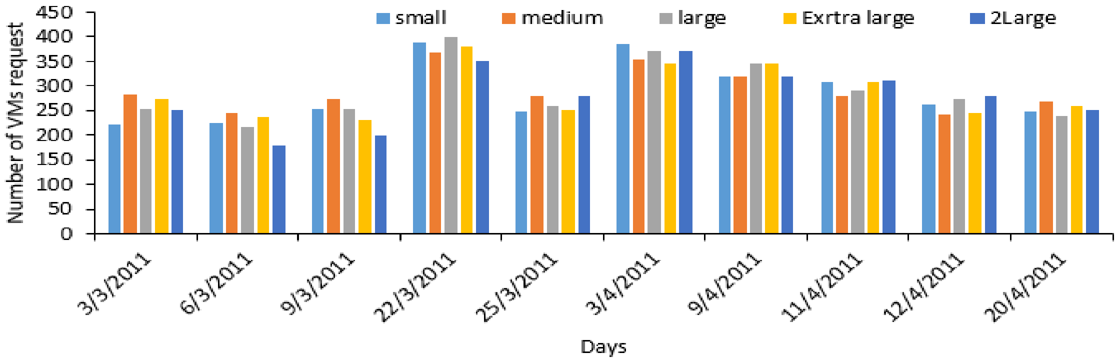

FIGURE 3. VMs requested at the data center.

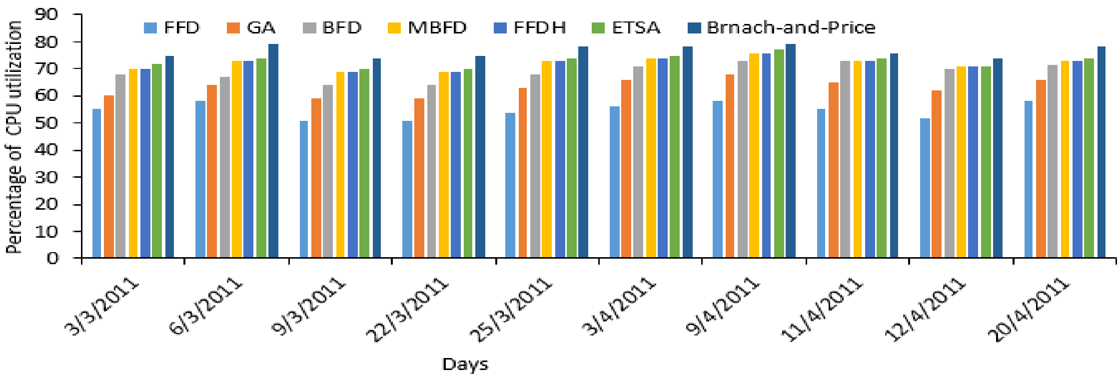

FIGURE 4. % of CPU utilization.

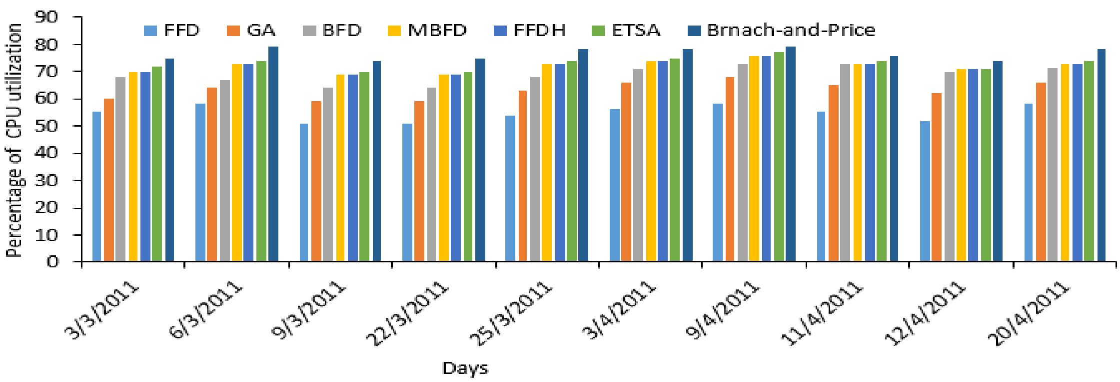

FIGURE 5. % of RAM utilization.

the proposed algorithm is that because the Branch-and-Price algorithm is an exact algorithm therefore, for VMs allocation, it selects the best PM from the datacenter by generating the branch in the search tree. Therefore, the resource wastage is less in the case of the proposed Branch-and-Price VMs allocation algorithm than other algorithms.

Among all the VM allocation algorithms, FFD and GA perform worst because FFD is an approximation algorithm that allocates the VM based on the first fit decreasing; therefore, FFD does not search for the best PM for the VM allocation. In the case of GA, it performs a random search using crossover and mutation operations; therefore, the probability of getting the best PM is low. Further, BFD and MBFD are the approximation algorithms that do not guarantee the optimal solution. The day-wise power consumption of the cloud data centre is shown in Fig. 7. The day-wise power consumption is less in the proposed Branch-and-Price algorithm than other VMs allocation algorithms because Branch-and-Price selects the minimum number of energy-efficient PMs using the search tree at the cloud data centre. The proposed Branch-and-Price VMs allocation algorithm consolidates more VMs into fewer energy-efficient

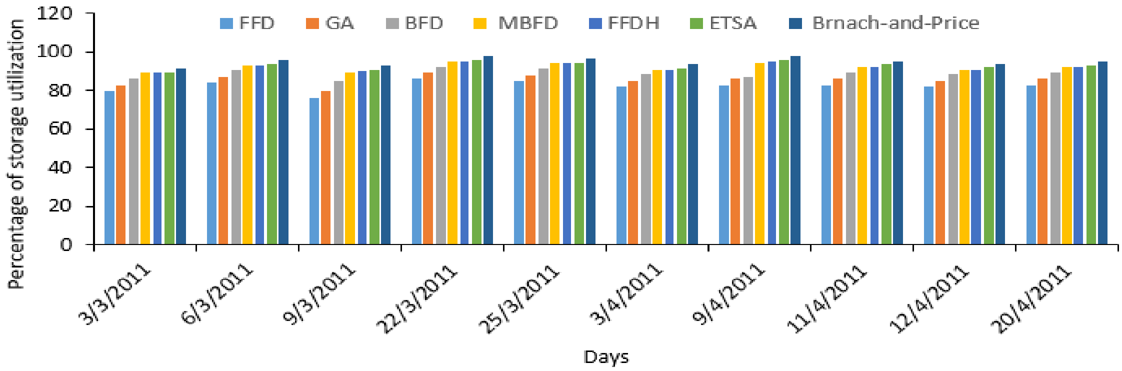

FIGURE 6. **% of storage utilization.**

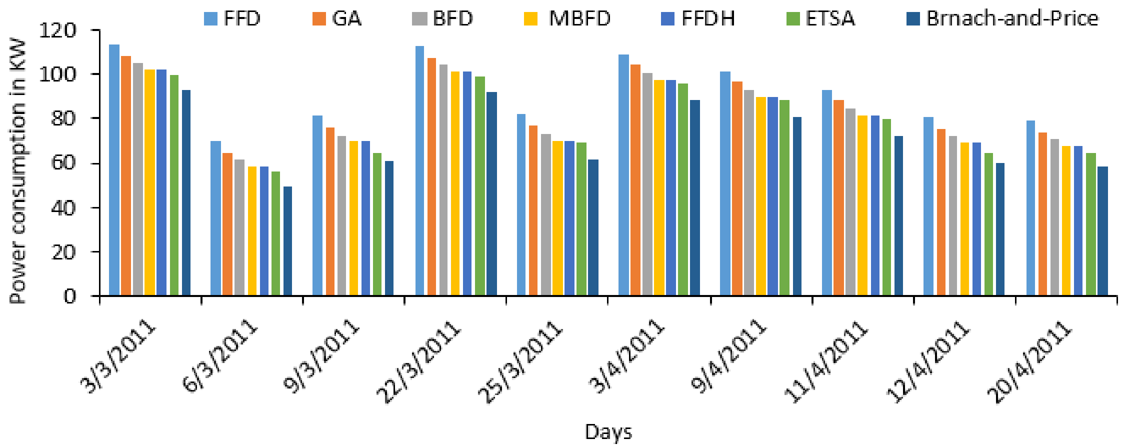

FIGURE 7. **Power consumption at the cloud data center.**

PMs resulting in more number of PMs in switched-off conditions at the cloud data centre. A larger number of switched-off PMs reduced the power consumption at the data centre.

After allocating VMs to PMs at the cloud data center, we evaluate the performance of our proposed multi-dimensional VMs migration algorithm. In this process, we applied the Google 2019 cluster workload dataset [39]. The Google 2019 cluster dataset consists of fields such as Workload ID, workload arrival time, CPU requirement, memory requirement, storage requirement, etc. To check the superiority of the proposed VMs migration technique, we compared our proposed VMs migration algorithm with recently proposed state-of-the-art VMs migration algorithms such as CPU utilization-based VMs migration, Inter Quartile Range (IQR) [40], SLA aware energy-efficient virtual machine selection (MECpVmS) [21], Look-ahead Energy Efficient VM Allocation (LAA) [38].

Fig. 8 shows the day-wise energy consumption at the cloud data center. The energy consumption of the cloud data center for the proposed MDVMM migration policy is less than other state-of-the-art VMs migration techniques because the proposed VMs migration technique migrate the VMs from source PM to destination PM by taking into account all the resources such as CPU, RAM, and Storage utilization. Further, in the proposed VMs migration algorithm, a PM selection for the VMs migration is based on the energy efficiency and current workload of the PM. Therefore, an optimal number of VM migrations are performed in the proposed VMs migration algorithm. In the case of other existing VMs migration algorithms the selection of VMs is based on the CPU utilization only; therefore aggressive number of VMs migration lead to more energy consumption and SLA violation at the cloud data center.

In our proposed VMs migration algorithm all the VMs from source PM (underutilized state) to destination PM (energy-efficient state) are performed at the cloud data center. After migrating all the VMs from source PM to destination PM the source PM is switched off and thereby, we reduced the energy consumption at the cloud data center.

Since, the VMs allocation and migration problems are not only related to energy efficiency but also related to SLA violation therefore, we evaluated different performance metrics: energy consumption, number of VMs migration, SLA violation, and ESV metrics. Fig. 9 shows the cloud data centre's min, max, and median energy consumption.

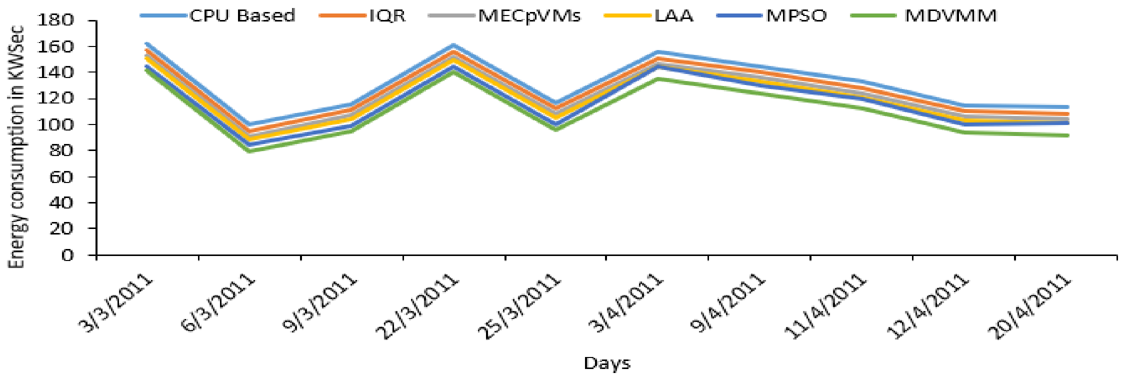

FIGURE 8. Day-wise energy consumption at the cloud data center.

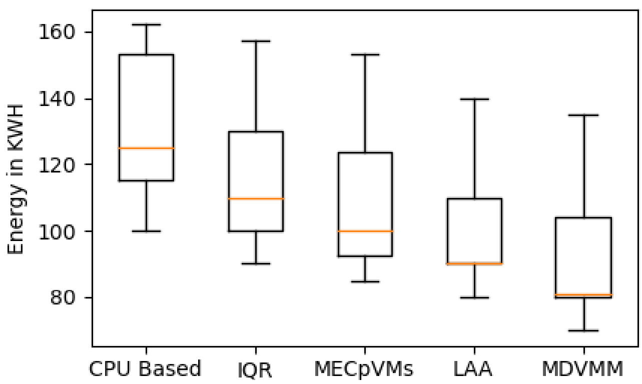

FIGURE 9. Energy consumption.

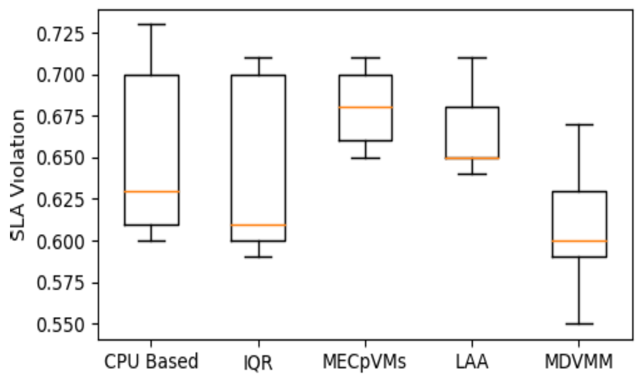

FIGURE 11. SLA violation.

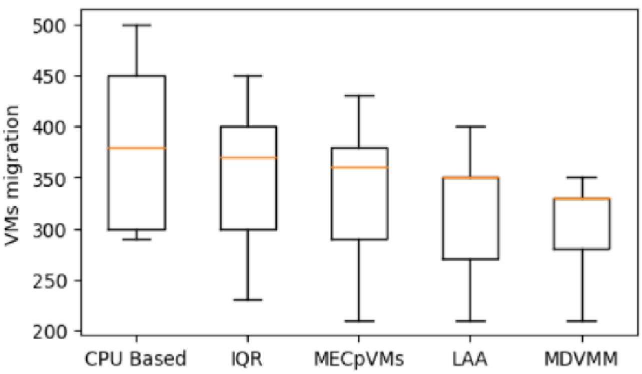

FIGURE 10. VMs migration.

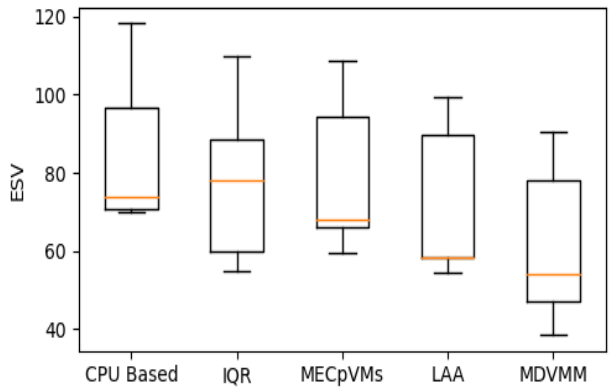

FIGURE 12. ESV metric.

The cloud data centre's min, maximum, and median energy consumption is low in the case of the proposed algorithms because the proposed VMs allocation and migration techniques consolidate the maximum number of VMs into a minimum number of energy-efficient PMs at the cloud data center. The medan power consumption in the case of proposed MDVMM is very close to min power consumption represents the selection of energy-efficient PM as a destination PM during the VMs migration. Fig. 10 shows the min, max, and median number of VMs migrations during the experiment. The min, max, and median number of VMs migration is low in the case of proposed MDVMM algorithm as compared to existing VMs migration algorithms. The Inter quartile Range (IQR) difference between the first quartile range (Q1) and third quarterly range (Q3) i.e. (Q1-Q3) is low in the case of MDVMM as compared to other existing algorithms. The low Inter quartile Range describes the stability of our proposed VMs migration algorithm as compared to the algorithms.

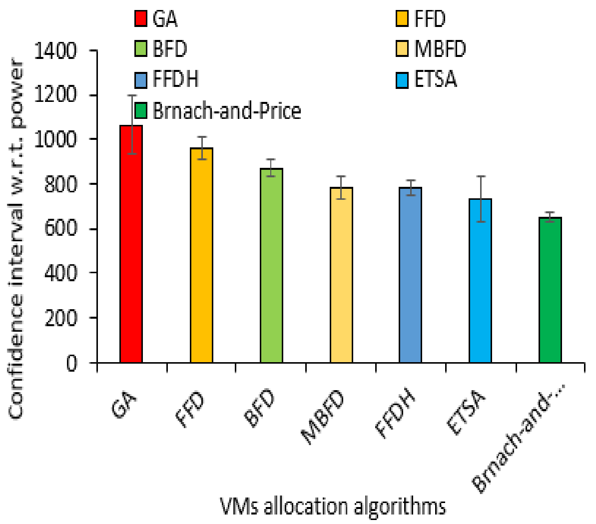

FIGURE 13. Confidence interval w.r.t power.

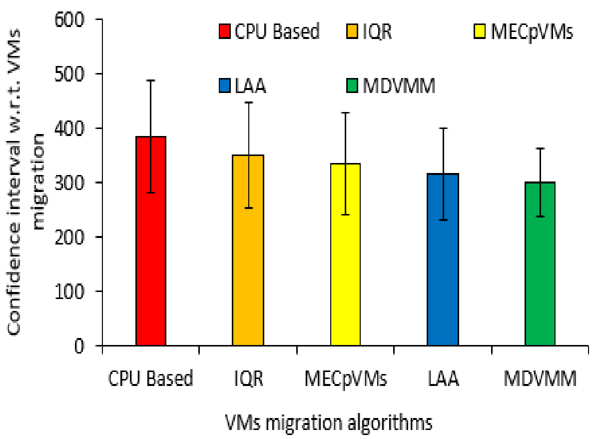

FIGURE 14. Confidence interval w.r.t VMs migration.

Fig. 11 shows the SLA violation of different VMs migration techniques. The SLA violation is low in the case of the proposed MDVMM technique because the proposed approach migrates the VMs by considering all the resources, such as CPU, RAM, and Storage. Hence, an optimal number of VMs migrations promotes less SLA violation at the cloud data center. In the case of other VMs migration technique they only considered the CPU utilization; therefore aggressive (more number of VMs migrations) leads to more SLA violations at the cloud data center.

Fig. 12 shows the performance of different VMs migration algorithms on a combined performance metric ESV. The ESV metric shows the performance of the data center in terms of two Pareto optimal objectives, such as energy consumption and SLA violation. The min, max, and median values are low in the case of the proposed approaches as compared to other state-of-the-art techniques. Since the proposed method performs well over both energy and SLA parameters, the proposed system performs well on the ESV metric.

To check the applicability of our proposed VMs allocation algorithm and migration policy, we conducted our experiment 10 times and calculated the confidence interval of all the algorithms. In this process, with 95% confidence, we calculate upper and lower limits w.r.t. mean power consumption and mean VMs migration. We added and subtracted the margin of error with the point estimate (mean power consumption) to calculate the upper and lower limits. In our experiment, we set the significance value ($\alpha = 0.05$), and the number of experiments (n=10) and calculate the margin of error.

Fig. 13 shows the confidence interval of all the algorithms w.r.t. mean power consumption at the cloud data center. Fig. 13 horizontal bar represents the mean power consumption during ten experiments. A small vertical line on the horizontal bar represents the algorithm's error margin. In our experiment, the mean power consumed by the cloud data center over different VMs allocation algorithms is FFD(1050.6 KW), GA(1020 KW), BFD(865.3 KW), MBFD(820.8 KW), FFDH(819.7 KW), ETSA (798.2 KW), Branch-and-Price(705.9 kW). Further, a margin of error is calculated for GA, FFD, MBFD, FFDH, ETSA, and Branch-and-Price algorithms are 130.15, 49, 38.3, 32.9, 110, and 22, respectively. Among all the VM allocation algorithms, the margin of error is high in the case of GA and ETSA because GA and ETSA are based on the randomized search, resulting in a higher deviation of power consumption during different experiments. Thus, GA and ETSA show low confidence. Furthermore, the proposed Branch-and-Price is an exact algorithm and uses the search tree to calculate the best solution; therefore, the margin of error in the case of Branch-and-Price is low. A low margin of error in the case of the proposed Branch-and-Price VMs allocation algorithm shows the high acceptability of our proposed VMs allocation algorithm for the commercial cloud data center.

Fig. 14 shows the 95% confidence interval w.r.t. mean number of VMs migration. The mean number of VMs migrations calculated for CPU Based, IQR, MECpVMs, LAA, and MDVMM are 1232, 1168, 1094, 1063, and 969, respectively. Further, the margin of error for CPU Based, IQR, MECpVMs, LAA, and MDVMM are 77, 65, 55, 40, and 37, respectively. The optimal number of VMs migrations with a low margin of error shows the superiority and adaptability of the proposed VMs migration algorithm at the cloud data center.

In the proposed MDVMM VMs migration technique, we predicted a PM's over-utilization and under-utilization state by the least square method. The least square method is simple to apply, computationally efficient, and flexible in modelling. Therefore, it is helpful in real-time VMs migration for commercial cloud data centers. But the critical limitation of the least square method is that it assumes linearity and does not consider the seasonality in the time series data.

V. CONCLUSION AND FUTURE WORK

This paper presented a Branch-and-Price-based VMs allocation and MDVMM technique in two phases. In the first phase, a Branch-and-Price-based VMs allocation algorithm is presented, reducing energy consumption and increasing resource utilization. Further, PlanetLab-based VMs were created in the data center. The proposed Branch-and-Price-based VMs allocation algorithm reduced approximately 31% power consumption and increased 20.7% average resource utilization over other state-of-the-art techniques. Further, we evaluated the performance of the proposed MDVMM technique by applying Google cluster load to the data center. The proposed MDVMM technique reduced energy consumption and avoided the SLA violation at the cloud data center. The proposed MDVMM technique migrated an optimal number of VMs with a median value of 3300, which is low compared to other methods. The proposed MDVMM technique also performed well on performance metrics such as energy, SLA violation, and ESV with 85 KWH, 0.61, and 52 median values, respectively. Therefore, in the future, we predict a PM's over-utilization and under-utilization state by applying different statistical models and machine learning algorithms. In the future, we predict a PM's over-utilization and under-utilization state by applying different statistical models and machine learning algorithms. We will also work on reducing the network device's energy consumption by taking the networking device's (switches) power consumption at the cloud data center.

REFERENCES

[1] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, "Cutting the electric bill for internet-scale systems," in *Proc. ACM SIGCOMM Conf. Data Commun.*, Aug. 2009, pp. 123–134.

[2] (Jan. 2023). *U.S. Environmental Protection Agency*. [Online]. Available: http://www.hightech.lbl.gov/documents

[3] Z. Zhou, M. Shojafar, M. Alazab, J. Abawajy, and F. Li, "AFED-EF: An energy-efficient VM allocation algorithm for IoT applications in a cloud data center," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 658–669, Jun. 2021.

[4] (2011). *EPA Report on Server and Data Center Energy Efficiency*. Accessed: Jan. 2023. [Online]. Available: http://www.energystar.gov

[5] M. Pedram, "Energy-efficient datacenters," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 31, no. 10, pp. 1465–1484, Oct. 2012.

[6] M. Dabbagh, B. Hamdaoui, M. Guizani, and A. Rayes, "Energy-efficient resource allocation and provisioning framework for cloud data centers," *IEEE Trans. Netw. Service Manage.*, vol. 12, no. 3, pp. 377–391, Sep. 2015.

[7] L. Guo, C. Lu, and G. Wu, "Approximation algorithms for a virtual machine allocation problem with finite types," *Inf. Process. Lett.*, vol. 180, Feb. 2023, Art. no. 106339.

[8] Y. Gao, H. Guan, Z. Qi, T. Song, F. Huan, and L. Liu, "Service level agreement based energy-efficient resource management in cloud data centers," *Comput. Electr. Eng.*, vol. 40, no. 5, pp. 1621–1633, Jul. 2014.

[9] X. Ruan, H. Chen, Y. Tian, and S. Yin, "Virtual machine allocation and migration based on performance-to-power ratio in energy-efficient clouds," *Future Gener. Comput. Syst.*, vol. 100, pp. 380–394, Nov. 2019.

[10] S. Wang, Z. Liu, Z. Zheng, Q. Sun, and F. Yang, "Particle swarm optimization for energy-aware virtual machine placement optimization in virtualized data centers," in *Proc. Int. Conf. Parallel Distrib. Syst.*, Dec. 2013, pp. 102–109.

[11] N. K. Sharma and G. R. M. Reddy, "Novel energy efficient virtual machine allocation at data center using genetic algorithm," in *Proc. 3rd Int. Conf. Signal Process., Commun. Netw. (ICSCN)*, Mar. 2015, pp. 1–6.

[12] K. H. Kim, A. Beloglazov, and R. Buyya, "Power-aware provisioning of virtual machines for real-time cloud services," *Concurrency Comput., Pract. Exper.*, vol. 23, no. 13, pp. 1491–1505, Sep. 2011.

[13] R. A. C. da Silva and N. L. S. da Fonseca, "Topology-aware virtual machine placement in data centers," *J. Grid Comput.*, vol. 14, no. 1, pp. 75–90, Mar. 2016.

[14] D. Borgetto, R. Chakode, B. Depardon, C. Eichler, J. M. Garcia, H. Hbaieb, T. Monteil, E. Pelorce, A. Rachdi, A. Al Sheikh, and P. Stolf, "Hybrid approach for energy aware management of multi-cloud architecture integrating user machines," *J. Grid Comput.*, vol. 14, no. 1, pp. 91–108, Mar. 2016.

[15] V. Ebrahimirad, M. Goudarzi, and A. Rajabi, "Energy-aware scheduling for precedence-constrained parallel virtual machines in virtualized data centers," *J. Grid Comput.*, vol. 13, no. 2, pp. 233–253, Jun. 2015.

[16] N. J. Kansal and I. Chana, "Energy-aware virtual machine migration for cloud computing—A firefly optimization approach," *J. Grid Comput.*, vol. 14, no. 2, pp. 327–345, Jun. 2016.

[17] S. Ilager, K. Ramamohanarao, and R. Buyya, "ETAS: Energy and thermal-aware dynamic virtual machine consolidation in cloud data center with proactive hotspot mitigation," *Concurrency Comput., Pract. Exper.*, vol. 31, no. 17, Sep. 2019, Art. no. e5221.

[18] H. Li, Y. Zhao, and S. Fang, "CSL-driven and energy-efficient resource scheduling in cloud data center," *J. Supercomput.*, vol. 76, no. 1, pp. 481–498, Jan. 2020.

[19] A. A. Khan, M. Zakarya, R. Buyya, R. Khan, M. Khan, and O. Rana, "An energy and performance aware consolidation technique for containerized datacenters," *IEEE Trans. Cloud Comput.*, vol. 9, no. 4, pp. 1305–1322, Oct. 2021.

[20] J. Masoudi, B. Barzegar, and H. Motameni, "Energy-aware virtual machine allocation in DVFS-enabled cloud data centers," *IEEE Access*, vol. 10, pp. 3617–3630, 2022.

[21] R. Mandal, M. K. Mondal, S. Banerjee, G. Srivastava, W. Alnumay, U. Ghosh, and U. Biswas, "MECpVmS: An SLA aware energy-efficient virtual machine selection policy for green cloud computing," *Cluster Comput.*, vol. 26, pp. 651–665, Jul. 2022.

[22] Z. Li, K. Lin, S. Cheng, L. Yu, and J. Qian, "Energy-efficient and load-aware VM placement in cloud data centers," *J. Grid Comput.*, vol. 20, no. 4, p. 39, Dec. 2022.

[23] A. Ullah and A. Chakir, "Improvement for tasks allocation system in VM for cloud datacenter using modified bat algorithm," *Multimedia Tools Appl.*, vol. 81, no. 20, pp. 29443–29457, Aug. 2022.

[24] P. Memari, S. S. Mohammadi, F. Jolai, and R. Tavakkoli-Moghaddam, "A latency-aware task scheduling algorithm for allocating virtual machines in a cost-effective and time-sensitive fog-cloud architecture," *J. Supercomput.*, vol. 78, no. 1, pp. 93–122, Jan. 2022.

[25] D. Li, Y. Yu, W. He, K. Zheng, and B. He, "Willow: Saving data center network energy for network-limited flows," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 9, pp. 2610–2620, Sep. 2015.

[26] M. Polverini, A. Cianfrani, S. Ren, and A. V. Vasilakos, "Thermal-aware scheduling of batch jobs in geographically distributed data centers," *IEEE Trans. Cloud Comput.*, vol. 2, no. 1, pp. 71–84, Jan. 2014.

[27] Y. Huang, H. Xu, H. Gao, X. Ma, and W. Hussain, "SSUR: An approach to optimizing virtual machine allocation strategy based on user requirements for cloud data center," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 670–681, Jun. 2021.

[28] P. Lv, Z. Zhang, Y. Deng, L. Cui, and L. Lin, "HVMM: A holistic virtual machine management strategy for cloud data centers," *IEEE Trans. Netw. Service Manage.*, early access, Jul. 3, 2023, doi: 10.1109/TNSM.2023.3291890.

[29] W. Cerroni and F. Esposito, "Optimizing live migration of multiple virtual machines," *IEEE Trans. Cloud Comput.*, vol. 6, no. 4, pp. 1096–1109, Oct. 2018.

[30] A. H. Anwar, G. Atia, and M. Guirguis, "A game-theoretic framework for the virtual machines migration timing problem," *IEEE Trans. Cloud Comput.*, vol. 9, no. 3, pp. 854–867, Jul. 2021.

[31] B. Nikzad, B. Barzegar, and H. Motameni, "SLA-aware and energy-efficient virtual machine placement and consolidation in heterogeneous DVFS enabled cloud datacenter," *IEEE Access*, vol. 10, pp. 81787–81804, 2022.

[32] H. Li, G. Zhu, C. Cui, H. Tang, Y. Dou, and C. He, "Energy-efficient migration and consolidation algorithm of virtual machines in data centers for cloud computing," *Computing*, vol. 98, no. 3, pp. 303–317, Mar. 2016.

[33] H. Li, G. Zhu, Y. Zhao, Y. Dai, and W. Tian, "Energy-efficient and QoS-aware model based resource consolidation in cloud data centers," *Cluster Comput.*, vol. 20, no. 3, pp. 2793–2803, Sep. 2017.

[34] A. Beloglazov and R. Buyya, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers," *Concurrency Comput., Pract. Exper.*, vol. 24, no. 13, pp. 1397–1420, Sep. 2012.

[35] IBM-Power-Model. (2014). *Power Consumption Specification of Servers*. [Online]. Available: http://www-03.ibm.com/systems/power/hardware/
[36] P. Cloud Workload. (Jan. 2023). *PlanetLab Cloud Workload Cluster Trace*. [Online]. Available: http://cloudbus.org/cloudsim/
[37] J. Lu, W. Zhao, H. Zhu, J. Li, Z. Cheng, and G. Xiao, ''Optimal machine placement based on improved genetic algorithm in cloud computing,'' *J. Supercomput.*, vol. 78, pp. 3448–3476, Jul. 2021.
[38] S. Talwani, J. Singla, G. Mathur, N. Malik, N. Z. Jhanjhi, M. Masud, and S. Aljahdali, ''Machine-learning-based approach for virtual machine allocation and migration,'' *Electronics*, vol. 11, no. 19, p. 3249, Oct. 2022.
[39] G. C. Workload. (Jan. 2023). *Google Cluster Workload Cluster Trace*. [Online]. Available: https://research.google/tools/datasets/google-cluster-workload-traces-2019/
[40] İ. Çağlar and D. T. Altílar, ''Look-ahead energy efficient VM allocation approach for data centers,'' *J. Cloud Comput.*, vol. 11, no. 1, pp. 1–16, 2022.

NEERAJ KUMAR SHARMA received the B.E. degree from JIT, Khargone, Madhya Pradesh, India, in 2005, the M.E. degree from IET, DAVV, Indore, Madhya Pradesh, in 2010, and the Ph.D. degree from the Cloud Computing, National Institute of Technology Karnataka, Surathkal, India, in 2018, with a focus on energy efficient resources management. He is currently an Assistant Professor with the Department of Computer Science and Engineering, SRM University AP, Amaravati, India. He has published many research articles in high-impact factor journals of IEEE Transactions, Elsevier, and Springer. He has published many papers in international conferences and book chapters. His research interests include machine learning, deep learning, cloud computing/green data centers, computer vision, and NLP.

SRIRAMULU BOJJAGANI (Senior Member, IEEE) received the B.Tech. degree from JNTU Hyderabad, India, the M.Tech. degree in information technology from the Andhra University College of Engineering, Visakhapatnam, India, and the joint Ph.D. degree from the School of Computer and Information Sciences, University of Hyderabad, and the Institute for Development and Research in Banking Technology (Established by the Reserve Bank of India), in 2019, with a focus on design, testing and formal verification of secure mobile payment protocols. He is currently an Assistant Professor with the Department of Computer Science and Engineering, SRM University AP, Amaravathi, India. He has published in many SCI/SCIE-indexed journals and international conference proceedings and was the inventor of several Indian patents. His research interests include the secure protocol design and formal verification of mobile payment applications, IoV, the IoT, and IoMT. He is a Reviewer of many reputed journals, such as IEEE Internet of Things, IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Intelligent Vehicles, IEEE Access, Elsevier, Springer, and Wiley. He has been involved in the research community as a Technical Program Committee (TPC) member for more than a dozen international conferences of high repute. He received a cash prize for the IDRBT Annual Awards 2019 for his research work on ''A secure end-to-end proximity NFC-based mobile payment protocol" published in CSI-Elsevier.

Y. C. A. PADMANABHA REDDY received the B.Tech. degree from MITS, Madanapalle, the M.Tech. degree from RGMCET, Nandyal, Kurnool, and the Ph.D. degree from JNTUA, Anathapuramu, in 2019, with a focus on fast semi-supervised clustering methods to find arbitrary shaped clusters. He is currently an Associate Professor with the Department of CSE, B V Raju Institute of Technology, Narsapur, Telangana, India. His research interests include machine learning, deep learning, and pattern recognition.

MANOJKUMAR VIVEKANANDAN received the Ph.D. degree from the National Institute of Technology, Tiruchirappalli, India, (IDRBT Fellowship) in 2021, with a focus on designing and developing authentication protocols for various applications. He is currently an Assistant Professor with the Department of Computer Science and Engineering, SRM University AP, Amaravati, India. He has published many SCIE-indexed journals with good impact factors, many conference papers, and book chapters of international repute. His research interests include mobile cloud computing, the Internet of Things, authentication, cryptography, and blockchain. He is a Reviewer of some reputed journals, including *The Journal of Supercomputing* and the *Journal of Ambient Intelligence and Humanized Computing* (AIHC).

JAGADEESAN SRINIVASAN received the B.E. degree in computer science and engineering from the Dhanalakshmi Srinivasan Engineering College, Anna University, in 2006, the master's degree in software engineering from Anna University, Tiruchirappalli, in 2009, and the Ph.D. degree in wireless ad-hoc networks from Anna University, Chennai, India, in 2019. He has been in the teaching profession, since 2009. He is currently an Assistant Professor Senior Grade 2 with the School of Information Technology and Engineering, Department of Software and Systems Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India. He has published many SCIE-indexed journals with good impact factors, many conference papers, and book chapters of international repute. His research interests include network security, intrusion detection, cross-layer design, wireless ad-hoc networks, and cyber security.

ANUP KUMAR MAURYA received the B.Sc. degree in mathematics from DDU Gorakhpur University, the master's degree in computer application from Uttar Pradesh Technical University, the M.Tech. degree in artificial intelligence from the University of Hyderabad, and the joint Ph.D. degree in secure wireless sensor networks and Internet of Things from the Institute for Development and Research in Banking Technology (IDRBT, established by RBI) and the University of Hyderabad. He was selected as a Visiting Researcher at The State University of New York at Buffalo (Centre for Unified Biometrics and Sensors), USA, and has worked there on biometric-based efficient user authentication protocols. He is currently an Assistant Professor with the Goa Institute of Management. He has various research publications, including one special mention, where he got the Best Research Paper Award from the International Symposium on Security in Computing and Communication Manipal. He is a reviewer of various reputed international journals.