

A Virtual Machine Placement Strategy Based on Virtual Machine Selection and Integration

Denghui Zhang^{1,*} and Guocai Yin²

¹College of Information Science, Zhejiang shuren university, Hangzhou, 310015, China

²School of Beihua Institute of Aerospace Technology, University of Langfang, Langfang, 065201, China

*Corresponding Author: Denghui Zhang. Email: dhzhang@zjsru.edu.cn

Received: 01 August 2021; Accepted: 15 October 2021

Abstract: Cloud data centers face the largest energy consumption. In order to save energy consumption in cloud data centers, cloud service providers adopt a virtual machine migration strategy. In this paper, we propose an efficient virtual machine placement strategy (VMP-SI) based on virtual machine selection and integration. Our proposed VMP-SI strategy divides the migration process into three phases: physical host state detection, virtual machine selection and virtual machine placement. The local regression robust (LRR) algorithm and minimum migration time (MMT) policy are individual used in the first and section phase, respectively. Then we design a virtual machine migration strategy that integrates the process of virtual machine selection and placement, which can ensure a satisfactory utilization efficiency of the hardware resources of the active physical host. Experimental results show that our proposed method is better than the approach in Cloudsim under various performance metrics.

Keywords: Cloud data centers; virtual machine selection; virtual machine placement; migration; energy consumption

1 Introduction

Recently, the construction and use of large-scale energy-saving data centers has become an critical issue that the government and major IT companies have paid more and more attention [1–3]. A cloud data center is usually configured with a large number of physical hosts. Virtualization is the key technology of the cloud data center. Virtualization encapsulates application tasks and data in the form of virtual machines, and dispatches virtual machines to specific physical nodes for execution through virtual machine allocation strategies. To save energy consumption in data centers, cloud service providers also adopt virtual machine migration strategies to complete the selection and placement of virtual machines on various physical servers. The ultimate goal is to save energy consumption in cloud data centers, improve service quality, save physical space and improve reliability. Virtual machine migration is a process of redeploying virtual machines based on changes in the physical host load that occur during the operation of the data center. The core of virtual machine migration is virtual machine placement, so it is urgent to adopt a feasible virtual machine placement algorithm to optimize the virtual machine migration strategy.

At present, the Cloudsim project developed by Dr. Anton is in the world's leading position in the energy-saving research of cloud data centers [4]. In [4], Cloudsim divides the virtual machine migration process into three steps: physical host status detection, virtual machine selection and virtual machine placement, and finally completes the entire virtual machine migration. Physical host status detection can be judged by observing its resource usage. If the CPU usage efficiency exceeds 90% or less than 10%, it can be considered as over-utilized or under-utilized. In this case, the virtual machine of the physical node will be selected and placed on another physical node at the same time. The meaning of virtual machine selection is to judge the running status of the physical host, and then select a suitable candidate to migrate



the virtual machine. The function of virtual machine placement is to relocate the virtual machines selected in the virtual machine selection algorithm to other most suitable physical nodes in the cloud data center according to a certain algorithm.

Obviously, virtual machine selection and virtual machine placement belong to two different steps, and the process of these two steps can be optimized by designing the corresponding algorithms. Virtual machine selection can be optimized by virtual machine size and resource utilization efficiency. Virtual machine placement is essentially a classical packing problem (CPP), that is, placing a large number of candidate virtual machine VMs on a large number of physical nodes. There are also some literatures that call this problem stable matching or hospital bed allocation problem [5–6].

The current research on virtual machine migration strategy focuses on the virtual machine selection algorithm and the virtual machine placement algorithm itself, while ignoring the distinct relationship between them. In fact, virtual machine selection can affect virtual machine placement. At least in a private cloud, both virtual machine selection and virtual machine placement can be controlled by one organization. This paper focuses on virtual placement phase, and introduce an efficient virtual machine placement strategy (VMP-SI) based on virtual machine selection and integration, and completes the efficient resource utilization of the virtual machine and the physical host by designing a stable matching algorithm. In the physical host status detection and virtual machine selection phase, our proposed VMP-SI strategy uses the basic ideas and methods in Cloudsim. In the virtual machine placement phase, we propose our own integration strategy. Similar to the allocation of hospital beds, VMP-SI determines that the virtual machine and the physical host are stable and reliable according to the priority of each individual of the matching parties. This mechanism can ensure that the physical resources of the active host have a satisfactory utilization efficiency. Using practical workload data, the VMP-SI strategy is implemented and simulated through Cloudsim. The simulation results show that the VMP-SI strategy can save the energy consumption of the cloud data center better than the ordinary virtual machine migration strategy in Cloudsim, reduce the rate of service level agreement violations, and ensure the quality of service.

2 Related Work

2.1 Virtual Machine Selection

Regarding the problem of virtual machine selection in data centers, there have been quite a lot of researches at home and abroad, and most of the research focuses on the computing power [7–9] and hardware resources of the physical host, such as memory size [10–11]. The research goal of these works is to seek a balance between performance and the cost of virtual machine.

There are also some studies that put the virtual machine selection algorithm in the minimum calculation cost of the virtual machine. Some researchers [12–15] believe that the cost of a virtual machine is proportional to the use time of the virtual machine, while some studies [16] believe that the cost should be calculated by the use of consumption resources or the use of physical resources under long-term lease. In the Cloudsim project, the minimum migration time (MMT) based virtual machine selection strategy is adopted, that is, a virtual machine that can be migrated in the shortest time is selected as a candidate for migration. Literature [4] shows that this strategy is significantly better than other virtual machine selection strategies. The method proposed in this paper draws on the MMT strategy at this phase.

2.2 Virtual Machine Placement

There are also many studies on virtual machine placement issue. Some research focuses on the computing power of physical hosts and the computing load of virtual machines [17–19]. Some other studies focus on the placement factors of physical resources in terms of memory size, disk space size, network bandwidth size, and I/O communication capabilities [20–23]. One of the most important factors in virtual machine placement research is the number of turn-on PMs, because it can determine the energy consumption to the greatest extent [17]. The other studies take into account the number of PMs overloads of the dynamic energy consumption of the physical host [24]. In the virtual machine placement, if the

resources are relatively single and only the target is placed under the reduced number of physical hosts, this type of problem can be classified as a classical packing problem, which has no optimal solution, only sub-optimal solution.

2.3 Combination of Virtual Machine Selection and Virtual Machine Placement

There is also part of the work that does not separate virtual machine selection and virtual machine placement, but discovers the correlation between them. For example, literature [25] selects the optimal virtual machine size to allocate virtual machines according to the characteristics of the task. Its purpose is to reduce energy consumption by minimizing the use of resources, regardless of its load. Each virtual machine has a fixed size, and the difference between virtual machine sizes and the overall size of its load are specified. Literature [26] maps the application components in the software-as-a-service cloud platform to the virtual machine. It also focuses on the size of the virtual machine and the sharing capabilities of the virtual machine. This strategy also puts forward suggestions for virtual machine placement algorithms, which virtual machine placement can be processed statically, and which virtual machine placement needs to be dynamically processed, but in fact the placement of these virtual machines on the physical host has not been completed, but an external operating algorithm is used. Compared to literature [26], this paper mainly focuses on the stable matching of the virtual machine selection process and the virtual machine placement process, and completes the efficient resource utilization of the virtual machine and the physical host by designing a stable matching algorithm.

3 Preliminary of Our Proposed Virtual Machine Migration Strategy

3.1 Overview

Our proposed VMP-SI divides the virtual machine migration process into physical host load detection, virtual machine selection process and virtual machine placement process, and finally completes the entire virtual machine migration process. The details are as follows.

Step 1: periodically detecting the over-utilized or under-utilized status of the physical host in the cloud data center.

Step 2: determining whether to perform virtual machine migration operations.

Step 3: completing virtual machine selection based on reliable matching.

Step 4: placing candidate virtual machines for virtual machine placement operations.

According to the usage threshold of physical resource utilization efficiency, the existing physical host status detection methods in the Cloudsim project include the five common ones: local regression, local regression robust, median absolute deviation, static threshold, and inter quartile range.

In the above Step 3, the existing virtual machine selection algorithms in the Cloudsim project include the following types:

1) Maximum migration: that is, the virtual machine with the highest correlation with the CPU usage efficiency on the same physical host is selected as the object.

2) Minimum migration time: that is, to migrate a virtual machine that can be completed in the shortest time as the object of choice.

3) Minimum utilization: that is, to migrate a virtual machine with the least use efficiency.

4) Random selection: that is, a virtual machine is randomly selected on the physical host for migration.

The difference between the VMP-SI virtual machine migration strategy proposed in this paper and the Cloudsim project is that in the subsequent virtual machine placement process, the combination of virtual machine selection and virtual machine placement is integrated, and the virtual machine and physical host are considered to match each individual's priority rather than classical packing method.

3.2 System Model and Metric

The utilization efficiency of physical hosts reflects the resource utilization status of the cloud data center, and the average utilization efficiency of all physical hosts can be expressed as

$$\bar{u} = \frac{\sum_{i=1}^m u_i}{m} \quad (1)$$

where u_i is the utilization efficiency of i -th physical host. m is the total number of physical hosts.

In order to calculate the energy consumption of the cloud data center, assuming that the server is idle, the percentage of its energy consumption is k , P_{full} represents the energy consumption of the server when it is fully loaded, and the energy consumption of the entire cloud server is expressed as

$$E = \sum_{i=1}^m kP_{full} + (1-k)P_{full}u_i \quad (2)$$

3.3 The Proposed VMP-SI Strategy

The VMP-SI virtual machine migration strategy mainly works under the condition of infrastructure as a service. Assuming that a cloud data center has N heterogeneous physical server nodes, in a given period of time, multiple clients submit applications to the platform. The application is encapsulated into M virtual machines by virtualization technology to the platform. In the next, the M virtual machines will be allocated to specific physical servers for execution. According to the previous description, it should include processes such as physical host status detection, virtual machine selection, and virtual machine placement.

In the first two phases, the VMP-SI uses LRR to complete physical host status detection, and MMT to complete virtual machine selection, respectively. In the third phase, the virtual machine placement phase, VMP-SI adopts a method that takes into account the voluntary priority conflicts of both the virtual machine and the physical host. The specific procedures of the algorithm are as follows. The notations are described in Table 1.

Table 1: Notations used in VMP-SI

Symbols	description
u_i	CPU utilization efficiency of current virtual machine V_i
U_j	CPU utilization efficiency of current virtual machine P_j
m_i	Computing power of virtual machine V_i (MIPS)
M_j	Computing power of physical host P_j
U_{ij}	CPU utilization efficiency of the physical host after completing the virtual machine placement
V_{list_j}	According to the priority of the physical host, a list of virtual machines that may be migrated

Step 1: Determine the highest priority physical host for each virtual machine in the candidate migration list.

a) Assuming that there are $\alpha \leq M$ virtual machines to be placed on $\beta \leq N$ physical hosts, the set of virtual machines is defined as V , $V = \{V_1, V_2, \dots, V_\alpha\}$, defining the set of available physical hosts as P , $P = \{P_1, P_2, \dots, P_\beta\}$.

b) For each virtual machine V in the set V_i , calculate U_{ij} in the physical host set P , which is expressed as

$$U_{ij} = \frac{(U_j M_j + u_i m_i)}{M_j} \quad (3)$$

c) Calculate the difference between the ideal target utilization efficiency U_{th} and the evaluation utilization efficiency U_{ij} . If $U_{th} > U_{ij}$, the difference value calculation formula is

$$\Delta U_{ij} = U_{th} - U_{ij} \tag{4}$$

Otherwise, it indicates that the physical host is not suitable as the physical node to be placed.

d) The physical host $P_{k(i)}$ is a physical node that is easily placed first by the virtual machine V_i , then there is

$$k(i) = \arg \min \Delta U_{ij}, 1 \leq j \leq \beta \tag{5}$$

According to the priority, multiple virtual machines can also select the same physical host, where V_{list_j} is defined as the candidate placement list that is most easily placed on the physical host P_j , which is sorted based on priority.

Step 2: Match virtual machine to physical host.

- a) For each physical host in the P set, match P_j to V_{list_j} , so that you can get a minimum ΔU_{ij} .
- b) After each matching process is completed, the V_{list_j} list is discarded, and V_i is deleted from the V set.

Step 3: If V is not equal to the empty set, repeat Steps 1 and 2; otherwise, terminate the entire placement step.

The steps show that the difference between the utilization efficiency of physical resources and the ideal utilization efficiency is the smallest, and the utilization efficiency of the physical host is relatively high. After each matching is completed, the matched virtual machine will be deleted from the candidate virtual machine list. The steps are executed until all virtual machines in the candidate virtual machine list are matched, and finally a virtual machine migration map is returned.

4 Experimental Results and Performance Analysis

4.1 Simulation Setup

In order to put the VMP-SI model in the virtual machine selected by the fusion virtual machine proposed in this paper, the Cloudsim simulator is used. The latest version of Cloudsim provides many experimental data of energy consumption models of existing virtual machine migration strategies, which facilitates the comparison and research of energy-related virtual machine migration algorithms for cloud data centers by researchers [5]. All experiments are performed independently 1000 times, and then averaged [27–28]. 1000 simulations are enough to ensure the convergence of the final result, so we choose this number of times.

According to the three phases steps of virtual machine migration, the better method in Cloudsim is to use the LRR strategy to detect the state of the physical host, combine the MMT strategy to complete the virtual machine selection, and then combine the CPP to complete the virtual machine placement. The above method is called a hybrid strategy, and its implementation results refer to [4–5]. The strategy proposed in this paper is mainly compared with the hybrid algorithm. The simulated cloud data center is mainly composed of two types of physical servers. The total number of physical servers is 800. The physical server configuration is shown in Table 2.

Table 2: Physical host hardware configure of cloud data center

Server nickname	Number of servers	Number of kernel	Computing power (MIPS)	RAM/GB
HP Proliant M1110G Xeon 3040 G4	400	2	1860	4
HP Proliant M1110G Xeon 3075 G5	400	2	2660	4

Simulating the application access of the cloud client, using 10,000 task data with correct CPU and memory requests, running a web application or another application with different types of workloads, a total of about 4 types of virtual machines, as described in Table 3.

Table 3: Four types of virtual machine of cloud data center

Virtual machine type	Network bandwidth (Mbit/s)	Required computing power (MIPS)	Required memory size (GB)
1	100	2500	2.5
2	100	2500	2.5
3	100	1000	2.5
4	100	500	2.5

The energy model is the energy consumption model described in Section 2.2, which is also a commonly used model in Cloudsim. The virtual machine migration cycle is set to 5 minutes. This setting means that the load detection of the virtual machine will run every 5 minutes for a total of 24 h. The energy consumption in one day is counted each time, and the operation is repeated 5 times in a week.

4.2 Evaluation Metrics

According to the research ideas of the Cloudsim project, there are 4 main metrics for evaluating the virtual machine migration strategy: 1) the overall energy consumption of the cloud data center; 2) number of virtual machine migrations; 3) average SLA violation rate analysis; 4) energy and SLA violations (ESV).

4.3 Simulation Results

In this paper, the simulation of VMP-SI, a virtual machine migration strategy that combines virtual machine selection and virtual machine placement, is implemented on Cloudsim. The resource utilization efficiency threshold is adjusted to 0.8, compared with the hybrid strategy. The difference between the method in this paper and the hybrid method is that a new virtual machine and physical host matching algorithm is used in the virtual machine placement phase. In the VMP-SI algorithm, the result of virtual machine selection is integrated, which reflects the combination of virtual machine selection and virtual machine placement, so that the virtual machine and the physical host can achieve a stable match.

Fig. 1 compares the overall energy consumption of cloud data centers. As can be seen from Fig. 1 the overall energy consumption of the VMP-SI strategy is lower than that of the hybrid strategy. On average, at least 30% of total energy consumption can be saved from Monday to Friday. The reason is that VMP-SI aims to improve the utilization efficiency of physical host resources. If the utilization efficiency is improved, it will directly reduce resource energy consumption, rather than as in a hybrid strategy. The classical packing method and the recursive classical method are used to reduce the number of active physical hosts as the goal.

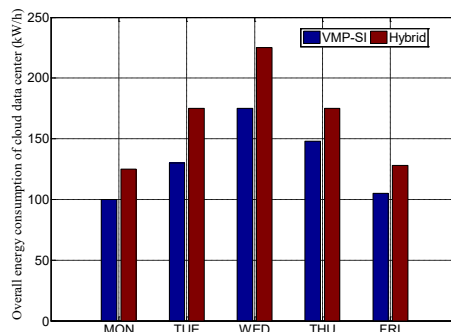


Figure 1: Total energy consumption performance results of cloud data center

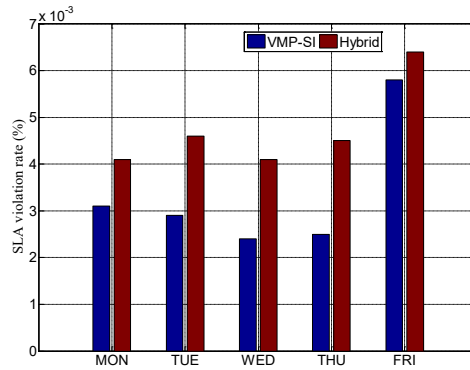


Figure 2: SLA violation rate performance results of cloud data center

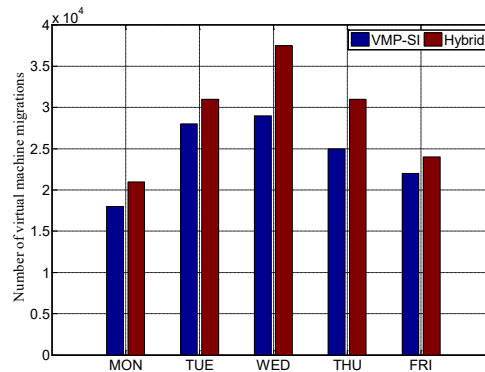


Figure 3: Virtual machine migration numbers results of cloud data center

Fig. 2 shows the SLA violation rate of cloud data centers. The SLA violation rate of the VMP-SI strategy is lower than that of hybrid strategy, and its SLA violation rate is concentrated between 0.002% and 0.007%. This means that for 10,000 cloud client requests, only about 0.2 to 0.7 physical resources cannot be allocated. Since each virtual machine migration affects the rate of system energy consumption and SLA violations, the number of virtual machine migrations is also very important. Fig. 3 shows the experimental results of the VMP-SI strategy that integrates virtual machine selection and virtual machine placement. The number of virtual machine migrations under the VMP-SI is significantly reduced. It is precisely because of the small number of migrations that the reliability and energy-saving performance of the cloud data center will be further improved.

In terms of energy and SLA violations (ESV), the smaller the value of ESV, the better the overall performance of the cloud data center. It can be seen from Fig. 4 that the VMP-SI is lower than hybrid strategy. This is because after adopting the migration strategy that combines virtual machine selection and virtual machine placement, the CPU utilization efficiency of all physical servers can be controlled very close to an ideal range, which increases the utilization efficiency of virtual machine resources. The VMP-SI strategy can shut down those idle physical servers, which ultimately saves the central energy consumption of the transportation bureau.

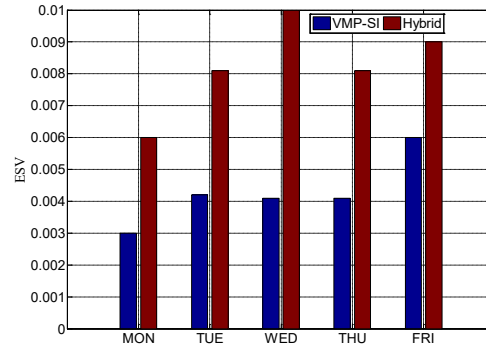


Figure 4: Total ESV performance results of cloud data center

5 Conclusions

This paper proposes the virtual machine placement strategy VMP-SI for the virtual machine selection of the integrated virtual machine in the operation data center. This strategy mainly works in the virtual machine placement phase, which targets the stable matching of both the virtual machine and the physical host, and controls the utilization efficiency of the physical resources of the cloud data center within an ideal threshold range. Experiments show that the strategy proposed in this paper has better performance under many performance evaluation metrics.

Acknowledgement: The dataset used in this paper are available from the corresponding author upon request.

Funding Statement: There is no specific funding to support this research.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

- [1] Z. Lei, E. Sun, S. Chen, J. Wu and W. Shen, "A novel-copy algorithm for live migration of virtual machine," *Future Internet*, vol. 9, no. 3, pp. 1–13, 2017.
- [2] X. Jin, F. Zhang, L. Wang, S. Hu, B. Zhou *et al.*, "Joint optimization of operational cost and performance interference in cloud data centers," *IEEE Transactions on Cloud computing*, vol. 5, no. 4, pp. 697–711, 2015.
- [3] J. Su, Z. Sheng, A. X. Liu and Y. Chen, "A partitioning approach to RFID identification," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2160–2173, 2020.
- [4] 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 and Computation: Practice and Experience*, vol. 24, no. 13, pp. 1397–1420, 2012.
- [5] G. Kim and W. Lee, "Stable matching with ties for cloud-assisted smart tv services," in *Proc. of IEEE Int. Conf. on Cloud Computing Technology and Science*, pp. 163–171, 2011.
- [6] A. Alashaikh, E. Alanazi and A. A. Fuqaha, "A survey on the use of preferences for virtual machine placement in cloud data centers," *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–39, 2021.
- [7] W. Li, J. Tordsson and E. Elmroth, "Modeling for dynamic cloud scheduling via migration of virtual machines," in *Proc. of IEEE 3rd Int. Conf. on Cloud Computing*, pp. 163–171, 2011.
- [8] Z. A. Mann, "Allocation of virtual machines in cloud data centers—A survey of problem models and optimization algorithms," *ACM Computing Surveys*, vol. 48, no. 1, pp. 1–34, 2015.
- [9] R. Buyya, S. Srirama, G. Casale, R. Calheiros and H. Shen, "A manifesto for future generation cloud computing: research directions for the next decade," *ACM Computing Surveys*, vol. 51, no. 5, pp. 1–38, 2018.

- [10] U. Lampe, M. Siebenhaar, R. Hans, D. Schuller and R. Steinmetz, "Let the clouds compute: cost-efficient workload distribution in infrastructure clouds," in *Proc. of Int. Conf. on Grid Economics and Business Models*, pp. 91–101, 2012.
- [11] D. Oliveira, K. Ocana, F. Baiao and M. Mattoso, "A provenance-based adaptive scheduling heuristic for parallel scientific workflows in clouds," *Journal of Grid Computing*, vol. 10, no. 3, pp. 521–552, 2012.
- [12] L. Bittencourt, E. Madeira and N. Fonseca, "Scheduling in hybrid clouds," *IEEE Communications Magazine*, vol. 50, no. 9, pp. 42–47, 2012.
- [13] J. Tordsson, R. Momtero, R. Vozmediano and I. Llorente, "Cloud brokering mechanisms for optimized placement of virtual machines across multiple providers," *Future Generation Computer Systems*, vol. 28, no. 2, pp. 358–367, 2012.
- [14] Z. Mann, "Multicore-aware virtual machine placement in cloud data centers," *IEEE Transactions on Computers*, vol. 65, no. 11, pp. 3357–3369, 2016.
- [15] P. Hoenisch, C. Hochreiner, D. Schuller, S. Schulte, J. Mendling *et al.*, "Cost-efficient scheduling of elastic processes in hybrid clouds," in *Proc. of IEEE 8th Int. Conf. on Cloud Computing*, pp. 17–24, 2015.
- [16] T. Genez, L. Bittencourt and R. M. Madeira, "Work flow scheduling for SaaS/PaaS cloud providers considering two SLA levels," *Computer Science Review*, vol. 30, pp. 31–54, 2018.
- [17] D. Breitgand and A. Epstein, "SLA-aware placement of multi-virtual machine elastic services in compute clouds," in *Proc. of 12th IEEE Int. Sym. on Integrated Network Management and Workshops*, pp. 161–168, 2011.
- [18] M. Guazzone, C. Anglano and M. Canonico, "Exploiting VM migration for the automated power and performance management of green cloud computing systems," in *Proc. of Int. Workshop on Energy Efficient Data Centers*, pp. 81–92, 2012.
- [19] Z. Mann, "Rigorous results on the effectiveness of some heuristics for the consolidation of virtual machines in a cloud data center," in *Future Generation Computer Systems*, vol. 51, pp. 1–6, 2015.
- [20] O. Biran, A. Corradi, M. Fanelli, L. Foschini and E. Silvera, "A stable network-aware VM placement for cloud systems," in *Proc. of 12th IEEE/ACM Int. Sym. on Cluster, Cloud and Grid Computing*, pp. 498–506, 2012.
- [21] Y. Gao, H. Guan, Z. Qi, Y. Hou and L. Liu, "A multi-objective and colony system algorithm for virtual machine placement in cloud computing," *Journal of Computer and System Sciences*, vol. 79, no. 8, pp. 1230–1242, 2013.
- [22] M. Mishra and A. Sahoo, "On theory of VM placement: Anomalies in existing methodologies and their mitigation using a novel vector based approach," in *Proc. of IEEE 4th Int. Conf. on Cloud Computing*, pp. 275–282, 2011.
- [23] L. Tomas and J. Tordsson, "An autonomic approach to risk-aware data center overbooking," *IEEE Transactions on Cloud Computing*, vol. 2, no. 3, pp. 292–305, 2014.
- [24] P. Svard, W. Li, E. Wadbro, J. Tordsson and E. Elmroth, "Continuous datacenter consolidation," in *Proc. of IEEE 7th Int. Conf. on Cloud Computing Technology and Science*, pp. 387–396, 2015.
- [25] S. Piraghaj, A. Dastjerdi, R. Calheiros and R. Buyya, "Efficient virtual machine sizing for hosting containers as a service," in *Proc. of IEEE World Cong. on Services*, pp. 31–38, 2015.
- [26] R. Ganesan, S. Sarkar and A. Narayan, "Analysis of SaaS business platform workloads for sizing and collocation," in *Proc. of IEEE 5th Int. Conf. on Cloud Computing*, pp. 868–875, 2012.
- [27] J. Su, R. Xu, S. Yu, B. Wang and J. Wang, "Idle slots skipped mechanism based tag identification algorithm with enhanced collision detection," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 5, pp. 2294–2309, 2020.
- [28] J. Su, R. Xu, S. Yu, B. Wang and J. Wang, "Redundant rule detection for software-defined networking," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2735–2751, 2020.