Acta Technica 62 No. 2A/2017, 187–198

Virtual technology of cache and data real time allocation in cloud computing data center[1]

Wenrong Bai[2]

Abstract. In order to develop a new technology used in data analysis and caching, a virtual technology of cache and data real time allocation is proposed. As a new commercial computing service model, users can obtain information services anytime and anywhere through cloud computing. Cloud computing data center has thousands of physical nodes, the application of virtualization technology allows the system to complete the dynamic migration between nodes and data resources on-demand distribution, thus supporting the data center resources, cache and distribution. However, with the expansion of cloud computing, the inherent dynamics of the system and the complexity of management are gradually increasing. A heuristic cloud computing data center resource management approach is proposed to reduce energy costs and ensure customer service quality. The minimum migration strategy is used to keep the number of VMS transferred at a minimum. In order to evaluate the improved algorithm comprehensively, the energy evaluation, SLA violation rate, virtual machine transfer cost and average migration time are compared and analyzed. The results show that the proposed RFFD method can reduce the cost of energy and reduce the cost of virtual machine migration in the case of lower SLA violation rate. Based on the above finding, it is concluded that this new real-time data distribution virtual technology plays an important role in data caching and analysis of cloud computing center.

Key words. Cloud computing, data centers, virtual technology, resource allocation.

1. Introduction

Cloud computing is a product of the integration of distributed computing, parallel computing, grid storage and virtualization, and it has great commercial value. In recent years, with the advent of the networked information age, cloud computing has received widespread attention and rapid development as an emerging business model. It is seen as the third wave of IT following changes in computers and the

[1]The authors acknowledge the National Natural Science Foundation of China (Grant: 51578109), the National Natural Science Foundation of China (Grant: 51121005).

[2]Inner Mongolia Electronic Information Vocational Technical College, 010070, Hohhot, Inner Mongolia, China

internet. However, virtualization is the key technology for cloud computing. Cloud computing data center obtains a lot of advantages, such as flexible management, high resource utilization and strong scalability, making the data center management more convenient and efficient [1].

At the same time, with the rapid development of cloud computing, the surge of energy consumption of basic equipment, such as network equipment and servers, has led to high operating costs, resulting in serious resource and environmental problems. Therefore, at this stage, the urgent problem is to develop new technologies to manage and optimize cloud computing data center resources. Sharkh studied resource allocation problem in cloud computing center under network virtual environment, and designed a new algorithm. Xiao Zhen used real-time migration technology to dynamically manage and integrate virtual machine resources [2]. A resource management method for heuristic cloud computing, data center caching and real-time data allocation based on frame is proposed around the theme of cloud computing data center resource management. Through the evaluation and analysis of cloud computing, data center structure, energy model and related performance indicators, virtual machine management method based on frame is proposed after the improvement of MBFD algorithm. Finally, the simulation is carried out in order to improve the utilization rate of cloud computing data, liberate human resources and compress management costs.

2. Experimental procedure

2.1. Cloud computing and data center for cloud computing

At present, the most authoritative definition of cloud computing comes from American National Institute of Standards and Technology. Cloud computing is a model of computing resources that can be independently controlled. Users can access the resources in the model with a convenient, on-demand way via Internet. These resources exist in a dynamically shared pool of resources and can be obtained and released in a quick and intelligent manner [3]. Cloud computing is not a new technology appeared suddenly, but gradually evolved by combining the traditional computing technology and the network technology. As shown in Fig. 1, compared with traditional computing technology, they are both continuous and distinct. They are the aggregation of advanced computing and service technology.

Cloud computing can be divided into three categories from the service type. The first is to provide services based on virtual infrastructure, that is, infrastructure services (IaaS). The second is to provide software development platform based services, that is, platform services (PaaS). The last one is the service that provides the application software directly to the user, namely software service (SaaS). Because of the diversity of cloud computing services, its architecture is relatively complex. As shown in Fig. 2, from the bottom layer to the top layer, it can be roughly divided into four layers: hardware resource layer, virtualization resource layer, management component layer and service interface layer [4].

Data center (DC) is a complete set of complex facilities that include not only

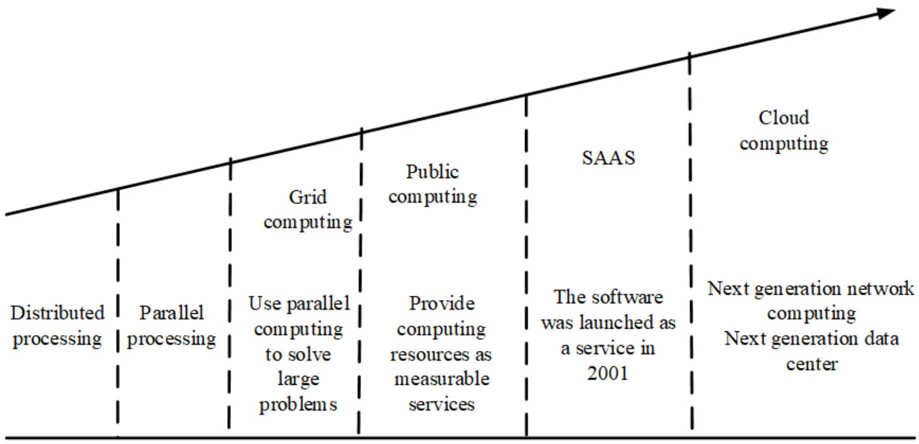

Fig. 1. History of cloud computing

the basic computer system and the physical equipment, but also the monitoring, scheduling and security mechanisms responsible for its normal operation [5]. It is the ultimate bearer of cloud computing services, providing basic support for customer applications such as resources and services. The physical resources such as computing, storage and network in data center are constructed into dynamically adjusted virtual resource pools, so that the basic service units change from physical hosts to virtual machines. The automatic deployment, dynamic expansion and on-demand allocation of cloud computing resources are realized. Users can access cloud computing services on demand and out-of-the-box.

2.2. Overview of virtualization technology

Virtualization technology is the foundation of cloud computing, including network virtualization, storage virtualization, server virtualization, desktop virtualization, application virtualization, presentation virtualization. In general, virtualization is an abstraction layer that separates physical hardware from the operating system to provide greater IT resource utilization and flexibility. Processor, memory, storage and network hardware resources are abstracted into standardized virtual hardware. It is packaged in a hardware-independent virtual machine along with a complete operating environment, including operating systems and applications. The server virtualization principle is depicted in Fig. 3.

By using virtualization software, virtualized data center software abstracts the underlying hardware device and the upper operating system, and manages the resources using appropriate methods. Through virtualization, virtual can create multiple virtual machines on a server. At this point, the virtual machines are isolated from each other, while the applications, operating systems and hardware devices are encapsulated so that the system can be quickly backed up and deployed quickly.

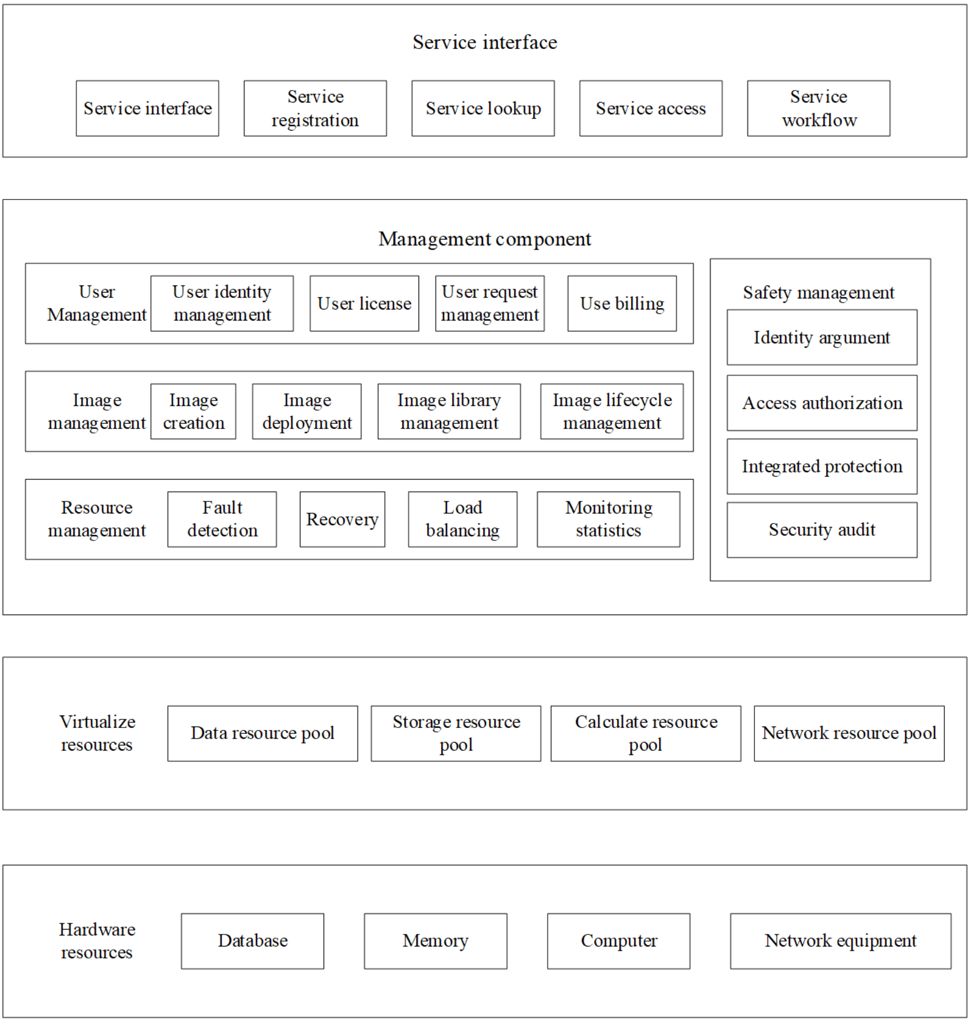

Fig. 2. Cloud computing architecture diagram

2.3. The system model of virtual cloud computing center

Starting from the basic composition of data center, the composition of data center structure is analyzed. It mainly includes host, frame, cooling system and network switch and other physical equipment. At the same time, the energy calculation method of data center and the index of evaluating data center performance are put forward. The energy cost of cloud computing data centers is divided into three parts, namely, host, cooling equipment and network [6]. The energy cost of the host is mainly generated by CPU, memory, hard disk and network interface, in which the energy cost of CPU accounts for most of the total overhead. In order to calculate the energy cost of the host more accurately, the energy overhead of the host is collected

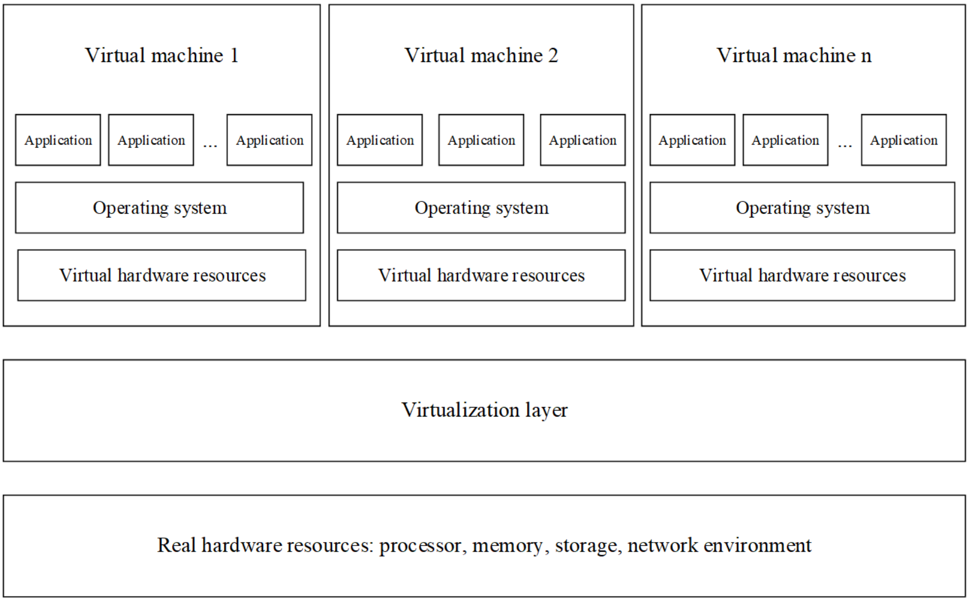

Fig. 3. Server virtualization principle

every 10 % of the host CPU utilization from 0 % to 100 %, with a total of 11 nodes of data. For hosts with CPU utilization falling at each 10 % interval, a specific energy cost can be calculated from the linear relationship

$$P_i = P_{\mathrm{pre}} + (u_i - o_{\mathrm{pre}}) \times (P_{\mathrm{after}} - P_{\mathrm{pre}}) \times 10\,. \tag{1}$$

In formula (1), P_i represents the host energy cost, and u_i is the CPU utilization of the host. P_{pre} represents the energy overhead of the previous node when the CPU utilization is u_i. P_{after} represents the energy overhead of the latter node when the CPU utilization is u_i.

Let P_{TE}, P_{Agr}, and P_{CR} represent the energy overhead of the rack upper-layer switch, aggregation layer switch, and core routing, respectively. Then

$$P_{\mathrm{Net}} = P_{\mathrm{TE}} + P_{\mathrm{Agr}} + P_{\mathrm{CR}}\,. \tag{2}$$

The total energy cost of the rack is $P_{\mathrm{R}} = \sum_{j=1}^{n} P_j$ (n is the total number of running racks), and the total energy cost of the host is $P_{\mathrm{S}} = \sum_{i=1}^{m} P_i$ (m is the total number of running hosts). The total energy cost of the cloud computing data center can be expressed as

$$P_{\mathrm{DC}} = P_{\mathrm{S}} + P_{\mathrm{R}} + P_{\mathrm{Net}}\,. \tag{3}$$

There are four main metrics for quantifying the performance of a virtualized cloud computing data center, namely, energy overhead, SLA breach rate, virtual machine migration cost and average migration time. The cost of energy is calculated as described above. SLA violation rate refers to the ratio of the total time of all virtual machine SLA violations in the data center to the sum of the time the virtual machines run. Let symbol n represent the number of virtual machines in the data center. $ViolateTime(VM_i)$ is the virtual machine SLA collision time, and $Runtime(VM_i)$ is the virtual machine running time. Then

$$sla\_violate\_rate = \frac{\sum_{i=1}^{n} ViolateTime(VM_i)}{\sum_{i=1}^{n} Runtime(VM_i)}\,. \tag{4}$$

Virtual machine migration costs is represented by the sum of all virtual machine migration costs. It is related to the specific transfer cost of each virtual machine and the selected destination node. When the destination node and the virtual machine are in the same frame, the transfer cost is the number of migration. When not in the same frame, the migration cost is multiplied by a constant K (the constant value is higher than 1)

$$migrate\_cost = \sum_{i=1}^{n} migrateCost(VM_i)\,. \tag{5}$$

Here, $migrateCost(VM_i)$ is the total cost of the virtual machine being migrated. The average migration time of a virtual machine represents the average time spent in the migration of the virtual machine. Symbol n is the total number of virtual machine migration times.

Let $migrateTime_(i)$ represent the time it takes to migrate. Then

$$average\_migration\_time = \frac{\sum_{i=1}^{n} migrateTime_i}{n}\,. \tag{6}$$

2.4. Virtual machine management algorithm and improvement

The core problem of virtualization cloud computing data center resource management is the migration of virtual machines and virtual machine placement. Virtual machine migration will virtually increase the overhead of CPU and network bandwidth. Therefore, the minimum migration strategy is adopted to reduce the migration of the virtual machine [7]. The core idea is that the least amount of the virtual machine is used to control the resource utilization of the host within the threshold. The virtual machine placement problem is improved by the best fit decreasing algorithm. First, the descending order is sorted according to the MIPS required by the virtual machine, and then the best placement host is found from the resource utilization higher than the threshold lower value and the host set below the threshold upper limit. If there is no host that meets, it is searched from the collection of hosts whose resource utilization is below the threshold. If there is no host that conforms, a host is turned on and the virtual machine is placed on it. On this basis, virtual machine placement algorithm based on frame is proposed. First, a lower limit is set for rack utilization. When the rack utilization is below the threshold, all virtual machines on the rack are migrated. After that, the host on the rack is turned off, and the rack itself and the network device are used to save energy.

It is assumed that V_j is the set of virtual machines allocated on host j, and $u_{a(v)}$ represents the proportion of CPU allocated to virtual machine v. The purpose of least migration strategy is to find the set $V_{mm} \in V_j$ that satisfies the formula

$$V_{mm} = \begin{cases} \left\{V \left| V \in V_j, u_j - \sum_{v \in V} u_{a(v)} T_u, |V| \to \min \langle u_j \, \mathrm{T}_u \rangle \right.\right. \\ V_j \langle u_j \, \mathrm{T_u} \rangle , \\ \emptyset \langle \text{other} \rangle . \end{cases} \tag{7}$$

In order to evaluate the performance of the resource management method proposed in the virtualized cloud computing data center, the rack components were expanded in the simulation platform. At the same time, five algorithms of MBFD, IMBFD, RFFD, R & D virtual machine placement algorithm, NURBN (non-low load rack optimal node algorithm) and NURMG (non-low load rack minimization interval algorithm) are implemented [8]. Simulation experiments are conducted to verify the performance of different algorithms in terms of energy cost, SLA violation rate, virtual machine transfer cost and average migration time. Meanwhile, a comparative analysis is carried out.

3. Results and discussion

3.1. Total energy cost evaluation

Figure 4 compares the total energy costs of five different virtual machine resource management methods for MBFD, IMBFD, RFFD, NURBN, and NURMG, while Fig. 5 shows the total transfer cost of data center virtual machines. Compared with the MBFD algorithm, the energy utilization rates of IMBFD, RFFD, NURBN and NURMG have been improved by 1.7 %, 8.1 %, 6.3 % and 10.1 % respectively. The RFFD, NURBN and NURMG can achieve a relatively large increase in efficiency because these rack-based virtual machine management methods can reduce the number of racks running, thereby reducing the cooling system and rack energy costs [9]. With the increase of the number of virtual machines, the improved MBFD algorithm is more efficient than virtual machine placement algorithm based on frame, especially the rack based virtual machine placement method, which makes the energy saving more obvious.

3.2. Cost analysis of virtual machine migration

Both NURBN and RFFD have the preference to migrate virtual machines to rack hosts whose frame load is greater than the threshold limit. This reduces the number of racks that load is below the threshold, thus reducing the migration of the virtual machine. As a result, the cost of their virtual machine migration is reduced by 2.7 % and 0.2 %, respectively, compared with MBFD. With the increasing number of virtual machines, the cost of IMBFD, NURMG and NURBN transfer is in a

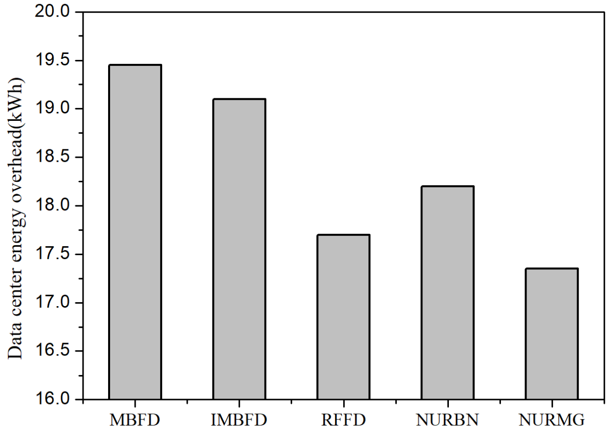

Fig. 4. Comparison of energy cost of data center under different resource management methods

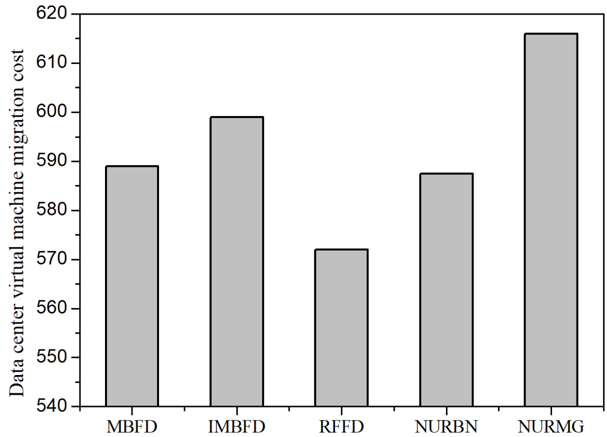

Fig. 5. Comparison of total transfer cost of data center virtual machines under different resource management methods

relatively stable state with the MBFD contrast. On the contrary, RFFD increases as the number of virtual machines increases, and the cost of virtual machine migration decreases further.

3.3. Comparison of service grade contract violation rate

As shown in Fig. 6, NURMG has the highest SLA violation compared to other methods. Because it minimizes the distance between the host's resource load and the set host load upper limit, it is easy to cause SLA violation because of changes in the resource load. NURBN attempts to find the destination host on the rack where all the resources load is greater than the threshold threshold. This can reduce the number of hosts, but also lead to relatively high host resource load. Therefore, compared with RFFD, its SLA violation rate is still much higher [10]. With the increase of the number of virtual machines, the SLA violation rate of resource management method proposed by MBFD has been improved and the range of increase is in a steady state with sufficient resources. Figure 7 compares average migration time of data center virtual machine under different resource management.

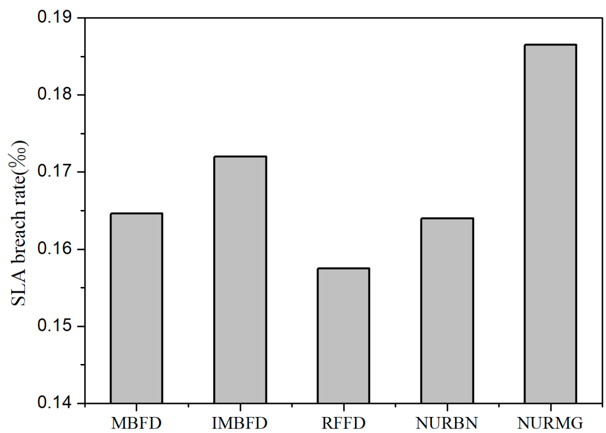

Fig. 6. Comparison of data center SLA violation rates under different resource management methods

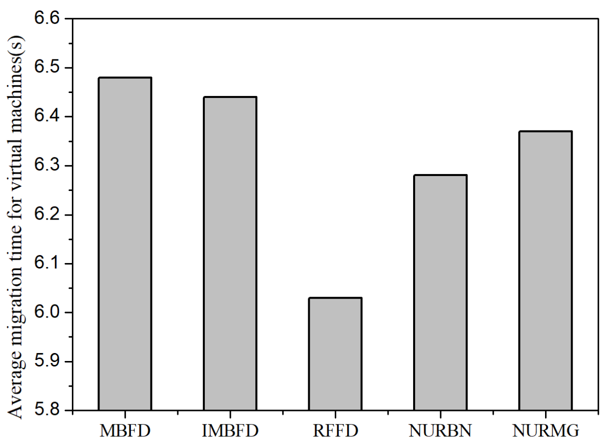

Fig. 7. Comparison of average migration time of data center virtual machine under different resource management

3.4. Calculation of average migration time of virtual machines

The virtual machine placement algorithm based on frame can reduce the average migration time of the virtual machine. Because these algorithms first look for the destination node on the rack node which is not below the threshold, thus increasing the probability that the virtual machine migrates to the frame. It takes less time for virtual machines to migrate between nodes within the rack. In these methods, RFFD has the least average migration time. With the increase of the number of virtual machines, the average migration time of virtual machines is relatively stable relative to MBFD.

RFFD is the best choice from the comprehensive performance of energy cost, SLA violation rate, virtual machine transfer cost and average migration time. Because the RFFD algorithm preferentially puts the virtual machine on the other physical nodes of the frame. If the other nodes of this frame cannot place the virtual machine, it is arranged in descending order according to the utilization ratio of the rack, and then the first frame for placing the virtual machine can be selected. This not only reduces the number of racks, but also reduces the virtual machine migration costs and average migration time.

4. Conclusion

Aiming at the technology of data center cache and real-time data allocation in cloud computing, a heuristic cloud computing data center resource management method is proposed. The structure and energy calculation model of cloud computing data center are analyzed. The least migration strategy is used to screen the virtual machines, and the number of VMS is guaranteed to be the least. Then, in view of the placement of virtual machines, a virtual machine placement algorithm based on frame is proposed according to the improved heuristic algorithm. By comparing the performance of several algorithms, the results show that compared with the previous methods, the proposed heuristic cloud computing data center resource management method reduces the energy consumption in the case of guaranteeing low SLA collision rate.

Due to the complexity and complexity of cloud computing resources, there are many influencing factors besides the energy cost, SLA breach rate, migration cost and average migration time, which need to be considered in the management of cloud computing data center. Subsequent research can improve the virtual machine migration progress and access to better cloud computing data management performance. At the same time, it can improve the stability of the system in order to develop the virtual technology in the cloud computing data center for further application.

References

[1] A. Beloglazov, J. Abawajy, R. Buyya: *Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing.* Future Generation Computer Systems *28* (2012), No. 5, 755–768.

[2] M. A. Sharkh, M. Jammal, A. Shami, A. Ouda: *Resource allocation in a network-based cloud computing environment: Design challenges.* IEEE Communications Magazine *51* (2013), No. 11, 46–52.

[3] Z. Xiao, W. Song, Q. Chen: *Dynamic resource allocation using virtual machines for*

cloud computing environment. IEEE Transactions on Parallel and Distributed Systems *24* (2013), No. 6, 1107–1117.

[4] S. Mohapatra, S. Mohanty, S. Pattanayak, A. Hota: *Comparison of various platforms in cloud computing.* International Journal of Computer Applications *162* (2174), No. 7, 28–33.

[5] H. T. Dinh, C. Lee, D. Niyato, P. Wang: *A survey of mobile cloud computing: architecture, applications, and approaches.* Wireless Communications and Mobile Computing *13* (2013), No. 18, 1587–1611.

[6] R. Moreno-Vozmediano, R. S. Montero, E. Huedo, I. M. Llorente: *Cross-site virtual network in cloud and fog computing.* IEEE Cloud Computing *4* (2017), No. 2, 46–53.

[7] F. Al-Haidari, M. Sqalli, K. Salah: *Impact of CPU utilization thresholds and scaling size on autoscaling cloud resources.* International Conference on Cloud Computing Technology and Science (CloudCom), 2–5 December 2013, Bristol, UK, IEEE Conference Publications *2* (2013), 256–261.

[8] D. Yang, C. Deng, Z. Zhao: *Dynamic scheduling method of virtual resources based on the prediction model.* International Conference on Collaborative Computing: Networking, Applications and Worksharing, 10–11 November 2016, Beijing, China, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering *316* (2016), 384–396.

[9] W. Tian, Y. Zhao: *A toolkit for modeling and simulation of real-time virtual machine allocation in a cloud data center.* Optimized Cloud Resource Management and Scheduling (2015), 217–243.

[10] S. Hosseinimotlagh, F. Khunjush, R. Samadzadeh: *SEATS: Smart energy-aware task scheduling in real-time cloud computing.* Journal of Supercomputing *71*, (2015), No. 1, 45–66.

Received July 12, 2017