International Journal of Engineering and Management Research e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-13, Issue-3 (June 2023)
https://ijemr.vandanapublications.com https://doi.org/10.31033/ijemr.13.3.24

QoS-Aware Virtual Machine Migration in Cloud Environment- A Review

Vijay Bhadouria[1], Sandeep Kumar Jain[2] and Avneesh Vashishtha[3]
[1]Research Scholar, Department of Computer Science, Dr. Bhimrao Ambedkar University, Agra, INDIA
[2]Associate Professor, Department of Computer Science, Dr. Bhimrao Ambedkar University, Agra, INDIA
[3]Assistant Professor, Department of CSE, Geeta University Panipat, Haryana, INDIA

[1]Corresponding Author: vijay81nov@gmail.com

Received: 23-05-2023 **Revised:** 09-06-2023 **Accepted:** 26-06-2023

ABSTRACT

Cloud computing is an evolving field of information technology that allows for online access to computational resources that are maintained in huge data centers. These data centers have physical machine or servers but to fulfill the user request efficiently one physical machine may have several virtual machines. Virtualization technique with the help of hypervisor software serves this purpose. The paper investigates the role of virtualization in providing resources effectively and various parameters involved in live VM allocation. We have also discussed available methods and techniques of virtual machine allocation. Our aim for this review is to boost the chances of delivering unbiased research that accurately reflects the current viewpoint for allocating virtual machines in cloud data centers.

***Keywords*--** Cloud Computing, Virtualization, Virtual Machine, Migration Parameters, Resource Utilization

I. INTRODUCTION

Cloud computing is an emerging field of today's era and it is used in every field. Various cloud service providers fulfill the user requests of resources through data centers. Data centers have physical machines or servers. One physical machine may have several virtual machines. A virtual machine is a logical representation of physical machine or server. At server side due to significant ups and down in workload or user demand, VMs are moved dynamically from one physical machine to another to fulfill user request effectively. This process is known as virtual machine migration. Hypervisor or virtual machine monitor is used for this purpose. This paper focuses on virtual machine migration in respect of Quality of Service (QoS).

We have tried to find out answers of following research questions throughout our review paper in various sections.

a. How does resource allocation work in virtual machines?
b. What are the parameters generally considered in Virtual Machine allocation/re-allocation?
c. What are the resource allocation challenges in the cloud data center?
d. What are available methods and techniques for virtual machine allocation?
e. What are future research directions for VM Migration?

The paper is divided in following parts. Research methodology for QoS-aware VM migration in cloud environment is illustrated in Section 2. Virtualization technology, resource type and different approaches are described in Section 3. Classification of various approaches is discussed and analyzed in section 4. Analysis of review is presented in section 5. Section 6 presents benefits for researchers at last conclusion at section 7.

II. RESEARCH METHODOLOGY

Research methodology for a review paper is a process of classification and analysis of approaches of available literature. Therefore we review evaluate and interpretation of all available research studies on virtual machine migration in cloud.

2.1 Planning Review

The first step in planning is to determine the requirements for a review and the final step is creating and validating the review procedure. We searched Springer link, Science direct, Google Scholar, and IEEE Xplore digital libraries with following keywords like "cloud computing" and "virtual machine allocation/ migration" and found the need for a review for quality of service aware VM Migration.

2.2 Selected Data Source

We have selected some important indexing like IEEE Xplore, ACM Digital Library, Science direct, Springer link, and Google Scholar were used for this review. We gathered total 973 research papers for this survey and after applying exclusion criteria finally we have selected 63 studies as shown in given table 2.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-13, Issue-3 (June 2023)
https://doi.org/10.31033/ijemr.13.3.24

Table-2: Number of selected studies

S. No	Databases Directories	Selected
1	Springer Link	10
2	Science Direct	11
3	Google Scholar	12
4	ACM Digital Library	12
5	IEEE Xplore Digital Library	18
Total		**63**

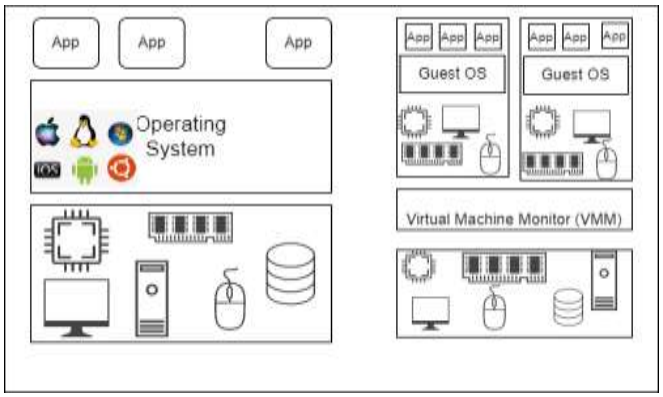
2.3 Criteria for Inclusion and Exclusion

The retrieved entries as outcome were reviewed using inclusion criteria. Research articles from peer-reviewed journals and international conferences have been included Research articles are written in English and focused on virtual machine allocation; technical debt. White papers, non-peer-reviewed papers, non-English scripts. Abstracts, tutorials, presentations, essays, or editorial articles. Research papers not focusing on a virtual machine in IaaS or not explicitly showing methods, approaches, and tools to facilitate virtual machine allocation.

To address our first question the studies were examined based on how the resources have been allocated to virtual machines and for this fundamental mechanism of virtualization and the requirement of resources to be provisioned is being explained in Section 3.

III. VIRTUALIZATION

It is actual technique behind cloud computing. It creates virtual version of resources like server, storage, operating system or even network. This is possible by *hypervisor* [8] [26]. Hypervisor provides an environment to guest operating system or virtual machine as they are running on independent physical machine (host machine) and manages requests of VMs to access physical host's resources like CPU, RAM, storage etc. as shown in figure 1.



(a) Without Virtual Machine (b) With Virtual Machine

Figure 1

Hypervisor can be further categorized into two types. First is Type-1 hypervisor or bare-metal hypervisor which runs on system hardware directly. The main advantage of this type of hypervisor is that it does not affect other running virtual machines. Second is Type-2 hypervisor or hosted hypervisor that runs on a present operating system but the problem with this type of hypervisor is that if host operating system fails it disturbs the hypervisor also.

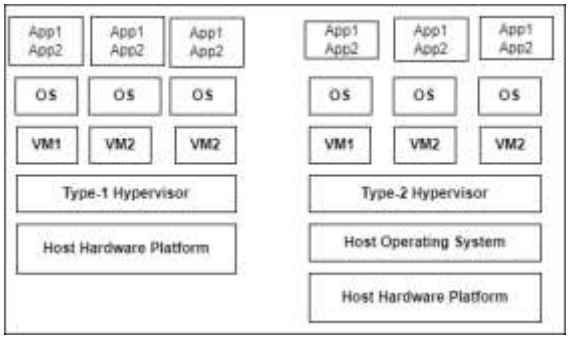


Figure 2: Hypervisors Types

Virtualization provides server consolidation (run different operating system simultaneously) and encapsulation (also provides security, reproducibility, monitoring, migration, legacy systems) etc. The challenges in virtualization include mapping between physical machine and virtual machine, virtual machine live migration, horizontal and vertical virtual machine scaling.

3.1 Different Resources

The main resources for cloud that required to be provisioned for virtualization [ ]

Compute Recourses: Compute resource means set of actual physical resource like CPU cores, Memory, Virtual Network Interface Card (VNIC), and local input-output which have capability for processing data.

Network Resources Multiple host or physical machines are packaged into racks in data centers. These hosts are connected with a high-speed network bandwidth based on fiber cables.

Storage Resources Amazon, Rackspace, Google Cloud offers different storage services. These storage is in the form of virtual disks and storage area network. Storage capacity is increased or decreased according to dynamic nature of workload.

Power Resources Energy efficiency also a major issue to run data centers because data centers and network infrastructure need a significant amount of power consumption that also lead to carbon di-oxide emission.

3.2 Parameters Considered for VM Allocation

To address our second question the studies were examined which parameters, in general, considered for the allocation to the virtual machine: we find out which parameters in general, considered for allocation to the virtual machine.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-13, Issue-3 (June 2023)
https://doi.org/10.31033/ijemr.13.3.24

(a) Service Level Agreement (SLA) SLA is a legal agreement or a signed contract between customer and service providers. SLA ensures the performance and availability of the virtual machine to the customer. If any one of them not fulfill the SLA terms he has to pay penalty according to the SLA clause [6]. SLA focuses on execution time, availability, responsiveness, jitter and billing etc. [7][9]

(b) Load Balancing Load balancing means managing the user requests by dividing workload among various available virtual machines in such a manner that none of the VMs remains idle. Effective load balancing assures throughput, migration time, scalability and overall performance etc.(A.V.n.d) (Narayan 2016).

(c) Quality of Service (QoS) QoS is a major challenge that holds different issue like budget, response time, deadline, availability of overall infrastructure. Different QoS policies are required for different types of application running on a virtual machine.[7][9].

(d) Workload The workload is the number of requests that the VM has been allocated to process at a given time. The workload is determined by the arrival of number of jobs on a virtual machine at a particular time. For better performance the workload on VMs must be dynamically adjusted to ensure that each VM gets the capacity it needs. [11]

3.3 Resource Allocation Challenges

To answer our third question we have identified several resource allocation challenges which are being discussed as follows:

Workload Prediction: If demand of allocated resources is greater than existing service instance capacity, this may result in a possibility of decreasing performance of application. The VMs should be capable of performing the listed tasks as under.(i) Predicting future incoming workloads (ii) allocation of jobs to VMs (iii) Provisioning necessary virtual machines in advance on self-adaptive basis.

High Uncertainty: Even today, it is a major difficulty for cloud providers to choose the configuration that avoids SLA breach. The uncontrolled change in IaaS refers to unexpected increase in demand of resources. Violation of SLA occurs on account of high uncertainty of resources. This as QoS of service parameter the application performance is decreased.

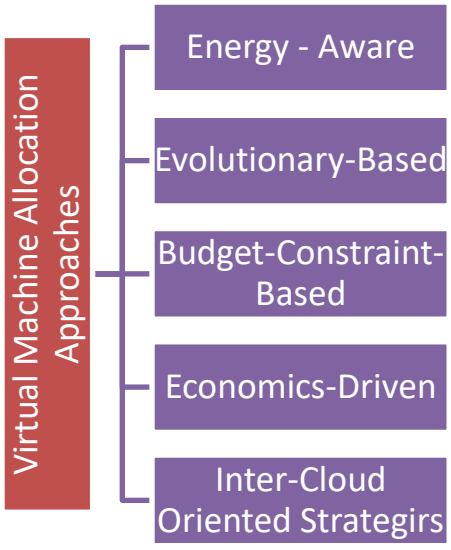
Resource- Intensive Energy Consumption: This is also a big concern as the higher energy consumption increase operational cost and creates environment related issues. There is a increase in the energy required to operate a data center with the no. of VMs increased. To reduce the energy consumption of data centers, cloud service providers are using green computing aware resource allocation algorithm.

Resource Optimization: Virtualized memory, applications, processor, or a server are major resource components of cloud based on the user request. As the cloud is highly dynamic, elastic and without human intervention, with high level QoS the allocation of various resources is a greater challenge. A self adaptive approach is required to maximize the resource optimization more accurately.

Economic Driven Decision Users can definitely achieve efficient resource utilization and optimization if all VMs operating on all PMs are utilized continuously. Clearly, VMs will not be fully exploited if the no. of users is not enough at a point of time. This result in decrease in the profit and then increase the financial burden of service providers. A self-adaptive method built on past data can be used to solve this issue. When tracking and analyzing the number of customers joining or leaving VMs over time, which is critical from a revenue perspective, the cloud service providers' debt burden can be reduced.

IV. CLASSIFICATION OF APPROACHES

To answer our fourth question we have identified, compared, and classified the existing methods and techniques that support virtual machine allocation decisions as given in the figure below.



4.1 Energy-Aware Strategies

Several studies have already done on energy-aware strategies. We have shown some important papers in tabular form below;

Table 4: Energy-Aware Strategies

Authors Name	Proposed Approach	QoS Parameters Considered
Chimakurthiet et al.[18]	Energy-efficient mechanism	Throughput, Response time
Qian et al.[19]	Energy efficiency	Cost, dynamic Voltage
Beloglazovet et al.[20]	Energy-aware resource Management system	Operational Cost
Y.Chen et al.[21]	Energy-aware	Cost, Energy
S.Zhang et al.[22]	Fair scheduling policy and Energy-aware policy	Cost, Energy, Performance
L.Chen et al.[23]	Spatial-awareness approach	Cost, Energy
I.S.Moreno et al.[24,229]	The dynamic resource provisioning mechanism	Energy, Cost
S.K. Tesfatsion et al.[25]	Energy-aware	Energy, Cost

4.2 Evolution Based Strategies

Various meta-heuristic and genetic algorithms on evolution based strategies is proposed by different authors are listed below;

Table 5: Evolutionary-Based Strategies

Authors Name	Proposed Approach	QoS Parameters Considered
Z.Xiao et al.[26]	Skewness algorithm	Energy
Wei et al.[27]	Evolutionary mechanism	Cost, time
Ferdauset al.[28]	Ant Colony Optimization	Power consumption
Liu et al.[29]	Ant Colony Optimization	Cost
Joseph et al.[30]	Genetic Algorithm	Cost, time, Energy
N.K.GholiAbad[31]	Genetic Algorithm	Performance
S.A.Hamad[32]	Genetic Algorithm	Cost, Speedup, Efficiency
J.Hu et al.[33]	Genetic Algorithm	Cost
S.Ho Jang et al.[34]	Genetic Algorithm	Performance
J.Gu et al.[35]	Genetic Algorithm	Cost

4.3 Budget-Constraint Strategies

Various resource management approaches which reduce time and cost given by different authors are listed below;

Table 6: Budget-Constraint Strategies

Authors Name	Proposed Approach	QoS Parameters Considered
S.Mehta et al.[36]	ReCon Framework	Cost
Sotomayor et al.[37]	Batch Processing/Cost-effective VM lease mgt.	Throughput, Run time
Tram Truong Huu[38]	Cost-based approach	Cost
Lee et al.[39]	Pricing model	Cost, Response time
Entrialgo et al.[40]	Cost-based approach	Cost, Performance
Zhu et al.[41]	Dynamic provisioning Technique for shared data Center	Cost
Li et al.[42]	RDCM a multicast approach	Throughput
Farahnakian et al.[43]	Hierarchical agent-based architecture	Cost, Enegy
Meng et at.[44]	Joint VM provisioning Approach	Performance constraint
N.Kumar et al.[45]	Demand-based preferential Resource allocation technique	Cost

W.Lin et al.[46]	Threshold-based dynamic Resource allocation	Cost
S.Schulte et al.[47]	Vienna Platform	Cost, Time

4.4 Economic Driven Strategies

Based on fuzzy logic algorithm H. Arabnejad et al.[48] proposed a dynamic learning approach. This approach deals with load traffic and able to efficiently scale up/down cloud resources while minimizing infrastructure and management cost. M.Fokaefs et al.[50] investigated and formalized the opportunities for economical as well as optical resource scaling decisions some other research papers are also listed below;

Table 7: Economic-Driven Strategies

Authors Name	**Proposed Approach**	**QoS Parameters Considered**
H.Arabnejad et al.[48]	Dynamic learning strategy based On a fuzzy logic algorithm	Cost
J.Chen et al.[49]	Tradeoffs between profit and Customer satisfaction	Profit, customer satisfaction and instance utilization.
M.Fokaefs et al.[50]	Economically optimal scaling Decisions	Cost, Profit
H.jin et al.[51][27]	Optimized fine-grained pricing Scheme	Profit
Y.Kouki et al.[52]	Cloud service level Agreement	Cost
Y.Kouki et al.[53]	Delta Scaling	Cost
A.Pandey et al.[54]	Hybrid planning approach	Response time
J.A.Pascual et.al.[55]	Reduction of the time required By VM management operations	Cost
D.Perez-Palacin et al.[56]	Synthesis of cost-effective Adaptation plan framework	Cost
C.Qu et al.[57]	Fault-tolerant mechanism-based Cost-effective auto-scaling system	Cost, Response time, availability

4.5 Inter-Cloud Oriented Strategies

Dynamic server migration algorithm are better than static resource allocation in terms of resource consumption and SLA violations. In IaaS cloud architecture, if virtualized resources intelligently allocated the revenue of cloud providers can be maximized. Various inter-cloud oriented strategies are listed below;

Table 8: Inter-Cloud Strategies

Authors Name	**Proposed Approach**	**QoS Parameters Considered**
Clark et al.[58]	Live Operating Systems Migrations	Service Management
Bobroff et al.[59]	VMs live migration	Energy Consumption
Song et al.[60]	Resource Optimization	Profit, Cost
Zhu et al.[61]	Dynamic provisioning tech. for shared data centers	Profit, Cost
Mishra et al.[62]	Novel Vector-based approach For VM placement	Energy
Nguyen et al.[63]	AGILE	Energy
Strunk et al.[64]	Live Migration	Energy
Chen et al.[65]	Resource provisioning Strategies	Energy
Mosa et al.[66]	Optimized VM allocation	Energy

V. ANALYSIS OF REVIEW

The aim for this review is to increase the possibility of producing unbiased exploration reflecting current perspective towards better performance SLA and QoS for virtual machine migration in the cloud data center. Moreover, this paper indicates how the virtual machine migration approaches share contributions to deal with virtual machine migration challenges in the cloud data center. We identified resource allocation challenges; different parameters used for virtual machine allocation; and classified virtual machine allocation approaches based on energy-aware, evolutionary-based, budget-constraint-based, economics-driven strategies, and inter-cloud strategies. Further, we find the research gap in the current VM allocation approaches and argued how we can fill this gap with our proposed approach.

VI. BENEFITS OF RESEARCHERS AND PRACTITIONERS

This section discusses our fifth question the future research directions for VM Migration.

Findings of the SLR show that researchers paid a lot of attention to the VM allocation in the past. The outcome of this SLR provided some key issues that could help in findings the future research. In the past, different VM allocation approaches had been studied that were based on different parameters like QoS, SLAs, power consumption, resource utilization, VM optimization, etc. [21]. This SLR identified numerous research gaps in the VM allocation approaches which are mostly related to dynamic VM adaptation decisions; uncertainties in the cloud operating environment; consumption of energy by the idle server. The dynamic workload on the VM encounters the problem of over-provisioning or under-provisioning. Over-provisioning of a VM results in scarcity of resources while under-provisioning violates SLA. Setting upper and lower CPU utilization limits is a fundamental concept; CPU utilization must be maintained within certain ranges. A VM must be regarded as over-provisioned if the CPU utilization drops below the lower threshold. In this instance, migrating the present VM is necessary to reduce idle power consumption. Similar to this, VM under-provisioning occurs when CPU utilization rises above a certain level. In this situation, the VM must be moved once more to avoid a possible SLA violation. We argue that we may pay a penalty for under-provisioned VM by considering the future cost instead of the current cost. Besides, CPU utilization we may consider memory and network bandwidth as other parameters also for a VM to be considered as either over-provisioning or under-provisioning.

REFERENCES

[1] R. Creasy, "The Origin of the VM/370 Time-Sharing System," *IBM Journal of Research and Development*, 1981.

[2] R. Goldberg, "Survey of Virtual Machine Research," *IEEE Computer Magazine*, 1974.

[3] P. Gum, "System/370 Extended Architecture: Facilities for Virtual Machines," *IBM Journal of Research and Development*, 1983.

[4] A. R. P. G. G. J Popek, "Formal Requirements for Virtualizable Third Generation Architectures," *ACM*, vol. 17, no. 7, pp. 412-421, 1974.

[5] B. J. a. R. Stadler, "Resource Management in Clouds: Survey and Research Challenges," *Springer*, pp. 1-53, 2014.

[6] Z. A. Mann, "A taxonomy for the virtual machine allocation problem," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 9, 2015.

[7] A. V. e. al., "economic-driven-strategies-for-virtual-machine-allocation-in-cloud-data-center".*https://deepai.org/publication/economic-driven-strategies-for-virtual-machine-allocation-in-cloud-data-center*.

[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, 2015.

[9] a. P. V. A.Vashsitha, "Economic-Driven Strategies for Virtual Machine Allocation in Cloud Data Center," *IJET*, pp. 10-15, 2019.

[10] S. A. K. S. a. B. U. Banerjee, "Development and Analysis of a New Cloudlet Allocation Strategy for QoS Improvement in Cloud," *Arabian Journal of Science & Engineering*, vol. 40, p. 1409–1425, 2015.

[11] C. Y. G. D. S. S. W. B. R. W. Z. X. a. H. C. Wang. Z., "AppRAISE:Application-Level Performance Management in Virtualized Server Environments," *IEEE Transactions on network and service management*, vol. 6, no. 4, pp. 240-254, 2008.

[12] S. K. P. V. a. R. P. A. Vashishtha, "A Self-Adaptive View on Resource Management in Cloud Data Center," in *8th International Conference on Cloud Computing, Data Science & Engineering, IEEE*, Noida, Inida, 2018.

[13] G. H. X. L. a. H. M. Y. Zhang, "Integrating resource consumption and allocation for infrastructure resources on-demand," 2010.

[14] a. R. S. B. Jennings, " Resource Management in Clouds: Survey and Research Challenges," *Journal of Network*

and System Management, Springer, vol. 23, no. 3, pp. 567-619, 2015.

[15] M. M. C. I. P. a. F. M. Filho, "Approaches for optimizing virtual machine placement and migration in cloud environments: A survey," *J. Parallel Distrib. Comput*, pp. 222-250, 2018.

[16] R. B. T. C. K. L. a. R. B. S. Kumar, "Multi-Tenant Cloud Service Composition Using Evolutionary Optimization," in *24th IEEE International Conference on Parallel and Distributed Systems*, 2018.

[17] S. a. R. M. Narayanan, "A Comprehensive Study on Load Balancing Algorithms in Cloud Computing Environments," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 13, no. 10, pp. 794-799, 2016.

[18] C. a. K. S. Lskrao, "Power efficient resource allocation for clouds using Ant Colony Framework," arXiv, 2011.

[19] H. a. M. D. Qian, "Data Center Resource management with temporal dynamic workload," in *In the proceedings of IFIP/IEEE International Symposium on Integrated Network Management*, Belgium .

[20] A. a. B. R. A. Belogiazov, "Energy efficient resource management in virtualized cloud data centers," in *In the proceedings of 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, Australia, 2010.

[21] Y. D. A. S. W. W. Q. a. G. N. Chen, "Managing server energy and operational costs in hosting centers," *ACM SIGMETRICS Perform. Eval. Rev., vol. 33(1):1, 303–314.*, vol. 33, no. 1, pp. 303-314, 2005.

[22] S. L. Y. W. B. a. Z. R. Zhang, "A novel resource allocation algorithm for a heterogeneous data center," in *In the Proceedings of International Conference Information Science and Applications (ICISA)*, Suwon, South Korea, 2013.

[23] L. a. S. H. Chen, "Consolidating complementary VMs with spatial/temporal-awareness in cloud data centers," in *In the Proceeding of IEEE INFOCOM*, Toronto, ON, Canada,, 2014.

[24] I. S. M. a. J. Xu, "Customer-aware resource overallocation to improve energy efficiency in realtime cloud computing data centers," 2011.

[25] E. W. a. J. T. Selome K Tesfatsion, "A combined frequency scaling and application elasticity approach for energy-efficient cloud computing," *Sustainable Computing: Informatics and Systems*, vol. 4, no. 4, pp. 205-214, 2014.

[26] W. S. a. Q. C. Z. Xiao, "Dynamic resource allocation using virtual machines for cloud computing environment," *IEEE transactions on parallel and distributed systems*, vol. 24, no. 6, p. 1107–1117, 2013.

[27] G. V. A. X. N. Wei, "Scheduling parallel cloud computing services: An Evolutional Game," in *In the proceedings of First International Conference on Information Science and Engineering*, China.

[28] M. M. M. C. R. a. B. R. Ferdaus, "Virtual machine consolidation in cloud data centers using ACO metaheuristic," in *In :Silva ., Dutra I., Santos Costa V. (eds) Euro-Par 2014 Parallel Processing*, vol. 8632, Euro-Par 2014. Lecture Notes in Computer Science, vol 8632. Springer, Cham, 2014.

[29] X. Z. Z. D. K. a. C. W. Liu, "Energy aware virtual machine placement scheduling in cloud computing based on ant colony optimization approach," in *In the Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, Canada, 2014.

[30] C. C. K. a. C. R. Josheph, "A novel family genetic approach for virtual machine allocation," in *In the proceedings of International conference on information and communication technologies*, 2014.

[31] N. K. GholiAbad, "RSCCGA: Resource Scheduling for Cloud Computing by Genetic Algorithm," *Computer Engineering and Intelligent Systems*, vol. 7, no. 6, pp. 14-22, 2016.

[32] S. A. H. a. F. A. Omara, "Genetic-Based Task Scheduling Algorithm in Cloud Computing Environment," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 4, pp. 550-556, 2016.

[33] J. G. G. S. a. T. Z. Jinhua Hu, "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment," in *3rd International Symposium on Parallel Architectures, Algorithms and Programming*, 2010.

[34] T. Y. K. J. K. K. a. J. S. L. Sung. Ho Jang, "The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing," *International Journal of Control and Automation*, vol. 5, no. 4, pp. 157-162, 2012.

[35] J. H. T. Z. a. G. S. Jianhua Gu, "A New Resource Scheduling Based on Genetic Algorithm in Cloud Computing Environment," *Journal of Computers*, vol. 7, no. 1, pp. 42-52, 2012.

[36] S. a. N. A. S. Mehta, "ReCon: A tool to recommend dynamic server consolidation in multi-cluster data centers," in *In the Proceedings of IEEE Netw. Oper.Manag.Symp. (NOMS)*, Salvador, Brazil, 2008.

[37] B. K. K. F. I. a. F. T. Sotomayor, "Enabling cost-effective resource leases with virtual machines," in *ACM/IEEE International Symposium on High-Performance Distributed Computing*, 2007.

[38] T. a. M. J. Huu, "Virtual resources allocation for workflow-based applications distribution of a cloud

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-13, Issue-3 (June 2023)
https://doi.org/10.31033/ijemr.13.3.24

infrastructure," in *In the proceedings of IEEE/ACM International Conference on Cluster, Cloud and Grid Computing* , Australia, 2010.

[39] Y. W. C. Z. A. a. B. Z. B. Lee, "Profit-driven service request scheduling in clouds," in *in the proceedings of 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, Australia, 2010.

[40] J. ,. D. J. L. G. J. G. M. a. G. D. Entrialgo, "Cost minimization of virtual machine allocation in public clouds considering multiple applications," in *In the proceedings of 14th International Conference, GECON*, Biarritz, 2017.

[41] Z. ,. B. J. Y. H. a. C. Y. Zhu, "SLA based dynamic virtualized resources provisioning for shared cloud data centers," in *In the proceedings of IEEE 4th International Conference on Cloud Computing*, USA, 2011.

[42] M. X. M. Z. C. G. Z. W. D. Li, "RDCM : Reliable data center multicast", in the Proceedings of IEEE INFOCOM," China, 2011.

[43] F. P. L. T. a. P. J. Farahnakian, "Hierarchical agent-based architecture for resource management in cloud data centers," in *In the proceedings of IEEE international conference on cloud computing,USA, 928-929.*, USA, 2011.

[44] X. ,. I. C. K. J. L. Z. ,. B. E. a. P. D. Meng, "Efficient resource provisioning in compute clouds via VM multiplexing," in *In the proceedings of the 7th International Conference on Autonomic Computing*, USA, 2010.

[45] N. a. S. S. Kumar, "A preference based resource allocation in cloud computing systems," in *In the proceedings of 3rd International conference on recent trends in computing*, 2015.

[46] J. W. C. L. a. D. Q. W. Lin, "A threshold-based dynamic resource allocation scheme for cloud computing. Procedia Engineering," 2011.

[47] S. S. D. H. P. ,. a. L. U. Schulte, "Cost-driven optimization of cloud resource allocation for elastic process," *International Journal of Cloud Computing*, vol. 1, no. 2, 2013.

[48] P. J. G. E. N. E. I. a. C. P. Hamid Arabnejad, "An auto-scaling cloud controller using fuzzy q-learning-implementation in," in *In Proceedings of the European Conference on Service-Oriented and Cloud Computing*, 2016.

[49] C. W. B. B. Z. L. S. Y. C. L. a. A. Y. Z. Junliang Chen, "Tradeoffs between profit and customer satisfaction for service provisioning in the cloud," in *In Proceedings of the 20th international symposium on High performance distributed computing*, 2011.

[50] C. B. a. M. L. Marios Fokaefs, "Economics-driven resource scalability on the cloud," in *In Proceedings of the 11th International Workshop on Software Engineering for Adaptive and Self-Managing Systems*, 2016.

[51] X. W. S. W. S. D. a. X. S. Hai Jin, "Towards optimized fine-grained pricing of iaas cloud platform," in *Cloud Computing, IEEE Transactions on*, 2015.

[52] F. A. D. O. S. D. a. T. L. Yousri Kouki, "A language support for cloud elasticity management," in *In Proceedings of the 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CC-Grid 2014)*, 2014.

[53] M. S. H. a. T. L. Yousri Kouki, "Delta scaling: How resources scalability/termination can be taken place economically?," in *In Proceedings of the IEEE World Congress on Services (SERVICES 2015)*, 2015.

[54] G. A. M. J. C. a. D. G. Ashutosh Pandey, "Hybrid planning for decision making in self-adaptive systems," in *In Proceedings of the 10th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO 2016)*, 2016.

[55] J. A. L. a. J. M.-A. Jose A Pascual, "Effects of reducing vms management times on elastic applications," *Journal of Grid Computing*, pp. 1-18, 2018.

[56] R. M. a. R. C. Diego Perez-Palacin, "Synthesis of adaptation plans for cloud infrastructure with hybrid cost models," in *In Proceedings of the 40th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA 2014)*, 2014.

[57] R. N. C. a. R. B. Chenhao Qu, "A reliable and cost efficient auto-scaling system for web applications using heterogeneous spot instances," *Journal of Network and Computer Applications*, vol. 65, pp. 167-180, 2016.

[58] C. C. e. al., "Live Migration of Virtual Machines," in *NSDI-05: 2nd Symposium on Networked Systems Design & Implementation*, 2005.

[59] A. K. a. K. B. N. Bobroff, "Dynamic placement of virtual machine for managing SLA violations," *IEEE*, pp. 119-128, 2007.

[60] M. H. a. E. H. B. Song, "A Novel Heuristic-based Task Selection and Allocation Framewok in Dynamic Collaborative Cloud Service Platform," in *in the proceedings of 2nd IEEE Conference on Cloud Computing Technology and Science*, 2010..

[61] J. B. H. Y. a. Y. C. Z. Zhu, "SLA Based Dynamic Virtualized Resources Provisioning for Shared Cloud Data Centers," in *in the proceedings of 4th international Conference on Cloud Computing*, 2011.

[62] a. A. S. M. Mishra, "On Theory of VM Placement: Anomalies in Existing Methodologies and Their

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-13, Issue-3 (June 2023)
https://doi.org/10.31033/ijemr.13.3.24

Migration using a Novel Vector Based Approach," in *in the proceedings of 4th international Conference on Cloud Computing*, 2011.

[63] Z. S. X. G. S. S. a. J. W. H. Nguyen, "AGILE: elastic distributed resource scaling for Infrastructure-as-a-Service".

[64] W. D. A. Strunk, "Does Live Migration of Virtual Machine Cost Energy?," in *in the proceedings of 27th International Conference on Advanced Information Networking and Applications*, 2013.

[65] a. H. S. L. Chen, "Consolidating Complementary VMs with Spatial/Temporal-awareness in Cloud Datacenters," in *in the proceedings of IEEE INFOCOM - IEEE Conference on Computer Communications*, IEEE.

[66] a. N. W. P. A. Mosa, "Optimizing virtual machine placement for energy and SLA in clouds using utility functions," *Journal of Cloud Computing: Advances, Systems, and Applications, pp. 1-17, 2016.*, pp. 1-26, 2016.