A Self-Optimization Based Virtual Machine Scheduling to Workloads in Cloud Computing Environment

Bhupesh Kumar Dewangan, Amit Agarwal, Venkatadri M, Ashutosh Pasricha

***Abstract*: Energy consumption in the cloud computing plays a vital role in operating cost to the service provider and cloud user. The cloud is scalable and it can provide the access as per demand, due to this resource access requests are increasing, and submitted to the server. To manage all request, scheduling is the solution to assign request with the quality of service. To avoid high operating cost, resource scheduling needs to be energy-aware. In this paper, energy-aware resource scheduling in the cloud is proposing. Total resource utilization of each resource has been calculating and energy is optimizing through antlion optimization algorithm to avoid high consumption of power. The resource is identify with its best utilization value and assign to submit workloads as per priority basis. The experimental results of the proposed work are analyzing with existing autonomic frameworks, and it is observed that proposed work is performing utmost.**

Index Term: Energy, Resource Cost, Execution Time, Performance, SLA Violation Rate

I. INTRODUCTION

To meet each changing business, require association need to contribute time and spending plan to scale-up IT framework, for example, equipment, programming, and administrations. In any case, if out-premises IT foundation, the scaling procedure can be moderate and associations are every now and again unfit to accomplish ideal usage of IT framework.

Distributed computing is a worldview, that giving figuring over the web. A distributed computing administration comprise of the profoundly altered virtual datacenter, which gives different programming, equipment and data assets for helpful required. Association can interface with the cloud and utilize the accessible asset on pay per utilize premise. The assets can be opt by asset administration strategies.

The formal resource management, there are two stakeholders, one side is cloud user and the other side is cloud service provider (CSP) and the requirements of these two are the computing resources. The cloud user would expect resources form cloud service provider for computing applications, the CSP is allocating resource with respect to maximize the revenue and resource utilization, so the resources become the point of context for the CSP.

On the other hand, when these resources are allocate to the cloud user, the user will consume the resource with an objective to maximize the performance in terms of minimizing the finish time and cost. In this context, the resource management is how these two objects have met that maximization of resource utilization and the minimization of time and cost. The resources can be categorize into the following:

A. Power Resources

This is the indirect resources, which are consumes by the user. The datacenter is consuming the very large amount of power [1], off-course how these powers can be manage, is a very typical task, and therefore reducing the power consumption is the objective of the research.

B. Network Resources

Network resources are how the machines are connects with each other within the datacenter. In this, high bandwidth is expect so that communication should not be fail.

C. Compute Resources

It contains the collection of the physical machine, like processors, memory, network interface, input and output devices, and total computational capacity of the cloud to use the resources.

Resource provisioning can be also categorize as follows:

- Provisioning of assets to third parties on rented, utilization based for an open cloud.
- Private foundation kept up and used by singular association for private cloud.
- Leasing open cloud assets to a third-party private cloud will be cloud blasting for the hybrid cloud.
- Resources contributed by multiple individuals or organizations for community cloud.

Revised Manuscript Received on December 22, 2018.
Bhupesh Kumar Dewangan, Department of Informatics, University of Petroleum and Energy Studies, Dehradun, India.
Amit Agarwal, Department of Cloud Computing and Virtualization, University of Petroleum and Energy Studies, Dehradun, India.
Venkatadri M., Department of CSE, Amity University, Madhya Pradesh, India.
Ashutosh Pasricha, Head Account, Schlumberger: Oil Field, New Delhi, India.

With these, CSP provides the resources to the cloud user. The cloud users are cloud enterprise user, who consume the more power for accessing the resources, where individual users are consuming the less power, but it is in large in numbers so, power consumption is more if cumulative add and commercial cloud user, use the open source application in large volume, so the power consumption [2] is more. In all, the objective of resource management is to allocating computing, storage, networking, and energy resources to meet the performance objectives.

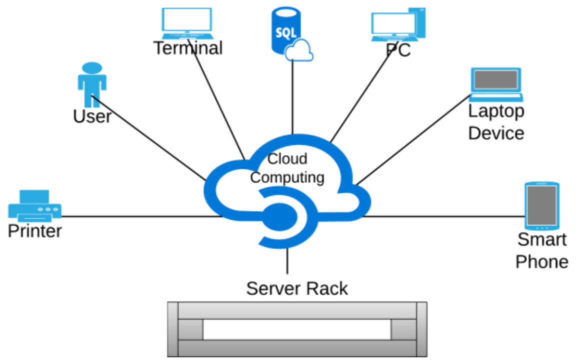

Fig. 1: Cloud Computing Model

The focus of this paper is to do state of art survey of different resource scheduling algorithm, which based on power consumption and energy efficiency analysis, with its findings, and develop energy-aware scheduling model to optimize energy of resources. The problems of the current mechanism have been analyze and modeled the scheduling architecture, and observation of the proposed model is performing better as compared existing models.

II. RELATED WORK

The state of art survey has been taken carried out on the relevant area and scope, and it is compared with some performance metric. The recent developments on energy-aware resource scheduling frameworks are:

In [3] demonstrated that the execution of the proposed Tri Objective Resource Optimization Algorithm (TROA) is the best calculation in view of the parameters of least or low power utilization of resources with the high usage, least cost, better makespan and less CO2 outflow prompts the idea of condition manageability. Where, authors [4] have presented a decentralized engineering of the dynamic resource management framework for Cloud server farms. One of the heuristics prompts critical reducing of the essentialness use by a Cloud server cultivate by 83% interestingly with a non-control careful structure and by 66%, conversely with a system, that applies just Dynamic Voltage and Frequency Scaling (DVFS) framework yet does not alter segment of VMs in run-time.

In [5], introduce two vitality cognizant errand combination heuristics, which mean to boost asset usage and unequivocally consider both dynamic and sit out of gear vitality utilization. The proposed movement [6] is general and goes past the present condition of the quality by limiting both the number of relocations required for combination and resource utilization in a solitary calculation with an arrangement of substantial disparities and conditions. The paper [7], about distributed computing, mainly focuses on performance, cost and execution time of resource scheduling approaches. In [8] propose a booking estimation for the cloud datacenter with a dynamic voltage repeat scaling methodology. Their planning calculation can productively build resource usage; consequently, it can diminish the resource utilization for executing employments.

In [9] examine the connection between foundation segments and power utilization of the distributed computing condition and talk about the coordinating of assignment composes and segment control modification technique, and afterward, we exhibit an asset booking calculation of Cloud Computing in light of vitality effective enhancement strategies. In this examination [10], they proposed EnaCloud, which is a vitality sparing application live position approach for the vast size of cloud stage.

This paper [11]goes for laying out, realizing and surveying a Green Scheduling Calculation fusing a neural framework marker for streamlining server control usage in Cloud enrolling. The outcomes demonstrate that the PP20 mode can set aside to 46.3% of power use with a drop rate of 0.03% on one load take after, and a drop rate of 0.12% with a power diminishment rate of 46.7% on the other. Authors [12] give an essentialness capable intense offloading and resource booking (editors) plan to reduce imperativeness use and abridge application-completing time. This work [13] accentuations the part of correspondence texture and presents a planning arrangement, named e-STAB, which considers movement necessities of cloud applications giving vitality effective occupation designation and activity stack adjusting in server farm systems. Researchers [14] examination and the show-cantered extents of idealize online deterministic counting for the single VM movement and dynamic VM issues. What's more, we propose novel adjust capable heuristics for dynamic relationship of VMs in context of an examination of genuine information from the preferred standpoint use by VMs. Cloud RAN [15] is another incorporated worldview in view of virtualization innovation that has developed as a promising design and proficiently tends to such issues.

In [16] play out an extensive examination of a framework supporting the distributed computing worldview with respect to vitality proficiency. Writers [17] propose another parallel bi-target half-and-half hereditary calculation that considers makespan, as well as vitality utilization. Authors [18] acquaint an interim number hypothesis with depict the vulnerability of the processing condition and a planning design to alleviate the effect of vulnerability on the errand booking quality for a cloud server farm.

In [19] proposes a model for evaluating the utilization of each virtual machine without committed estimation equipment.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Soccer [20] design energy-aware cloud management, which optimized energy by using the self-optimization technique. Resources are scheduled based on the energy threshold value and fulfil the SLA and QoS. They have opted four steps to schedule the resource: 1) monitoring, 2) analysing, 3) planning, and 4) execution. The model is implements and test in real cloud environment of Thapar University. Chopper [21] enhanced the continuation with the soccer model for energy optimization. In this paper, the authors applied self-characteristics [22] to solve the resource scheduling problems. This model aims to autonomic techniques through self_healing, self_optimization, self_protection, and self_configuration [23]. The autonomic characteristics are implement in the real cloud by using fuzzy and machine learning concepts. The workloads are categorize in different clusters according to its type [24].

III. PROBLEM STATEMENT

In the above study, it is observe that the cloud resources are managing through different optimization techniques applies to cost, SLA violation rate, resource utilization and energy consumption. In all, efficient resource management depends upon its cost, resource utilization, and QoS fulfilments. The following observation is carry out of a study on existing resource scheduling algorithms:

SLA violation rate is high, due to this customer, trust is decreasing, Energy consumption is increasing due to the scalability of IT infrastructures like data Centre and servers, Lack of fault recovery technique, therefore resource utilization is less, The failure rate of resources increases the execution time and cost, A large number of data Centre and server required an intelligent resource management system, Lacking QoS, that decrease the user satisfaction, and Protection from malicious workloads is missing.

The findings are solving in this paper by applying the autonomic resource management strategy by using the antlion optimization algorithm. The novelty of this work is utilizing the antlion in cloud resource scheduling which is successfully implement for network optimization. This work optimizes cost, energy and resource utilization, which will be explaining in results and analysis section. Table I is mentioned the multi-objectives consideration of the above survey of different energy-aware models.

Table I: Comparative Analysis of Existing Frameworks

	Analysis of different models	Performance Metrics						
		C	MK	RU	PM	SLA	QoS	ET
Energy-Aware Resource Scheduling Models	[3]	√	√	√	√			
	[4]			√		√		√
	[5]			√				√
	[7]	√			√			
	[8]			√	√			√
	[9]			√				
	[10]			√				√
	[11]	√		√				
	[12]	√			√			√
	[15]			√				√
	[13]	√		√	√			
	[14]			√		√		
	[16]	√		√				
	[17]		√	√				√
	[18]			√				√
	[19]	√		√	√			
	[20]				√	√	√	√
	[21]	√		√	√	√	√	
	Proposed	√		√	√	√	√	√

AT = Autonomic, RU = Resource Utilization, PM = Performance, Et = Execution Time, MK = Makespan, C = Cost, , QoS = Quality of service

IV. METHODOLOGY

The proposed framework is:

A. *Workload Submission*

Cloud user submit the request to the service provider for accessing the different resources from the cloud. This request is presenting as workloads. In this paper, workloads are classifying into four categories.

B. *Classification and Clustering of Workloads*

The workloads are submitting to the task manager, and it will be classify into different groups according to its type. The K-Mean clustering algorithms are applying for generating clusters. To generate workload clusters, the centroid value of each cluster are identifying and group the workloads accordingly. There are four clusters generated namely C1, C2, C3, and C4. All these clusters are leading web service demand, storage demand, computing demand, and performance demand.

C. *Architecture*

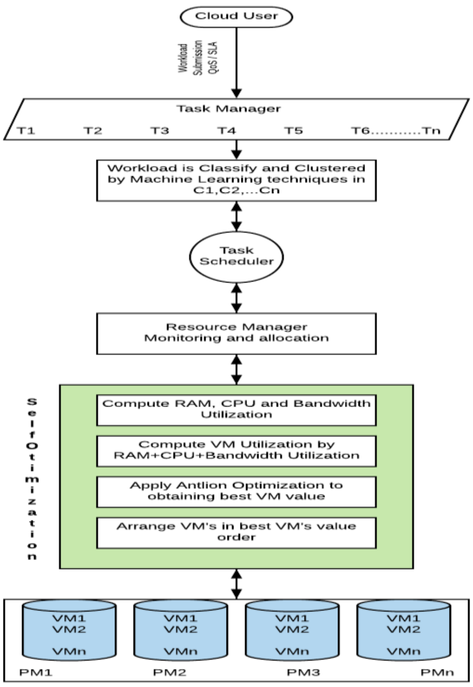

Fig. 2: Architecture of Proposed Model

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

D. *Ordering of Workloads*

The workloads are arrange according to its priority based on execution time and cost and ready to assign the demand resource. The best suitable resources need to be identify by the service provider to assign the different workloads.

E. *Identification of Best VM's*

In this paper for the best VM, selection used improved antlion optimization (ALO). After the centroid positions are finished in the grouping procedure, consider the hubs which are at the closest separation and furthermore the following closest separation from the centroid by utilizing ALO optimization. In this method, the initialization is making by using a random selection of antlions and ants (VM). After that, the multi-objective (resource utilization, energy, and cost) is calculated between each VMs, from that the fitness value of ant and ant lion is calculated and ant lion fitness value sorted to find the best ant lion is named as elite antlion (best VM). The VM value is identify by equation (1).

$$VM_U = VM.CPU_U + VM.BW_U + VM.RAM_U \quad (1)$$

Therefore, the first objective-function is,

$$F_1 = \max\{VM_u\} \quad (2)$$

F. *Scheduling*

In this, 100 workloads submitted to the task manager and it is classified and clustered by k-mean and grouped by types and sorted according to its priority based on time and cost. On the service provider side, 20 VM's are generate for allocation. The resources are identify and sort according to its best VM value, which is obtain by the antlion optimization algorithm. The resources are assign to the workloads by using round robin resource scheduling algorithm.

G. *Performance analysis*

For efficiency evaluation of the proposed method, performance evaluation is the key point. Here, after scheduling of VM's, we have computed the following performance metrics and compared with existing models.

Energy consumption

$$S = k * \max + (1-k) * \max * VM_u \quad (3)$$

Second objective-function is,

$$F_2 = \min\{s\} \quad (4)$$

Therefore, cost on resource is,

$$Cost\ C = c * S \quad (5)$$

Where c is the cost of 1 kW power.

The third objective function is,

$$F_3 = \min\{C\} \quad (6)$$

SLA violation rate

$$SVR = failurerate * \sum_{i=1}^{n} w_i \quad (7)$$

In continuation, to find the best VM values, resource utilization value of each VM is to be calculated by algorithm1.

Algorithm1: Resource Utilization

Step 1. Start
Step 2. Initialize maxiteration = 100, and Antlion Algo.
Step 3. if(i==Population)
Step 4. compute cpu_u =(Vm_Mips.get(i)/Pm_Mips);
Step 5. bw_u =(Vm_BandWidth.get(i)/Pm_Bw);
Step 6. ram_u =(Vm_Ram.get(i)/Pm_Ram);
Step 7. tot=(cpu_u+bw_u+ram_u);
Step 8. End if
Step 9. End

Energy consumption is one of the key feature of this research work, to obtain the best VM values, that is computing by algorithm2.

Algorithm2: Energy Consumption

Step 10. Start
Step 11. Initialize totalEnergy=0.0, max=1, k=0.5, and Antlion Algo.
Step 12. getPower(utilization)
Step 13. Compute s=k*max+(1-k)*max*utilization
Step 14. End

The third key feature of this research is resource cost, and that will be obtaining through algorith3.

Algorithm3: Resource Cost

Step 15. Start
Step 16. Initialize cost=600 and Antlion Algo.
Step 17. Compute cost=ini_cost*(energy);
Step 18. End

To achieve the best VM selection, all three components added and scheduled the VM's to the workloads submitted by the cloud user.

V. SIMULATION RESULTS AND ANALYSIS

The proposed work is simulating in cloudsim, the workloads submitted by cloud user sorted according to user priority. For VM's, best VM values are obtained by computing the energy consumption, resource cost, and SLA violation rate. According to VM's best values, it is sorted on the order and allocating to the workloads submitted by cloud user as per the user priority.

A. Energy Consumption Analysis

The energy consumption of each VM's are computing through equation (1), 20 VM's initializes for 200 workloads, and the consumption observation is 9.7 KWh, for the same workload, it is obtained 12 and 10.5 KWh for Chopper and Soccer. In continuous, workloads are increasing as 400, 600, 800, and 1000, consumption rate for proposed work are 13.5, 27 and 37.1 KWh, where consumption rate for other two existing models are 15, 26.3, 32 and 41.4 KWh and 14.7, 23.1, 29.4, and 38.9 KWh measured. Through this analysis, the results show that the proposed model consume less energy as compared to the other two existing models. The comparison is presenting in Fig. 3.

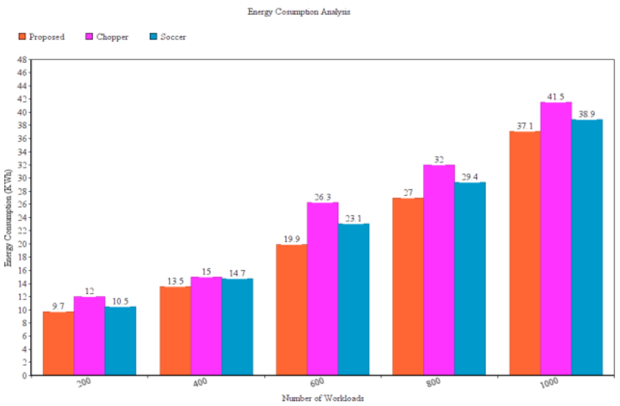

Fig. 3: Energy Consumption Ratio

B. Resource Cost Analysis

The resource cost of each VM's are computing through equation (5). 20 VM's initializes for 10, 20, 30, 40, 50, and 60 workloads, and the cost of each resource observation is $18, $20, $25, $38, $40, and $55 for same workloads it is obtained $12, $15, $26.3, $32 and $41.5 for Chopper and $25, $30, $40, $50, $60, and $70 for Soccer. While the workloads are increasing, simultaneously the resource costs are also increasing for each model. In Fig. 4.,it is clearly presenting that, proposed work compute less resource cost as comparing the other two existing models. The average cost of resources of the proposed model is less than the average cost of other existing models.

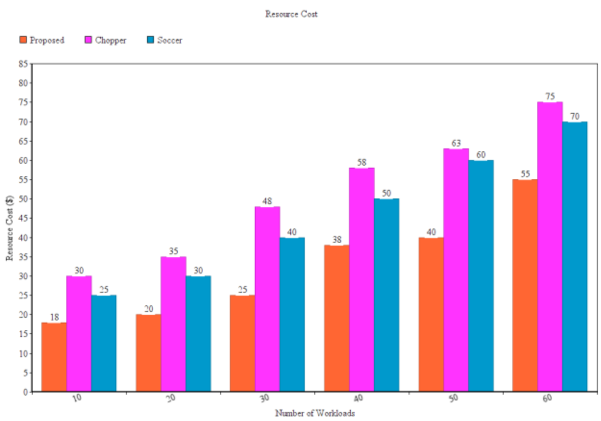

Fig. 4: Resource Cost Analysis

C. SLA Violation Rate and Energy Efficiency Analysis

SLA violation rate of each VM is pressing in Fig. 5., it is observing that SLA violation rates are increasing, when workloads are increasing, in continuous in Fig. 6., the energy efficiency of VM's are presenting, and it is observing that energy efficiency decreasing when workloads are increasing.

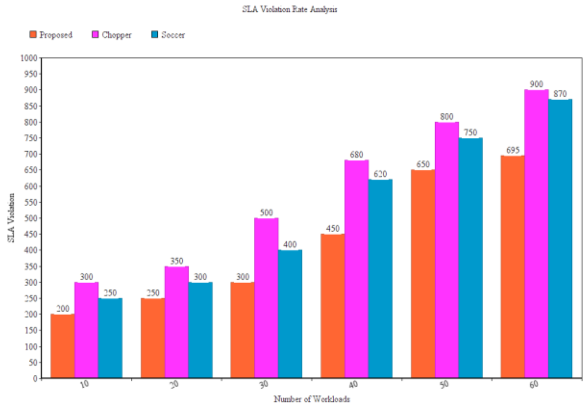

Fig. 5: SLA Violation Rate Analysis

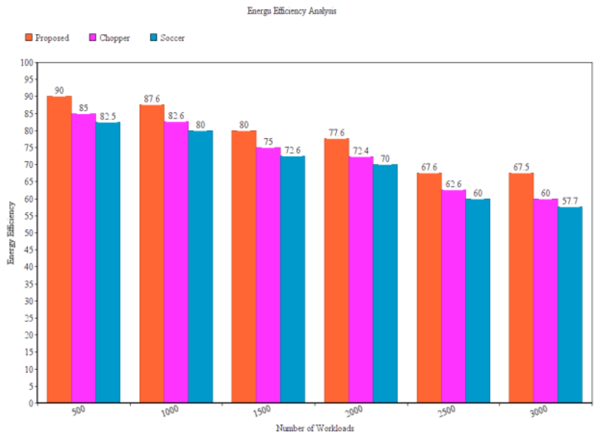

Fig. 6: Energy Efficiency Analysis

D. Performance Analysis

The average values of each performance metrics are calculating and presenting it in the graph, which is plotting in Fig. 7. The SLA violation rate, resource cost, energy consumption, and energy analysis observation is plotting. The average values of SLA violation rates are, 42.416, 58.833, and 53.166 for proposed and other existing models, which is evidence of better performance against SLA violation rate. The average resource cost is $32.6, $51.5, and $45.83 for proposed and other two existing models, which shows that the proposed model is observing less cost. The energy consumption and energy efficiency rates are obtaining and plotting in figure 6.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Which shows that the energy consumption of the proposed model is less and efficiency is greater than the other two models. The objective satisfying the requirements of energy optimization of the proposed model.

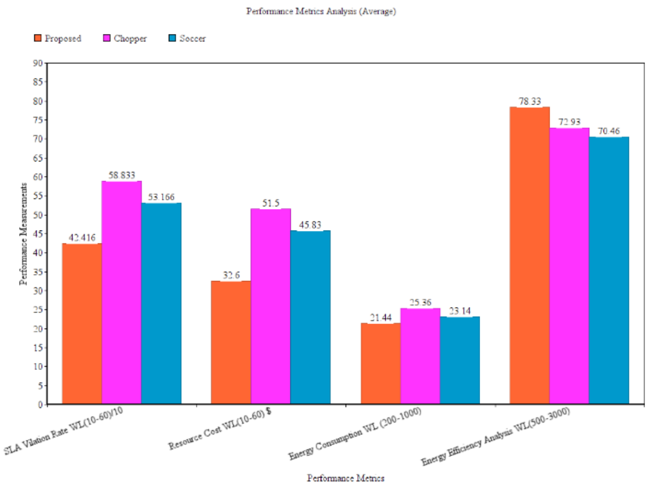

Fig. 7: Overall Performance Analysis (Average)

VI. CONCLUSION AND FUTURE SCOPE

In this paper, the proposed model displayed to enhance energy efficiency minimizes the resource cost and SLA violation rate. This model is simulates in the cloud environment. The outcome may shift when it will be execute in the real cloud environment. In this, it is introduced a point by point stream of proposed model that, how to apply the proposed structure to enhance the resource utilization, minimizes the SLA violation rate, execution time, and cost to boost the execution for scheduling in the cloud. The test comes about to demonstrate that the proposed model, that the different performance metrics are evaluated and test results examined that the proposed model is outperforming. The restriction of this work is, no fault tolerant, and protection mechanism is apply to find the faulty VM's and to reject the malicious workloads, this will be considered in future work.

ACKNOWLEDGMENT

The authors will like to thank anonymous reviewers for their constructive comments to improve the quality of this paper.

REFERENCES

1. Saxena, P., Sharma, S. K., & Verma, N. K. (2017). An Analysis of Energy Efficiency& Traffic Using Cloud Computing In Portable Devices. i-manager's Journal on Cloud Computing, 4(1), 1-6.
2. Naseera, S., & Jyotheeswai, P. (2015). A survey on energy aware job scheduling algorithms in cloud environment. i-manager's Journal on Cloud Computing, 3(1), 30-38.
3. Y. a. G. Alex. (2017). Comparison of Resource Optimization Algorithms in Cloud Computing. International Journal of Pure and Applied Mathematics, pp. 847-854.
4. Beloglazov, A., & Buyya, R. (2010, May). Energy efficient resource management in virtualized cloud data centers. In Proceedings of the 2010 10th IEEE/ACM international conference on cluster, cloud and grid computing pp. 826-831.
5. Lee, Y. C., & Zomaya. (2012). Energy efficient utilization of resources in cloud computing systems. The Journal of Supercomputing, 60(2), 268-280.
6. Ghribi, C., Hadji, M., & Zeghlache, D. (2013). Energy efficient vm scheduling for cloud data centers: Exact allocation and migration algorithms. In Cluster, Cloud and Grid Computing (CCGrid), 13th IEEE/ACM International Symposium on pp. 671-678.
7. Berl, A., Gelenbe, E., Di Girolamo, M., Giuliani, G., De Meer, H., Dang, M. Q., & Pentikousis, K. (2010). Energy-efficient cloud computing. The computer journal, 53(7), 1045-1051.
8. Wu, C. M., Chang, R. S., & Chan, H. Y. (2014). A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters. Future Generation Computer Systems, 37, 141-147.
9. Luo, L., Wu, W., Di, D., Zhang, F., Yan, Y., & Mao, Y. (2012). A resource scheduling algorithm of cloud computing based on energy efficient optimization methods. In Green Computing Conference (IGCC), 2012 International pp. 1-6. IEEE.
10. Li, B., Li, J., Huai, J., Wo, T., Li, Q., & Zhong, L. (2009). Enacloud: An energy-saving application live placement approach for cloud computing environments. In 2009 IEEE International Conference on Cloud Computing pp. 17-24. IEEE.
11. Duy, T. V. T., Sato, Y., & Inoguchi, Y. (2010). Performance evaluation of a green scheduling algorithm for energy savings in cloud computing. In Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), IEEE International Symposium on pp. 1-8.
12. Guo, S., Xiao, B., Yang, Y., & Yang, Y. (2016). Energy-efficient dynamic offloading and resource scheduling in mobile cloud computing. In INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, IEEE pp. 1-9.
13. Kliazovich, D., Arzo, S. T., Granelli, F., Bouvry, P., & Khan, S. U. (2013, August). e-STAB: Energy-efficient scheduling for cloud computing applications with traffic load balancing. In Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing pp. 7-13.
14. Beloglazov, A., & Buyya, R. (2012). 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, 24(13), pp. 1397-1420.
15. Pompili, D., Hajisami, A., & Tran, T. X. (2016). Elastic resource utilization framework for high capacity and energy efficiency in cloud RAN. IEEE Communications Magazine, 54(1), pp. 26-32.
16. Mastelic, T., Oleksiak, A., Claussen, H., Brandic, I., Pierson, J. M., & Vasilakos, A. V. (2015). Cloud computing: Survey on energy efficiency. Acm computing surveys (csur), 47(2), pp. 33-40.
17. Mezmaz, M., Melab, N., Kessaci, Y., Lee, Y. C., Talbi, E. G., Zomaya, A. Y., & Tuyttens, D. (2011). A parallel bi-objective hybrid metaheuristic for energy-aware scheduling for cloud computing systems. Journal of Parallel and Distributed Computing, 71(11), pp. 1497-1508.
18. Chen, H., Zhu, X., Guo, H., Zhu, J., Qin, X., & Wu, J. (2015). Towards energy-efficient scheduling for real-time tasks under uncertain cloud computing environment. Journal of Systems and Software, 99, pp. 20-35.
19. Kim, N., Cho, J., & Seo, E. (2014). Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems. Future Generation Computer Systems, 32, pp. 128-137.
20. Singh, S., Chana, I., Singh, M., & Buyya, R. (2016). SOCCER: self-optimization of energy-efficient cloud resources. Cluster Computing, 19(4), pp. 1787-1800.
21. Gill, S. S., Chana, I., Singh, M., & Buyya, R. (2017). CHOPPER: an intelligent QoS-aware autonomic resource management approach for cloud computing. Cluster Computing, pp. 1-39.
22. Bhupesh Kumar Dewangan, Amit Agarwal, Venkatadri M and Ashutosh Pasricha. (2018). Resource Scheduling in Cloud: A Comparative Study. International Journal of Computer Sciences and Engineering. 6(8). 167-173.
23. Bhupesh Kumar Dewangan and Anurag Jain. (2018). Cloud based Platform Independent Practical Evaluation Model for Technical University. International Journal of Scientific Research in Computer Science Applications and Management Studies. 7(4). 1-3.
24. Bhupesh Kumar Dewangan, Amit Agarwal, Venkatadri M and Ashutosh Pasricha. (2016). Credential and Security Issues of Cloud Service Models. 2nd IEEE International Conference on Next Generation Computing Technologies (NGCT). 14-16.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication