



Job Scheduling for Energy Efficiency Using Artificial Bee Colony through Virtualization

Kavita Arjun Sultanpure^{1*} Lakkam Siva Shanker Reddy¹

¹Koneru Lakshmaiah Education Foundation, India

* Corresponding author's Email: kavita.sultanpure@gmail.com

Abstract: Virtualization is used for saving energy, cost and is run for multiple applications using various operating systems on the same server at the same time. Thus, it increases the CPU utilization, efficiency and flexibility of the computer hardware. This paper presents the concept of job scheduling for energy efficiency using Artificial Bee Colony (ABC) by utilizing virtual migration concept in addition to broadcasting of requests. Due to the fact that the virtual migration concept helps a lot in management of the scheduling in cloud computing, so, the validations of job scheduling and service level agreement (SLA) violations with and without virtual migrations are opted in this work. In the initial phase, when server gets the requests for task execution, then the sub-servers (S_1, S_2, \dots, S_n), executes that requests (R_1, R_2, \dots, R_n) irrespective of the main server. When the load at sub-servers increased, then the usage of virtual migration machines is executed, so that the jobs can be implemented equally. Two types of energy has been obtained in the work, first, which is utilized for the execution of the task and second, when virtual migrations creates energy reduction of overall energy ($E = E_1, E_2, \dots, E_n$). So, for given interval of time, the creation energy for virtual migrations came out to be less than the allocation energy of virtual migrations for jobs, to have better accuracy. At last, the comparison of proposed work, with existing work is provided. In the existing work, we have considered the papers that evaluate SLA violation and job scheduling by using neural network (NN) and genetic algorithm (GA). The SLA violation of ABC algorithm with respect to GA is reduced by 10% whereas SLA violation of ABC algorithm with respect to NN is reduced by 15%. Job scheduling of ABC algorithm with respect to NN is improved by 15%.

Keywords: Cloud computing, Artificial bee colony, Broadcasting, scheduling, Service level agreement, Virtualization.

1. Introduction

Cloud computing allows [1] users to access large amount of data and various computing resources according to the use of customers. On the other hand, if we talk about use of virtualization in cloud computing it will works on multiple operating systems and applications in parallel. Hence virtualization increases the efficiency and reduces hardware cost.

For cloud applications the business sector has manifested an eagerness for the cloud suppliers. Similarly, owning server farms and holding the responsibility of their concerns demands vital means

of efforts and funds for organizations. In fact, providers wish to take benefits of cloud technology that leads to possess a prime focus on their business as the preference for organizations. Then again, by executing the higher count of recruitments, suppliers seek for the tactics to extend their overall revenues, though it depreciates the functional costs. Energy usage by the servers results in higher preparation cost. Moreover, limited spending policy marks the requirement for suppliers to build productive cloud frameworks which utilize the computational strengths of the clouds by reducing the energy usage and environmental impacts [2]. This indicates that it is necessary for server farms to enhance their energy usage.

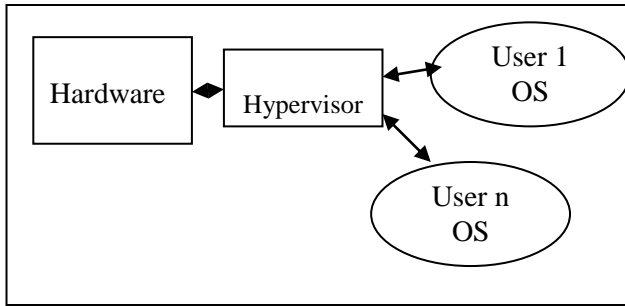


Figure.1 Architecture of virtualization

In a simple word one can say that virtualization [3] is way in which the resources of computers are divided into multiple execution environments. This is achieved using technologies like hardware and time sharing. Architecture of virtualization is explained in fig 1.

Using virtualization operating system (OS) has the overall control on hardware. This is done by using hypervisor and virtual machine manager (VMM). The hypervisor or VMM is a software layer that allows the end user to run more operating system simultaneously over on a single computer system. The difference between physical and virtual machine is only in terms of OS, underlying hardware etc. In physical machine it has less number of operating systems to access but virtual machine OS works independently on hardware. In physical machine OS and hardware are used in conjunction with each other whereas in virtual machine OS is not lying on hardware. Hypervisor used in virtualization is software which is used to handle all the activities of VMM. Hypervisor is again divided into two parts, Type 1 Hypervisor and Type 2 Hypervisor. Type 1 Hypervisor does not consists of any OS because they are used on the bare system. Type 2 hypervisor is software that emulates the devices with which a system normally interacts.

In this proposed work, the virtualization concept [4] simultaneously with broadcasting serves to perform the energy reduction. Suppose that the cloud system comprises of a set N of n machines. N can be expressed by:

$$N = \{Pm0, Pm1, Pm2 \dots Pmn\} \quad (1)$$

Here, N signifies the total number of physical machines that are represented by the $\{Pm0, Pm1, Pm2 \dots Pmn\}$.

An individual physical machine can uphold multiple combinations of CPU, memory and network cards. These complex physical machines in cloud data centres possess distinct capabilities and speed. Utilizing virtualization technology, a set B of

b virtual machines [5] operate on every physical machine to balance the load. B can be expressed by:

$$B = \{Vb0, Vb1, Vb2 \dots Vbb\} \quad (2)$$

Here, B signifies the total number of virtual machines that are represented by the $\{Vb0, Vb1, Vb2 \dots Vbb\}$.

In cloud data centre the virtual machines [6] on a physical machine can be revived, stopped and migrated to the other physical machines for load balancing [7]. Because of the heterogeneity, consider the hardware resource heterogeneity (including CPU, memory, and network cards) when placing VMs. Here, rcd is used to denote the computing capability of a dedicated reference server d. rcd can be measured in Millions of Instructions Per Second (MIPS). For an application p, it needs rpcvd of the computing capability of the dedicated reference VM dv. Thus, if the application p runs in another VM Sv, the computing proficiency rpcsv is given by

$$\frac{r_{pcvd} \cdot r_{cs}}{r_{cd}} \quad (3)$$

Here, rcs depicts the calculating proficiency of the server h. rcs can also be measured by MIPS.

In order to describe the difference between VMs' computing capabilities [8], usage of K×L Computing Capability Matrix (CCM) in which C is used to indicate the computing capabilities of K types of applications running on L types of physical servers through Eq. (3).

The requests in cloud network are managed through sub-server. Every request is supposed to be placed within the coverage set whereby cloud provider is requested for the services by the user in order to execute jobs. Generally, the system comprises two main constituents, specifically, User and Provider. The cloud user submits the job (or set of jobs) to the cloud provider to be performed, normally with a designated QoS (quality of service) requirement. The submission equation is presented as:

$$Job(QoS) \rightarrow DLp \quad (4)$$

Eq. (4) must be between the deadlines (DLp) for the user jobs. A Provider acquires the user's jobs, performs them at a particular data center, and later delivers the conclusion back to the users. Promptly, VM machines are utilized when the load at the sub-servers get increased. Here, the prime object of the work is to diminish the energy, but, some energy

can be utilized for task accomplishment. So, for assigned time period, the production energy for VM must be less than the allocation energy of VM for jobs.

Several QoS chosen for the accomplishment of jobs:

$$ssReqTime_{user} = \sum_{p=1}^{no.of\ jobs} Time_p \quad (5)$$

Where, $ssReqTime_{user}$ represents the time of request from user to server and $Time_p$ represents the processor time.

$$data\ ReqMemory_{user} = \sum_{p=1}^{no\ of\ jobs} RAM_p \quad (6)$$

$data\ ReqMemory_{user}$ represents the memory request by the users and RAM_p represents the processor RAM, which is used in job scheduling.

$$ReqStorage_{user} = \sum_{p=1}^{no\ of\ jobs} Storage_p \quad (7)$$

$ReqStorage_{user}$ signifies the storage request and $Storage_p$ represents the storage of processor.

The centre model which is proposed in the work consists of diverse PM. Every individual PM encloses multicore processors and they are described by the configuration presented as follows:

$$Storage_{PM} = Size\ of\ the\ storage\ in\ PM_p \quad (8)$$

$Storage_{PM}$ represents the storage of physical machine that is the total size of storage presented in physical machine

$$RAM_{PM} = RAM\ size\ in\ PM_p \quad (9)$$

RAM_{PM} Represents the RAM of physical machine which is the RAM size of physical machine.

Each PM can be allotted one or more VM. Each VM configuration is described as follows:

$$Storage_{VM} = Size\ of\ the\ storage\ in\ VM_p \quad (10)$$

$Storage_{VM}$ Represents the total storage size in virtual machine.

$$RAM_{VM} = RAM\ size\ in\ VM_p \quad (11)$$

RAM_{VM} Represents the size of RAM in virtual machine.

The proposed work has been already done by different authors in the field of VM migration. In [9] authors used GA to reduce SLA violation in job

scheduling technique. But we will find that the SLA violation has not acceptable in case of large number of jobs. This is due to the section of objective function of GA. Due to the objective function, the execution rate is slow this will increase the SLA violation. In [10] authors used artificial neural network (ANN) to allocate jobs. But we will find that they used insufficient properties of processors to train the artificial neural network. So, classification rate of ANN has been decreased therefore in case of large number of jobs SLA violation will be maximum and the probability of VM migration is high.

To resolve above problems of used techniques such as GA and ANN, we initiated load balancing concept with optimization technique to reduce the SLA violation in job scheduling technique. We have used ABC algorithm along with load balancing concept to maintain the SLA violation and execution of job by using the optimal objective function.

The organization of the paper is as follows a small description of cloud computing along with the role of virtualization in cloud computing is discussed in section 1. The description of energy utilization scenario in VM is discussed in section 2. In section 3, the work performed by different authors mainly in energy preservation is presented. Section 4 and section 5 presents the proposed model, algorithms and Result & discussion. At last, the conclusion is presented as per the results and discussions in section 6 followed by references.

2. Energy efficiency scenario in virtual machine

In cloud data centres [11], the energy utilization of physical machines is mostly determined by processor, memory, disk storage, and network interface controllers. The energy model as represented in this study is based on the fact that processor utilization holds a linear relationship with energy utilization. It can be stated that the CPU utilizes the essential part of the energy. The presumption has been employed by various other works. Consider that the physical machine functions in three different modes in the cloud data centres, namely, the idle mode, the active mode, and the sleep mode. Countless researches have revealed that an idle server utilizes approximately 70% of the energy utilized by the server streaming at the full speed on average. We should turn the idle servers into the sleep mode in order to diminish the entire energy utilization [12].

$$P(c) = kP \cdot max + (1 - kP) \cdot max \cdot c \quad (12)$$

In Eq. (12), P.max is the energy utilization at the peak load and k is the fraction of power consumed by the idle server, c is the CPU utilization [13].

3. Related work

There are several techniques that are already proposed by different researchers in their research work primarily on energy preservation. The comparison of the existing techniques has been done and presented in the tabular form for better understanding regarding the findings and limitations of their work. Table 1 shows the related work.

4. Proposed model

In some cloud server [15], scheduling and sustenance of the cloud server are extremely challenging and complicated jobs. There are various jobs at each cloud server that requires constant attention. The cloud server does not interact straight for the accomplishment of each job; instead, it is performed by sub servers [16].

Consider the following assumptions for proposed work:

- System has single main server
- Requests are managed through the sub-servers.
- Requests are managed at separate or same time by utilizing VM.
- Requests are required to manage inside the coverage area.
- Energy for the conception of VM must be lower than the energy required for the allocations of jobs.

Fig. 2 reveals the functioning of the proposed work.

Every single request is managed by the sub-server. All requests must lie within the frequency range, which is known as coverage set. The framework includes two basic fragments: User and Provider. The cloud buyer exhibits plenty of requests to the cloud supplier to be accomplished.

A provider satisfies the buyer's requests, accomplishes them at a specific sub-server and later broadcasts the results back to the buyers. Instantly, if the load at sub-servers gets increased then VM machines are utilized. Here, the main aim of this work is to reduce the energy, but some energy can be utilized for execution of the job. Therefore, for given time period, the creation energy for VM must be lower than the energy required for the allocation of VM for jobs.

The proposed algorithm initially finds the request handling capability that defines which job is

going to be performed and by which sub-server. For an individual job, there may be more than one particular sub-server, whereas there is also a probability that for the accomplishment of a job, physical machine or sub-server may not be granted. Therefore, requesting jobs must lie in the coverage area as defined in this section.

for $m = 1$: total physical machines
 for $w = 1$: total virtual machines
 if $(\sqrt{((X_{am} - X_{aw})^2 + (Y_{am} - Y_{aw})^2)}) \leq Cov - limit$

where X_a are requests.

Add Y_{ai} to coverage limit of physical sub server Id based on $[X_a, Y_a]$.

Find Cov_set as

$$d\lambda = (width * 35)/100 \text{ If } (\sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)}) < d\lambda$$

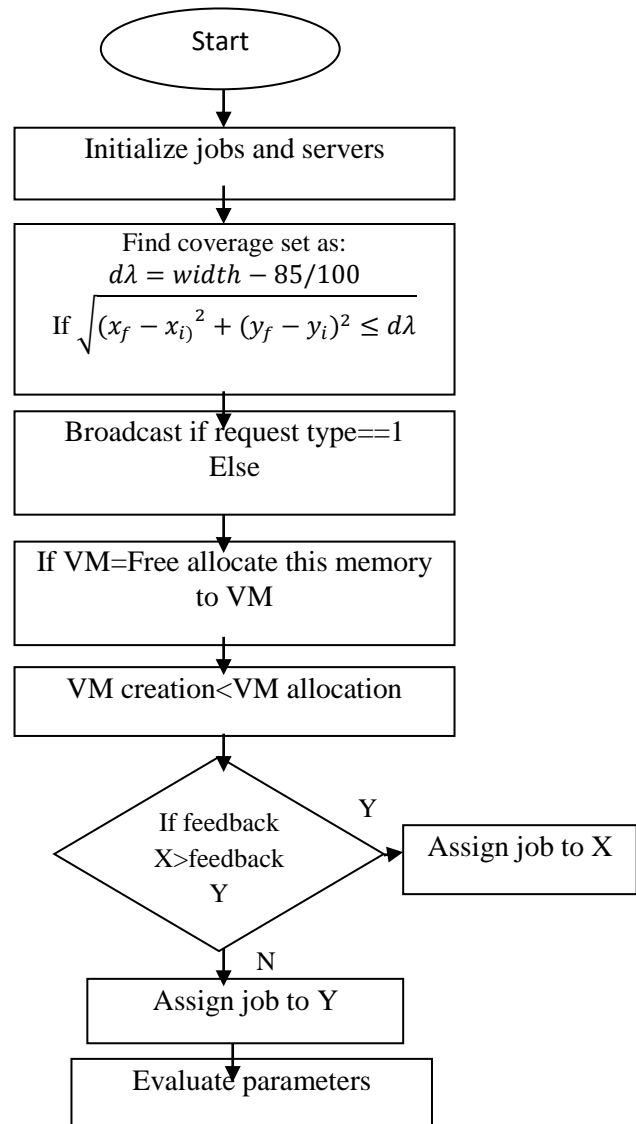


Figure. 2 Methodology flowchart

Table 1. Related work

Author	Method Used	Finding	Limitations
V. Do, Tien, C. Rotter [2]	Analytic performance model.	Obtains Performance measures such as the blocking probabilities, the average energy consumption and the heat emission.	Requires an appropriate allocation scheme in order to get better results.
Youwei Ding, et al. [5]	Energy efficient scheduling algorithm EEVS.	Reconfigured to consolidate the computation resources of the PMs to further reduce the energy consumption.	Requires further study based on practical cluster and real workload.
J. Díaz, J. Luis, J. Entrialgo, M. García, J. García, and D. Fernando Garcia [6]	Load Level based Optimization for Virtual machine Allocation (LLOOVIA).	LLOOVIA can handle problems in which hundreds of VMs of different types, multiple providers, and different kind of limits are used.	Lacks in new approximations and heuristics to avoid the limitations of integer linear programming.
J. Aroca, J. Arjona, A. Fernández Anta, M. Mosteiro, C. Thraves, and L. Wang [14]	Virtual Machine Assignment, Offline Approximation Algorithm and polynomial-time approximation schemes.	Virtual machine assignment (VMA) is utilized and the power consumption is minimized.	Lacks in energy-efficient task assignment in real data centers and VMs cannot be implemented fractionally
H. Cui, X. Liu, T. Yu, H. Zhang, Y. Fang and Z. Xia [15]	Cloud service Task Scheduling Model (TSS), An Ant Colony Optimization (ACO) and a Genetic Algorithm (GA)	The convergence speed and the output performance of proposed approach are optimal.	Requires Improvement in fitness function for hybrid optimization (ACO+GA) in order to increase convergence speed.
S. Kim, J. Byeon, H. Yu and H. Liu [16]	Adaptive Biogeography Based Optimization (BBO)	Performance of BBO is much better in comparison with other optimizations, when the problems are large	BBO does not works well for the middle sizes problems
Q. Liang, J. Zhang, Y. Zhang and J. Liang [17]	A concept of URM and AVMR Algorithm	URM can well represents the utilization ratio of VM's and host in same data	The computation of near optimal configuration is critical and requires in-depth research
M. Sheikhalishahi, R. Wallace, L. Grandinetti, and F. Guerriero [18]	Multi-resource scheduling system based on multi-capacity bin-packing.	Improved wait-time, slowdown and also provides better consolidation degree	Lacks heuristics to be applied on group of jobs to address capacity imbalance

Y. Gao, H. Guan, Z. Qi, T. Song, F. Huan, and L. Liu, [19]	Dynamic Resource management scheme	Proposed scheme can achieve power savings of 50.3%	Still Requires multi-tier applications in a prototype Cloud Data Center.
M. Vanitha and P. Marikkannu [20]	Novel load balancing method.	25% reduction in power consumption in comparison of other techniques.	Fitness level declines with the increase in the number of servers which increases CPU utilization.
S. Madni, M. Latiff, and Y. Coulibaly and S. Abdulhamid [21]	Resource scheduling schemes and algorithms.	Essential parameters and enhancement is requisite to improve the performance of existing schemes.	Scope for further improvement to provide the privacy and security with the resource scheduling.
A. Kaur and P. Nagpal [22]	Hybrid algorithm which is based on neural network with artificial bee colony and some concepts of scheduling has been used.	Hybrid algorithm which is based on neural network with artificial bee colony and some concepts of scheduling has been introduced in work.	Proposed algorithm is executed for less number of iterations.
L.Wu, S. Garg and R. Buyya [23]	Admission control and scheduling mechanism.	Proposed Mechanism provides substantial improvement over static server consolidation and reduces SLA violations.	Need to extend multi-core CPU architectures as well as network and memory conflicts.
S. Ahmed, K. Li, A. Ouyang, and Z. Li [24]	In this a new proactive workload management model is proposed for virtualized resources to inspect the workload behavior of the running Virtual Machines, and to assent an appropriate scheduling and resource consolidation schema.	Reactive systems result into poor performance and may lead to infrequent peak loads followed by lower system performance. To overcome this introduced an adapted version of Statistical Metrics Model (SMM).	Unable to handle intensive resource consolidation situations, and to test different scaling.
W. Lin, S. Xu, L. He and J. Li [25]	CloudSim is extended with a multi-resource scheduling and power consumption model.	Extended model allows more accurate valuation of power consumption in dynamic multi-resource scheduling. Extensive experiments on six combinations of task assignment algorithms and resource allocation algorithms demonstrate the powerful functionality and superior convenience of the extended CloudSim, MultiRECloudSim.	-
N. Taj and A. Basu [26]	Different values of the various factors are used in the calculation of energy of the data centre.	In this few iterated values are considered and have done the analysis and through graphs have shown that even though the parameters in equation changes still are able to calculate the better values of energy consumed in a data center when compared to other previously used equations.	-

S. Loganathan, R. Saravanan, S. Mukherjee [27]	Bacterial foraging optimization algorithm with genetic algorithm.	Used a novel hybrid approach consisting of Genetic Algorithm (GA) and Bacterial Foraging Optimization Algorithm (BFOA) resources are highly utilized with optimized performance result.	If the same job appears again without any execution the resources are not automatically allocated.
A. Bamini Antony Muthu and S. Enoch [28]	Job scheduling mechanism is proposed by considering job classification and pre-emption.	In this job is categorized in to three different types and assigned based on pre-emption policy with the earliest available time of the resource (VM) which is attached to a host. Thereby, energy consumption is reduced by making less number of hosts in the active state and increase the utilization of active host.	At high loads the energy savings is less.

```

Add sub servers to cov_limit
Executes request R.
If request < coverage set
Do same as above.
But
end (if)
end (for)
end (for)

```

The mathematical algorithm evaluates that the relation between PM and VM. Energy reduction will only be done if VM creation < VM allocation.

4.1 Broadcasting algorithm

The chief server broadcast the specification of the user, caches the acknowledgments from the physical machine, and later allocates the job to the sub-server according to their acknowledgments and the minimum energy required to perform the job.

```

Procedure
Repeat forever
call GetBalance to process load balance
    messages received
Jobs execution will be done using subservers
requests execute = Sub server 1 – sub server
If requests = coverage area
execute the jobs
Call GetDistribute to process distribution msg
    received
If (VM creation < VM allocation)
Call Distribute to route excess jobs through
    the queue
If (Vm creation > VM allocation)
Call Balance to initiate a load
    – balancing operation

```

```

Call UpdateLoad (TotalJobsQueued)
to set VM migration
Normal Processing
End
Repeat Procedure Update Load queue

```

4.2 Load balancing algorithm

Let $VM = \{VM_1, VM_2, VM_3, \dots, VM_K\}$ is a set of N virtual machines and $Requests = \{Request_1, Request_2, Request_3, \dots, Request_K\}$ is a set of K requests to be scheduled and processed in VM. All the machines are irrelevant but are correlated. Scheduling is non-preemptive which means that the processing of the requests on VMs cannot be disrupted. The flowchart of the VM scheduling and load balancing employing ABC algorithm is revealed in the Fig. 3. Virtualization in proposed work has been accomplished by applying fitness function of ABC approach. **Fitness value is 1 if $VM.resource.value > demand\ value$ else 0.**

5. Results and discussion

The results are obtained using CloudSim simulator. The prime focus of this work is on scheduling with VM utilization in Cloud Computing. For scheduling in cloud computing, ABC algorithm is utilized. Based on the simulation results it is analyzed that SLA violations with VM appears to be much better as compared with without VM results. Moreover, job scheduling with VM has sufficient rate with respect to without VM machines.

SLA violation can be measure using the following equation:

$$c(p, w, c) = r pr(p) + wr(w) + ar(a) \quad (13)$$

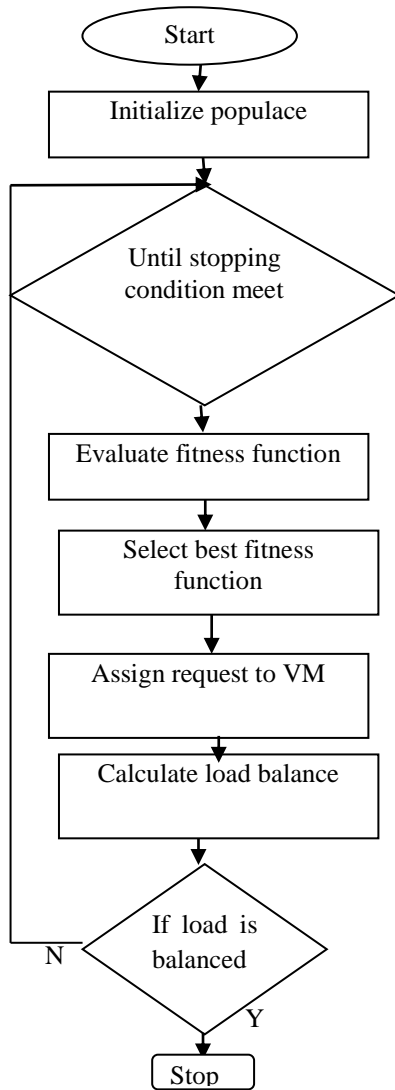


Figure. 3 Scheduling and load balancing algorithm using ABC

We assume functions pr , wr and ar for this evaluation with $pr(p) = 100p$, $wr(w) = 5w$, and $ar(a) = a$, for all r . The main purpose behind selecting these is to less violate the SLA.

5.1 Scenario 1: for 100 jobs

A VM is formed that intakes the memory of PM itself in order to decrease a load on a PM. Scheduling is a method or a job performed on a PM which is frequently popular in this modernized era of technological environment. From result simulations with and without VM the results are $\cong 80$ and $\cong 53$ respectively as revealed in Fig. 4.

The SLA violation with respect to VM machines is shown in Fig 5. SLA migrations have been found to be 0.29. SLA violation must be low in order to have good accuracy.

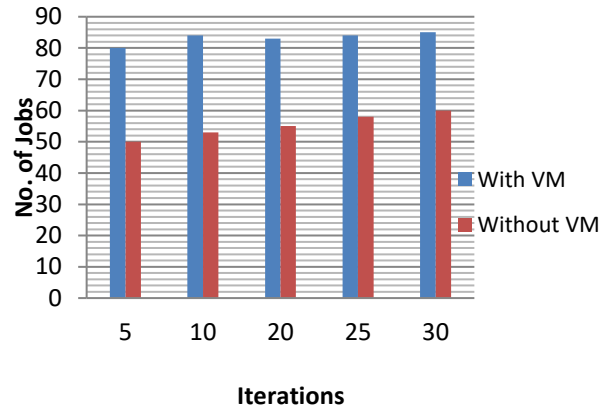


Figure. 4 Job scheduling with VM and without VM for 100 jobs

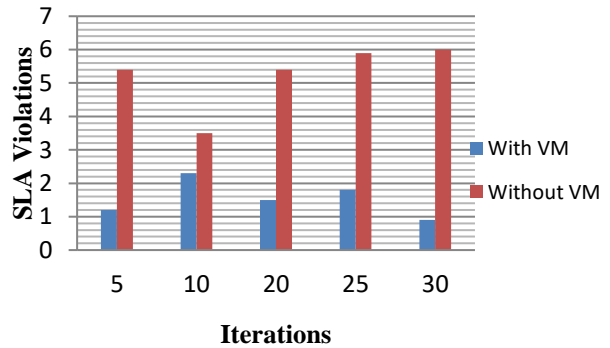


Figure. 5 SLA violations with VM and without VM for 100 jobs

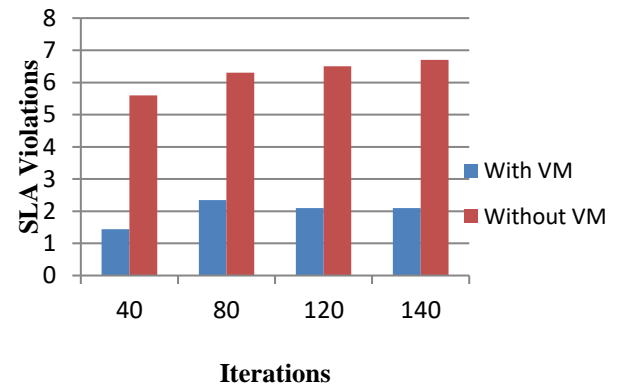


Figure. 6 SLA violations with VM and without VM for 150 jobs

5.2 Scenario 2: for 150 jobs

The SLA violations for 150 jobs with and without VM utilization is shown in fig 6. From graphical representation it has been seen that VM utilization leads to less SLA violation.

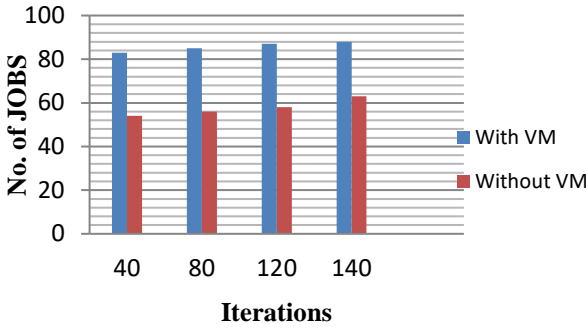


Figure. 7 Job scheduling with VM and without VM for 150 jobs

Table 2. Comparison of job scheduling of proposed work with Amndeeep Kaur [22]

Iterations	With Neural Network	With ABC
1	5	50
2	10	51
3	10	52
4	8	55
5	5	60

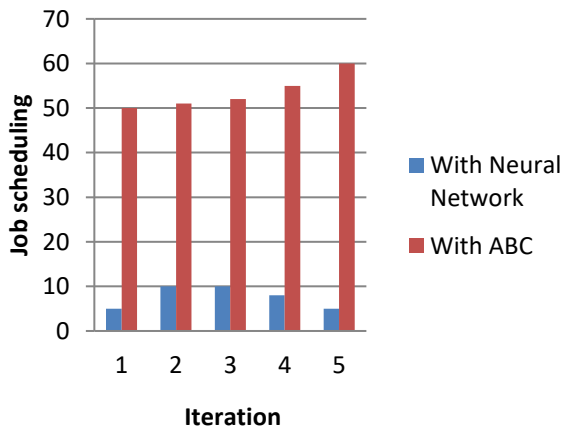


Figure. 8 Job scheduling with neural network and with ABC algorithm

The job scheduling with respect to VM machines is shown in fig 7. Job scheduling must be high in order to have good accuracy.

Note: For high job values, result analysis shows that SLA violation as well as job scheduling for 150 jobs has better numerical values with respect to 100 jobs because high number of jobs leads to increase number of job execution. Also, ABC works well for large number of sets.

5.3 Comparison with existing work

By comparing the present work with the Amndeeep Kaur work [22] it is concluded that the result obtained from the present work are better than the Amndeeep Kaur work [22]. In the present work

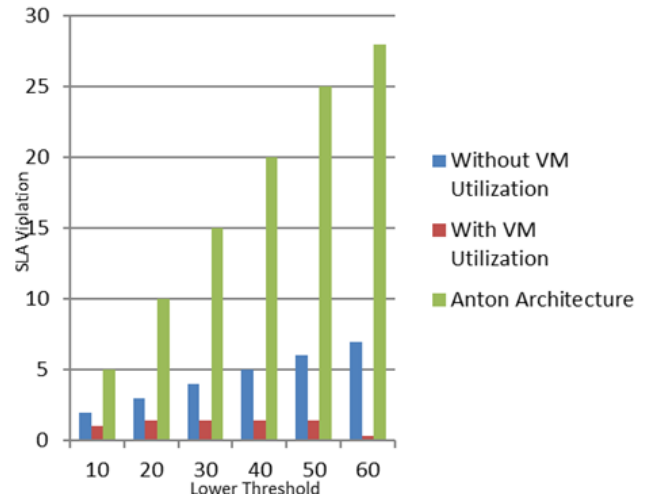


Figure. 9 SLA Violation without VM, with VM and Anton Architecture

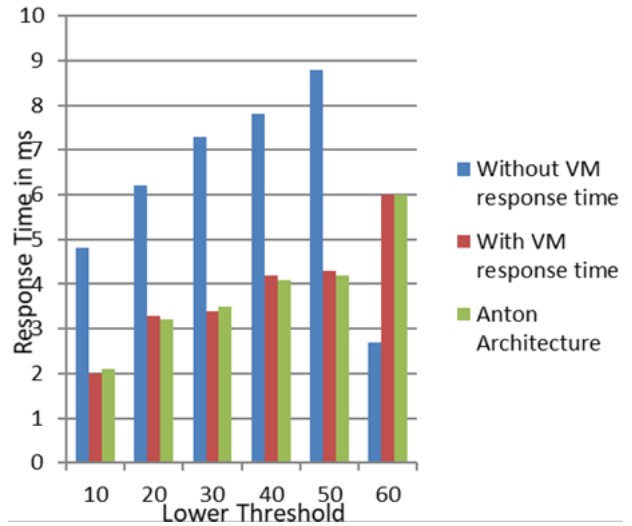


Figure. 10 Response Time without VM, with VM and using Anton Architecture

30 iterations are used whereas as in the Amndeeep Kaur work only five iterations are used to find the jobs. More iteration means more accurate are the results. Table 2 shows the comparison of job scheduling with Amndeeep Kaur work [22].

The Fig. 8 shows the job scheduling with ABC algorithm and with NN [22]. From graphical results in Fig. 9, it has been seen that SLA violation for proposed method is less with respect to Anton et.al [10] method and SLA violation must be low in order to have good accuracy. Average 1.13 SLA violations have been found by using VM, whereas the SLA violations without using VM are near about 4. Lower threshold represents the load in terms of jobs. Lower threshold value 10 represents an increment of 10% in the previous load.

Above graphical results in Fig. 10 shows the response time without VM, with VM and using Anton Architecture. In this, it has been seen that for

3 iterations, response time with VM = 3.87ms and without VM= 6.27ms is originated. For large iterations, the no. of migrations has been enhanced. Both Anton and Amandeep used the concept of allocation pretty well but the problem is allocating a VM to the Physical Machine does not ensure that no other better Physical Machine is present for the presented VM. The proposed work has utilized the concept of broadcast as well to ensure that the VM is allocated to the correct and best suitable Physical Machine and as a result the SLA violation of the proposed work is lower than the previous research works.

6. Conclusion and future work

In this proposed work, SLA and job scheduling for different iterations for checking the performance of the technique based on artificial bee colony algorithm has been taken place. From the experimental results, this is concluded that utilization of virtual migration in the proposed work is best from other policies, because it least violates SLA. In this work, SLA is calculated, by considering number of iterations. The comparison of ABC- with VM and without VM has been made on the job factors that are ultimately related to SLA violations. At last, comparisons of proposed work with various existing techniques have been provided. In the existing techniques, we have considered the neural network and Genetic algorithm and the comparison of SLA violation and job scheduling has been provided. The SLA Violation of the proposed architecture has been compared with famous Anton Architecture and a valid improvement has been noticed due to the broadcast architecture.

In future, to improve SLA violation and job scheduling in VM a hybridization of optimization algorithm along with classification algorithm namely Artificial Bee colony along with Genetic algorithm and neural network can be used.

References

- [1] M. Armbrust, A. Fox, and R. Griffith, "A View of Cloud Computing", *Communications. Communications of the ACM*, Vol.53, No.4, pp 50-58, 2010.
- [2] V. Do, Tien, and C. Rotter, "Comparison of Scheduling Schemes for On-Demand IaaS Requests", *Journal of Systems and Software*. Vol. 85, No.6, pp. 1400-1408, 2012.
- [3] M. Frincu, S. Genaud, and J. Goss, "On the efficiency of several VM provisioning strategies for workflows with multi-threaded tasks on clouds computing", *Computing*, Vol. 96, No. 11, pp. 1059–1086, 2014. 4.
- [4] F. Lombard and R. D. Pietro, "Secure virtualization for cloud computing", *Journal of Network and Computer Applications*, Vol. 34, No.4, pp. 1113-1122,2011.
- [5] Y. Ding, X. Qin, L. Liu, and T. Wang, "Energy efficient scheduling of virtual machines in cloud with deadline constraint", *Future Generation Computer Systems*, Vol. 50, pp. 62-74, 2015.
- [6] J. Díaz, J. Luis, J. Entrialgo, M. García, J. García, and D. Fernando Garcia, "Optimal allocation of virtual machines in multi-cloud environments with reserved and on-demand pricing", *Future Generation Computer Systems*, Vol. 71, pp.129-144,2017.
- [7] J. James and B. Verma, "Efficient VM Load Balancing Algorithm for a Cloud Computing Environment", *International Journal on Computer Science and Engineering*, pp.1658-1663, 2012.
- [8] U. Bhoi and P. Ramanuj, "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing", *International Journal of Application or Innovation in 184 Engineering & Management*, Vol.2, No.4, pp. 259-264, 2013.
- [9] S. Garg, A. Toosi, S. Gopalaiyengar, and R. Buyya, "SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter", *Journal of Network and Computer Applications*, Vol.45, pp. 108-120, 2014.
- [10] A. Beloglazov and R. Buyya, "Energy Efficient Resource Management in Virtualized Cloud Data Centers", In: *Proc. of the 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, pp. 826-831, 2010.
- [11] N. Serrano, G. Gallardo, and J. Hernantes, "Infrastructure as a service and cloud technologies," *IEEE Software*, Vol. 32, No.2, pp. 30-36, 2015.
- [12] X. Wu, M. Deng, R. Zhang, B. Zeng, and S. Zhou, "A Task Scheduling Algorithm based on QoS driven in Cloud Computing", *Information Technology and Quantitative management*, Vol. 17, pp.1162-1169, 2013.

- [13] Y. Lu, Q. Xie, G. Kliot, A. Geller, J. Larus, and A. Greenberg, "A novel load balancing algorithm for dynamically scalable web services", *Performance Evaluation*, Vol. 68, No. 11, pp. 1056-1071, 2011.
- [14] J. Aroca, J. Arjona, A. Fernández Anta, M. Mosteiro, C. Thraves, and L. Wang, "Power-efficient assignment of virtual machines to physical machines", *Future Generation Computer Systems*, Vol. 54, pp.82-94, 2016.
- [15] H. Cui, X. Liu, T. Yu, H. Zhang, Y. Fang, and Z. Xia, "Cloud Service Scheduling Algorithm Research and Optimization", *Security and Communication Networks*, Vol. 2017, 2017.
- [16] S. Kim, J. Byeon, H. Yu, and H. Liu, "Biogeography-based optimization for optimal job scheduling in cloud computing", *Applied Mathematics and Computation*, Vol. 247, pp. 266-280, 2014.
- [17] Q. Liang, J. Zhang, Y. Zhang, and J. Liang, "The placement method of resources and applications based on request prediction in cloud data center", *Information Sciences*, Vol. 279, pp.735-745, 2014.
- [18] M. Sheikhalishahi, R. Wallace, L. Grandinetti, and F. Guerriero, "A multi-dimensional job scheduling", *Future Generation Computer Systems*, Vol. 54, pp. 123-131, 2016.
- [19] Y. Gao, H. Guan, Z. Qi, T. Song, F. Huan, and L. Liu, "Service level agreement based energy-efficient resource management in cloud data centers", *Computers & Electrical Engineering*, Vol. 40, No.5, pp. 1621-1633, 2014.
- [20] M. Vanitha and P. Marikkannu, "Effective resource utilization in cloud environment through a dynamic well-organized load balancing algorithm for virtual machines", *Computers & Electrical Engineering*, Vol. 57, pp. 199-208, 2016.
- [21] S. Madni, M. Latiff, Y. Coulibaly, and S. Abdulhamid, "Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities", *Journal of Network and Computer Applications*, Vol. 68, pp. 173-200, 2016.
- [22] A. Kaur and P. Nagpal, "Efficient Cloud Server Job Scheduling using NN and ABC in cloud computing", *International Journal of Intelligent Engineering and Systems*, Vol.11, No.3, 2018
- [23] L. Wu, S. Garg, and R. Buyya, "SLA-based resource allocation for software as a service provider (SaaS) in cloud computing environments", In: *Proc. of the 2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 195-204, 2011.
- [24] S. Ahmed, K. Li, A. Ouyang, and Z. Li, "Proactive workload management in dynamic virtualized environments", *Journal of Computer and System Sciences*, Vol. 80, No. 8, pp. 1504-1517, 2014.
- [25] W. Lin, S. Xu, L. He, and J. Li, "Multi-resource scheduling and power simulation for cloud computing", *Information Science*, Vol. 397, pp. 168-186, 2017.
- [26] N. Taj and A. Basu, "Implementation of Energy Equation Used to Calculate the Energy Used By the Data Centre", *International Journal of Advances in Engineering & Technology (IJAET)*, Vol. 9, No. 2, pp. 240-246, 2016.
- [27] S. Loganathan, R. Saravanan, and S. Mukherjee, "Energy Aware Resource Management and Job Scheduling in Cloud Datacenter", *International Journal of Intelligent Engineering and Systems*, Vol. 10, No.4, pp. 175-184, 2017.
- [28] A. B. A. Muthu and S. Enoch, "Optimized Scheduling and Resource Allocation Using Evolutionary Algorithms in Cloud Environment", *International Journal of Intelligent Engineering and Systems*, Vol.10, No.5, pp.125-133, 2017.