Energy optimization methods for Virtual Machine Placement in Cloud Data Center

Esha Barlaskar[1], N. Ajith Singh[2] and Y. Jayanta Singh[1]
[1]Dept. Computer Science & Engineering and I.T,
School of Technology of Assam Don Bosco University,
Guwahati, India
eshabarlaskar[at]gmail.com, jayanta[at]dbuniversity.ac.in
[2]IT Consultant, AjithEx Technologies, Imphal India
ajithex[at]gmail.com

Abstract—The Information Technology industry has been upheaved by the influx of cloud computing. The extension of Cloud computing has resulted in the creation of huge data centers globally containing numbers of computers that consume large amounts of energy resulting in high operating costs. To reduce energy consumption providers must optimize resource usage by performing dynamic consolidation of virtual machines (VMs) in an efficient way. The problems of VM consolidation are host overload detection, host under-load detection, VM selection and VM placement. Each of the aforestated sub-problems must operate in an optimized manner to maintain the energy usage and performance. The process of VM placement has been focused in this work, and energy efficient, optimal virtual machine placement (E2OVMP) algorithm has been proposed. This minimizes the expenses for hosting virtual machines in a cloud provider environment in two different plans such as i) reservation and ii) on-demand plans, under future demand and price uncertainty. It also reduces energy consumption. E2OVMP algorithm makes a decision based on the gilt-edged solution of stochastic integer programming to lease resources from cloud IaaS providers. The performance of E2OVMP is evaluated by using CloudSim with inputs of planet lab workload. It minimized the user's budget, number of VM migration resulting efficient energy consumption. It ensures a high level of constancy to the Service Level Agreements (SLA).

Keywords: Cloud resource management; virtualization; dynamic consolidation; stochastic integer programming (SIP)

**Cite as: Esha Barlaskar, N. Ajith Singh, Y. Jayanta Singh, "Energy optimization methods for Virtual Machine Placement in Cloud Data Center" ADBU J.Engg.Tech., 1(2014) 0011401(7pp)*

1. Introduction

Cloud computing provide services in pay-as-you-go manner over the Internet. Therefore, service providers without making any advance investment in infrastructure could only rent resources from infrastructure providers. Customers can choose two payment plans which are offered by cloud providers namely reservation plan (pre-paid) and on-demand plan (pay per-use). The price of resources in on-demand plan is expensive than reservation plan. This can lead to an under provisioning problem when the quantity of reserved resources is unable to meet the demands fully. If a customer reserves resources more than required than an over provisioning problem occurs which cannot be neglected as well since the amount of reserved resources will be underutilized thus again increasing the cost which is referred to as oversubscribed cost. So it is necessary to minimize both on-demand and oversubscribed costs. The cost for maintenance by infrastructure providers and energy consumption increases rapidly. The energy usage could be minimized using a virtualization technology. This will allow using fewer physical servers with much higher per-server utilization. However it harbingers new challenges for the management of VM. So, VM management in cloud data centers must be provisioned and managed very productively and hence must pave the way for optimizing the energy-performance trade-off.

Dynamic assignments of physical resources to virtual machines are required to make energy efficient and improve the quality of service related to SLA. Dynamic VM consolidation as a vital control process is an efficient management system to amend energy efficiency [1]. The dynamic VM consolidation problem has four sub-problems.

1. Determine when a host is considered as being overloaded (host overloading detection). Live migration is required to migrate one or more VMs from the overloaded host;
2. Determine when a host is considered as being under-loaded. Here the host is ready for switching to sleep mode, allowing migration of its all VMs.
3. Determine which VMs must be selected to migrate from overloaded host.
4. Determine which hosts must be selected to place migrated VMs. The objective of this procedure is to perform in such a way that optimizes energy-performance tradeoff inside cloud data center. Each of aforestated sub-problems must operate in better optimized way. We proposed an algorithm for VM placement called Energy Efficient Optimal Virtual Machine Placement (E2OVMP) algorithm which minimizes the total cost.

This study explores optimal methods by considering both under provisioning and over provisioning problems of resource management and also to reduce the overall energy consumption. Taking IaaS model, E2OVMP algorithm can host a certain number of VMs taking into consideration the uncertainty of future demands and prices of resources between on-demand and oversubscribed costs. The result of E2OVMP algorithm is obtained as the gilt-edged solution from stochastic integer programming (SIP) formulation with two-stage recourse [2]. The performance of E2OVMP is evaluated by using CloudSim. The results show that an E2OVMP algorithm can reduce the total cost while meeting the requirements of both providers and customers. The new proposed algorithm reduces energy consumption and number of SLA violation resulting in maximization of overall performance. The remainder of the paper is systematized as follows. In Sect. 2 we present a literature survey. Sect. 3 provides an introduction of the proposed algorithm. Sect. 4 shows the experimental result. The section 5 provides the summary of the study and concludes the paper.

2. Related works

A cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements[3]. There is SLA between the provider and the consumer for getting services from the cloud on pay per user basis which must not be violated. The proposed eco-friendly cloud computing which not only diminished global warming but also will minimize operational cost by reducing power consumption. They have introduced novel algorithm (using honey bee and ant colony algorithm). In cloud computing environment some CPUs of IaaS are overloaded for processing consumers' services, some are under loaded, and some are totally idle. It can save the consumption of energy by turning these idle CPUs off and rescheduling services from overloaded CPUs to under loaded CPUs. The work will not only save energy but will also prevent potential SLA violation. But the problem lies in managing these idle CPUs, overloaded CPUs and under loaded CPUs efficiently; for this the study proposed Bee-Ants colony system. At the beginning, jobs are divided into two parts; 1st part looks after the proper control of overloaded, under loaded CPUs (service rescheduling), and 2nd helps to manage the idle CPUs (power consumption control). The study proposed bee colony algorithm for service rescheduling and ant colony algorithm for power consumption management emphasizing on an eco-friendly cloud computing. The study only focused on overload, under load and idle host detection but didn't consider VM-migration in their work [4]. There is a problem in load balancing of VM resource scheduling in such environment. It mainly considers the current state of the system but seldom considers system variation and previous data resulting to load imbalance. For such problem, Jinhua et. al. [5] proposed a scheduling strategy based on genetic algorithm that considers historical data and current state of the system. Hence computers ahead the influence it will have on the system after the deployment of the needed VM resources and then chooses the least-affective solution, through which it achieves the best load balancing. It solves the problem of load imbalance and high migration cost. However, the study was not suitable in situations where the user requirements differ with time. It becomes almost impossible to predict the future conditions if the user keep on changing the type of their requests. A possible solution for such situations is to maintain a record of available VMs. Thus, record keeping will make the process of resource allocation much faster by provisioning the VMs based on the cloudlet requirements as it arrives instead of predicting ahead by considering the historical data. Bobroff et. al. [6] proposed a new algorithm for preserving performance remapping the VM to PM for future resource demand. Barbagallo et al. described a bio-inspired algorithm based on methode of the scout-worker migration where some entities are allowed to move from one physical machine to another in order to identify cooperatively for a suitable destination for VMs. It focused on performance only [7]. Many software-oriented techniques have tried to reduce the consumed energy in Cloud by means of virtualization technology (VT)[8,9]. Scheduling and resource allocation by means of VMs is one of the most popular approaches to decrease the consumed energy of virtualized datacenters. The study of Ajith Singh et. al.[10] proposed a nature-inspired honey bee algorithm for solving the dynamic VM placement problem for energy-aware Infrastructures as a Service (IaaS) cloud computing environments. It has tackled the problem of power efficient management in virtualized datacenters to maximize the provider's profit by minimizing both power consumption and SLA violation. Bernardetta Addis et. al. [11] expressed five problem areas to be considered in the allocation policy. They have integrated all areas within a unifying framework, providing very efficient and robust solutions at various time-scales. Mohsen et. al. [12] Considered energy efficiency along with performance. They showed that the blind consolidation of VMs does not reduce the power consumption of datacenters, but it can also lead to energy wastage. They proposed an energy-aware scheduling algorithm using a set of objective functions in terms of consolidation fitness metric. The study assigns a set of VMs to a set of PMs aiming to minimize the total power consumption (PMs) in the entire datacenter.

The techniques of VM placement and consolidation which leverage min-max and shares features provided by hypervisors is proposed[13]. A dynamic consolidation method based on constraint programming for homogeneous clusters was developed [14]. Such work did not consider uncertainty of future demands and prices. The resource provisioning options were introduced [15]. A probabilistic advance reservation was proposed that relies on existing best effort batch schedulers which again cannot be guaranteed to be optimal [16]. A study [17] proposed a binary integer program to maximize resource providers' revenues and utilization and heuristic methods to solve this binary integer program. Such study did not consider uncertainty of future users' demands on the contrary uncertainty was taken into consideration. An optimization framework for resource provisioning was considered taking multiple client QoS classes under uncertainty of workloads, and the arrival pattern of workloads is estimated by using online forecasting techniques [18]. Stochastic programming has been developed to solve resource planning under uncertainty. In contradiction to the work in [19] specifies that demands given probability distributions.

3. The proposed Methodology

This study mainly considers IaaS as it has been recognized as the most promising model. Such Iaas represented by a large scale data center comprising large number of heterogeneous physical node. Each node is characterized by CPU performance, disk storage, amount of RAM and network bandwidth[4]. The proposed study architecture is adopted from [1] enhancing the efficient management of workload by assigning them in a proper manner[20]. Our proposed work is similar to the work done in[19] but we also consider energy efficiency in our work which is not being considered in[19]. Stochastic programming was applied to solve problems and even in resource provisioning; however, to the best of our understanding it has never been studied in energy efficiency for VM consolidation. The dynamic VM consolidation problem can be divided into four problems.

1. Checking whether the host is under loaded.
2. Checking whether the host is overloaded.
3. Selection policy to migrate VMs from overloaded host.
4. VMs placement for placing the VMs in allocation or migration to other active/reactivated node host.

Among all the aforementioned sub-problems, we have focused on VM placement and for the rest of the sub-problems we are using the existing algorithms[1]. For VM placement, we have proposed an energy efficient, optimal virtual machine placement (E2OVMP) algorithm which minimizes the total cost, while fulfilling the requirements of both providers and customers. Our proposed algorithm is based on SIP[19] and Bin Packing method [1, 21]where bins are described by Physical machines and the objects to be filled in the bin are represented by VMs. Bin Packing process enhanced the SIP algorithm allocating the VM to the server. The server reserved some resource for scalability. It has the global manager which implements the SIP algorithm to make the optimal decision for the users to reserve resources and host VMs to any PM in data center and the local manager. The VM Monitor (VMM) performs the actual placement, migration, resizing of VM etc after getting triggered by the global and local managers. Users are offered two payment plans, i.e., reservation and on-demand plans where the Cloud provider offer the price of resources which will be charged to the user when the resources are utilized or reserved, and price to provision resources in reservation plan is assumed to be cheaper than that in on-demand plan[19]. There are three phases of provisioning resources: reservation, utilization, and on-demand. In the reservation phase the users without knowing their exact future demand provision resources in the reservation plan. Next, the utilization phase starts when the reserved resources are used but if the demand exceeds the amount of reserved resources, the user has to pay for additional resources in the on-demand plan, and finally the on-demand phase starts. There are three costs affiliated with provisioning resources either reservation or utilization or on-demand. The sum of utilization and reservation costs for the same resource is generally less than that of an on-demand cost. The duty of the global manager is to be to minimize all the costs meeting the user's demands. It uses E2OVMP algorithm to obtain an optimal solution by reserving the optimum number of resources in the reservation phase. An optimal solution is achieved by solving and codifying a stochastic integer programming with two-stage remedy for VM placement which is discussed in III-A. There are two stages of decision making. First stage defines the number of VMs provisioned in phase of reservation, and the second stage defines the number of VMs allocated in both utilization and on-demand phases. The second stage represents the actual number of VMs required by the user and real prices defined for subscribing the VMs. The number of required VMs by the user is not known during the reservation of the resource is made due to uncertainty in demands.

A. VM Placement problem based on Stochastic Integer Programming(SIP)

The cloud providers supply a pool of resources to the user in the form of virtual machines. These VM can be divided into a set of classes with each class representing a different application type with the goal of minimizing all the costs at the provider's end by meeting the demands of users. The demands and prices can't be fixed. This value can be estimated based on a distribution curve. SIP is found to be useful for solving the problem of VM placement.

The problem statement is given below. Assuming a set of VM classes which represent application types and set of cloud providers who supply these VM classes as computing resources to users. Consider that the providers offer four types of resources-computing power, storage, network bandwidth and electric power which are denoted by the superscripts (p), (s), (n) and (e)respectively. The uncertainty of demands and prices are also taken into account. Thus, all these parameters are denoted as follows:

- $V=\{V_1,V_2,V_3,\ldots,V_{last}\}$ denotes the set of VM classes where a VM class represents an application type.
- $P=\{P_1,P_2,P_3,\ldots\ldots,P_{last}\}$ denotes set of cloud providers who supply a pool of resources to the user.
- $D_i=\{d_{i1},d_{i2},d_{i3},\ldots,d_{iv}\}$ denotes set of maximum number of required VMs of class V_i. The total number of required VMs D will be Cartesian product of all D_i over all i.
- $v_i(d)$ denotes the number of required VMs in class V_i if demand d is realized.
- $t_j^{(p)}$, $t_j^{(s)}$, $t_j^{(n)}$ -the maximum capacity of corresponding resource which cloud provider P_j can supply to user.
- $r_i^{(p)}$, $r_i^{(s)}$, $r_i^{(n)}$ denotes amount of corresponding resource required by a single VM under class V_i.
- $c_{j(rev)}^{(p)}$, $c_{j(rev)}^{(s)}$, $c_{j(rev)}^{(n)}$ denotes the prices of resources in reservation phase for provider P_j.
- $c_{j(util)}^{(p)}$,$c_{j(util)}^{(s)}$, $c_{j(util)}^{(n)}$ denotes the prices of corresponding resources in utilization phase for cloud provider P_j.
- $c_{j(od)}^{(p)}$, $c_{j(od)}^{(s)}$ $c_{j(od)}^{(n)}$ denotes the prices of corresponding resources in on-demand phase for cloud provider P_j and these costs can be random.
- $C_{j(od)}^{(p)}$, $C_{j(od)}^{(s)}$ $C_{j(od)}^{(n)}$ denotes set of possible prices of offered resources by provider P_j in on-demand phase.

The SIP algorithm has two stages. Stage1 defines the number of VMs to be provisioned in reservation phase and stage2 defines the number of VMs allocated in utilization and on-demand. The SIP formulation is expressed as:

$$\sum_{V_i \in V} \sum_{P_i \in P} c_{ij} X_{ij}^{(r)} + \xi\Omega[\rho(X_{ij}^{(r)}, \omega)] \text{..........}(1)$$

Where, $X_{ij}^{(r)}$ denotes number of VMs provisioned in stage1, and $\xi\Omega[\rho(X_{ij}^{(r)}, \omega)]$ denotes cost in stage2. $\omega \in \Omega = D \times \prod_{P_j \in P}$ (realizations-set of demands, prices).

The SIP formulation of OVMP algorithm[19] defined in equation (1) can be transformed into a deterministic integer program called deterministic equivalent formulation by introducing two variables $X_{ij}^{u}(m,d)$ and $X_{ij}^{o}(m,d)$. The deterministic equivalent SIP of OVMP algorithm is expressed in equations (2)-(8). When demand $d \in D$ and prices $m \in c_{j(od)}$ are achieved, then the two variables $X_{ij}^{u}(m,d)$ and $X_{ij}^{o}(m,d)$ denote the number of VMs allocated in utilization phase and on-demand phase. The goal is to reduce the function of equation1. Therefore, taking the probabilities of demands and costs being realized, the above equation(1) can be expanded.

$$\sum_{V_i \in V} \sum_{P_j \in P} c_{ij} X_{ij}^{(r)} + \sum_{V_i \in V} \sum_{P_j \in P} \sum_{m \in C_j} \sum_{d \in D} p_j(m) p(d) (c_{(util)} X_{ij}^{(u)}(m,d) +$$

$$c_{(od)}(m) X_{ij}^{(u)} + c_{(od)}(m) X_{ij}^{(o)}(m,d)) \text{................................}(2)$$

$$X_{ij}^{(u)}(m,d) \le X_{ij}^{(r)}, V_i \in V, P_j \in P, m \in C_j, d \in D \text{........}(3)$$

$$\sum_{P_j \in P} (X_{ij}^{(u)}(m,d) + X_{ij}^{(o)}(m,d)) \ge v_i(d), V_i \in V, m \in C_j, d \in D \text{.......}(4)$$

$$\sum_{V_i \in V} r_i^{(p)} (X_{ij}^{(u)}(m,d) + X_{ij}^{(o)}(m,d)) \le t_j^{(p)}, P_j \in P, m \in C_j, d \in D \text{.......}(5)$$

$$\sum_{V_i \in V} r_i^{(s)} (X_{ij}^{(u)}(m,d) + X_{ij}^{(o)}(m,d)) \le t_j^{(s)}, P_j \in P, m \in C_j, d \in D \text{.......}(6)$$

$$\sum_{V_i \in V} r_i^{(n)} (X_{ij}^{(u)}(m,d) + X_{ij}^{(o)}(m,d)) \le t_j^{(n)}, P_j \in P, m \in C_j, d \in D \text{.....}(7)$$

$$X_{ij}^{(r)}, X_{ij}^{(u)}(m,d), X_{ij}^{(o)}(m,d) \in \{0,1,.....\}, V_i \in V, P_j \in P, m \in C_j, d \in D \text{..}(8)$$

In eqn(2) has two probabilities p(d) and $p_j(m)$ where p(d) denotes the probability if demand d∈D is realized and $p_j(m)$ denotes the probability if prices $m \in C_{j(od)}$ offered by provider Pj. All constraints in (3)−(8) consider both realizations of *d* and *m*. The constraint in (3) ensures that the number of VMs in utilization phase does not exceed that in reservation phase. The demand under a realization ω is governed by the constraint in (4), where $v_i(\omega)$ denotes the VMs number required in stage2. In(4), the VM allocation of the second stage is governed by $v_i(d)$, where $v_i(d)$ denotes the number of required VMs in class V_i if demand d is achieved. The constraints in (5),(6), and (7) ensure that the allocation of VMs in class V_i must be bounded by the resource capacity offered by provider P_j. respectively. The list of constraint in (8) indicates that all decision variables take the values from a set of non-negative integer number.

B. VM placement problem based on Bin Packing method

The VM placement is viewed as a bin packing problem with variable bin sizes and prices, where bins represent PM, items are the VMs. Bin sizes are the available CPU capacities of the PM and prices correspond to the power consumption by the PM. A modified version of the Best Fit Decreasing (BFD) algorithm is used. All the VMs are sorted in the decreasing order of their current CPU utilization. And allocate each VM to a host that provides the smallest growth of the power consumption induced by the allocation. It allows the incorporation of PMs heterogeneity by firstly choosing the most power-efficient ones. The complexity of the algorithm is *nm*, where *n* is the number of physical machines and *m* is the number of VMs that have to be allocated and the algorithm is given below:

<u>Algorithm: Power Aware Best Fit Decreasing (PABFD)[1]</u>

1. ***Input:*** *hostList, vmList* ***Output:*** *allocation of VMs*
2. *Call vmList.sortDecreasingUtilization()*
3. *For each vm in vmList do*
4. *minPower MAX*
5. *allocatedHost NULL*
6. *For each host in hostList do*
7. *If host has enough resources for vm then*
8. *power estimatePower(host, vm)*
9. *If power < minPower then*
10. *AllocatedHost host*
11. *MinPower power*
12. *If allocatedHost 6= NULL then*
13. *Allocation.add(vm, allocatedHost)*
14. *Return allocation*

C. Proposed E2OVMP algorithm:

The proposed algorithm is the combination of SIP[19] and enhances version of BFD[1] VM placement algorithms (E2OVMP). At first the SIP algorithm is executed by global manager to perform the reservation as well as to meet the on-demand and utilization phases of the VM placement. Then the local managers execute the modified BFD algorithm to place the VMs in such a way that results in minimum usage of physical machines. Then the VMM performs the actual VM placement, migration and resizing activities after receiving commands from the global and local managers-VMs number, location and time. Once the VMM gets all the required information it starts its actual placement. E2OVMP's objective is to reduce the energy consumption of datacenter, optimize the overall cost and maximize their ROI.

D. Metrics Used For Measuring Energy Consumption and SLA Violation Due To Performance Degradation

The performance of our proposed work has been evaluated using existing metrics[1]. This algorithm is used to optimize two main parameters -energy consumption and SLA violation related to performance degradation. To portray the energy-performance tradeoff both the definition for energy consumption and performance degradation must be defined distinctly. In this study, the Energy Consumption (EC) by a server is defined as a linear function from CPU utilization, and performance is defined as a function that evaluates the SLA delivered to any VM deployed in an IaaS. The SLA violation is defined with the help of two metrics-SLA Violation Time per Active Physical machine (SLATAH) that rise with overload time period of the PM, and Performance Degradation due to Migrations (PDM) that rise due to live migration[1]. Hence, these metrics were defined with the assumption that the SLAs are delivered when 100% performance requested by any kind of applications inside a VM is provided at any time.

$$SLATAH = \frac{1}{N}\sum_{i=1}^{N}\frac{T_{si}}{T_{ai}} \text{..........(9)}$$

$$PDM = \frac{1}{M}\sum_{j=1}^{M}\frac{D_{dj}}{D_{rj}} \text{............(10)}$$

N is the number of active physical machines; T_{si} is total time during which physical machine i has experience CPU utilization of 100%; T_{ai} is total time during which physical machine i being in the serving VMs; M is the number of VMs; D_{dj} is an estimation of the performance degradation of the VM j caused by migration (10% for our study); D_{rj} is total CPU capacity requested by VM j during its lifetime. A metric for describing SLA violation can be defined:

$$SLAV = SLATAH \times PDM \text{...... (11)}$$

In consideration of aforementioned formulation SLA Time (SLAT) for each physical machine can be define as:

$$SLAT_i = \frac{T_{si}}{T_{ai}} \text{......(12)}$$

$$1 \leq i \leq N$$

4. Performance Evaluation

The proposed E2OVMP algorithm reserves the resources in VM placement by placing as many VMs that can accommodate in a single PM after considering the reserved resources. This study reduces the VM migration, SLA violation and Energy consumption. This method first reserves the resource while VM allocation with the certain amount as reserve resource on host and when the host reaches its total resource usage in utilization phase it then allows the VM to use the free resource which was being reserved earlier. We have used the Planetlab Workload traces which is available with the CloudSim package to perform our simulation which is run for 24hours with proposed VM placement algorithm along with the existing overload detection and VM selection algorithms[1]. The experimental result is shown in Figure 1, 2 and 3. The outcomes of the study are illustrated as follows:

1. Dynamic VM consolidation algorithms significantly outperforms static allocation policies like DVFS;
2. Combination of our method (E2OVMP) with LR-MC policy reduces the energy consumption significantly.
3. E2OVMP together with LRR-RS shows less number of SLA violation compared to other overload detection and VM selection policies combined with this process;
4. E2OVMP combined with LRR-MC has less number of VM migrations than with other overload detection and VM selection policies.

A. Energy Chart:

The proposed E2OVMP process has been implemented with four different parameters and five host overload detection algorithms. The four different parameters are

1. VM selection-MC (Maximum Correlation),
2. MMT (Min. Migration Time),
3. MU (Min. Utilization),
4. RS (Random Selection).

The five host overload detection algorithms are

1. THR (Static Threshold),
2. MAD(Median Absolute Deviation,
3. IQR (Inter quartile Range),
4. LR (Local Regression) and
5. LRR(Robust Local Regression).

The E2OVMP algorithm with LR-MC policy gives the better result for Energy compared to other policies, i.e., energy consumption was reduced significantly by applying E2OVMP with LR-MC (Local Regression and Maximum Correlation) policy. The detail outcome of this study is given as figure 1.

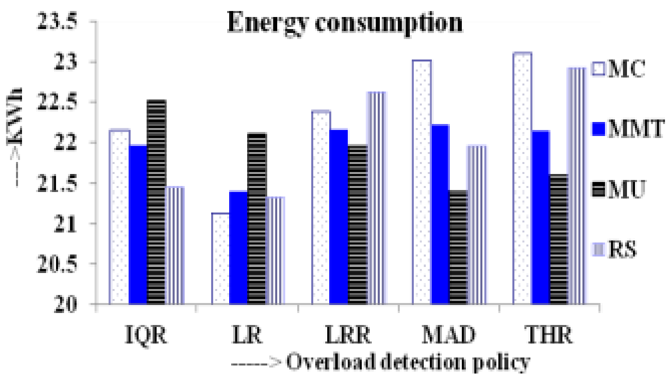

Figure1: Energy consumption

B. SLA Violation Chart:

The E2OVMP together with LRR-RS algorithms shows less number of SLA violation compared to other overload detection and VM selection policies. The detail outcome of this study is given as figure 2.

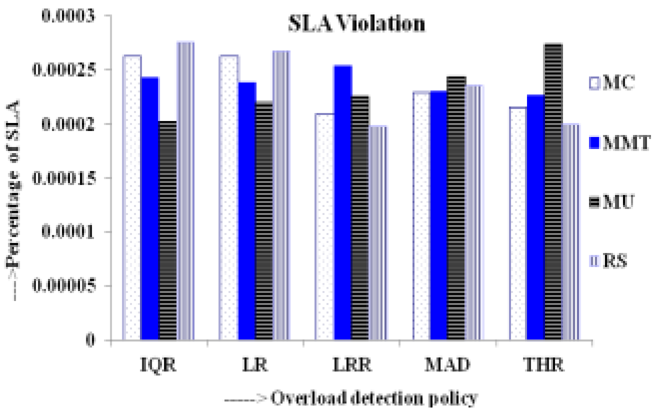

Figure2: SLA Violation

C. VM Migration Chart:

The E2OVMP combined with LRR-MC policy has less number of VM migration than with other overload detection and VM selection policies(figure3). The results indicated that if VM are selected based on the correlation between them then number of migrations can be minimized and outliers must be taken in consideration like LRR policy.

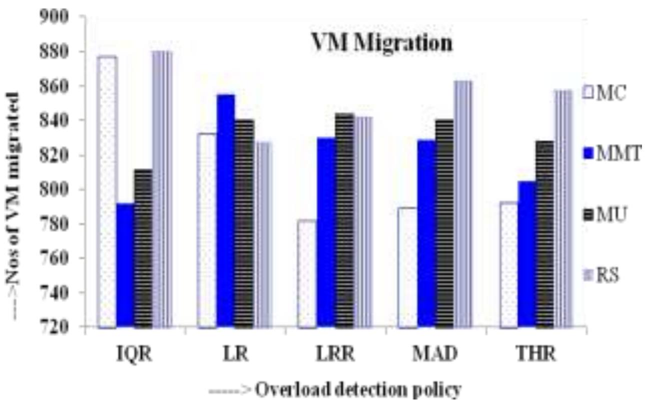

Figure3: Number of VM Migration.

5. Conclusion of the study

The cloud providers must implement energy efficient resource management techniques to maximize their ROI. Hence dynamic consolidation of VMs have come as a vital solution for this problem which is achieved by switching idle servers to power-saving modes. Our proposed algorithm E2OVMP reduces the energy consumption in data centre by implementing reservation technique and using enhancing BFD. E2OVMP algorithm shows better results with different combinations of overload detection policy and VM selection policy. It is evaluated through extensive simulations using workload traces from PlanetLab. E2OVMP algorithm gave better results with various combinations of VM selection and overload detection policies. We believe that our work can be extended with improvements and additional functionalities can be further added into it. Several studies in similar manners are under implementation for better procedures for VM consolidations in cloud computing environment.

References

[1] A.Beloglazov, R. Buyya, Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurr. Comput. Pract. Experience 24(13), 1–24 (2011)

[2] J. R. Birge, F. Louveaux, Introduction to Stochastic Programming, Springer-Verlag Newyork, Inc., 1997

[3] R Buyya, C. Shin Yeo and S Venugopal, “Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities”, 10th IEEE Int. Conf. on High Performance Computing & Communications, IEEE Computer Society, 2008.

[4] K.Mukherjee and G.Sahoo, “Green Cloud: An Algorithmic Approach”, Int. J. of Computer Apps, Vol9(9), 2010.

[5] J. Hu, J. Gu, G.Sun, T. Zhao, A Strategy on Load Balancing of VM Resources in Cloud Computing Environment, IEEE 3rd Int. Symposium on Parallel Architectures, Algo &Program. pp.89, 2010.

[6] N. Bobroff, A. Kochut, and K. Beaty, “ Dynamic placement of virtual machines for managing SLA violations”, In Proc. Int. Symposium on Integrated Network Management ’,2007.

[7] D. Barbagallo, E. Di Nitto, D. J. Dubois, and R. Mirandola, “A bio-inspired algorithm for energy optimization in a self-organizing data center,” In Proc. SOAR’09. Berlin, pp. 127–151, 2010.

[8] A. Beloglazovy, R. Buyya,Y. ChoonLee,A. Omaya, “A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems“, Tech. Report, CLOUDS-TR-2010-3, University of Melbourne, 2010.

[9] Nathuji R, Schwan K, “Virtual power: coordinated power management in virtualized enterprise systems”,ACM,pp265, 2007.

[10] A. Singh and M. Hemalatha, "Virtual Machine Placement by Using Honey Bee Forager Algorithm in Cloud Computing", Karpagam Journal of Computer Science, Vol7(4),pp 209, 2013.

[11] B. Addis, D. Ardagna, B. Panicucci, M.Squillante, Li Zhang: “A Hierarchical Approach for the Resource Management of Very Large Cloud Platforms”, IEEE Tran. On Dependable &Secure Comp, 2013.

[12] M. Sharifi, H.Salimi, M. Najafzadeh, “Power-efficient distributed scheduling of virtual machines using workload-aware consolidation techniques“, Journal of Supercomputing, Vol 61(1), pp. 46, 2011.

[13] M. Cardosa, M. R. Korupolu, and A. Singh, “Shares and Utilities based Power Consolidation in Virtualized Server Environments,” in Proc. of IEEE Integrated Network Management (IM),2009.

[14] F. Hermenier, X. Lorca, and J.-M. Menaud, “Entropy: A Consolidation Manager for Clusters,” in Proc.ACM SIGPLAN/ SIGOPS Int. Conf. on Virtual Execution Environments, 2009.

[15] W. Vogels, “Beyond Server Consolidation”, ACM QUEUE, vol. 6, no.1, pp. 20-26, January/February 2008.
[16] D. Nurmi, R. Wolski, and J. Brevik, “VARQ: Virtual Advance Reservations for Queues”, in International Symposium on High Performance Distributed Computing, 2008.
[17] A. Filali, A.S. Hafid, and M. Gendreau, “Adaptive Resources Provisioning for Grid Applications and Services”, in IEEE International Conference on Communications, 2008.
[18] D. Kusic and N. Kandasamy, “Risk-Aware Limited Lookahead Control for Dynamic Resource Provisioning in Enterprise Computing Systems”, in IEEE Int. Conference on Autonomic Computing, 2006.
[19] S. Chaisiri, B. Lee, and D. Niyato. Optimal Virtual Machine Placement across Multiple Cloud Providers, Service Comp. Conf. APSCC 2009
[20] Y. Jayanta Singh, Yumnam Somananda Singh, Ashok Gaikwad, and S. C. Mehrotra, “Dynamic Management of Transactions in Distributed Real- Time Processing System”, International Journal of Database Management System(IJDMS), Vol. 2, No. 2 , pp. 161-170, May, 2010, arXiv preprint arXiv:1005.5435
[21] The SPECpower benchmark website.[Online].2013

Brief about the Authors

Esha Barlaskar is research Scholar at Department of Computer Science & Engineering and Information Technology, School of Technology of Assam Don Bosco University, Assam, India. She has acquired B.E. (Computer Science) degree in the year 2010 from North-Eastern Hill University. Her major research areas are Cloud Computing and Artificial Intelligence. She has published one research paper at National Conference.

Ajith Singh. N completed M.C.A, Ph.D in computer science and currently working as IT consultant under his own firm AjithEx Technologies. Nine years of experience in industry and teaching and published nine international journal paper and four international and one conference publication. Area of research is cloud computing, big data & network. Also national advisor and reviewer in international conference.

Y. Jayanta Singh is working as Associate Prof. and Head of department of Computer Science & Engineering and Information Technology, School of Technology of Assam Don Bosco University, India. He has received Ph.D from Dr Babasaheb Ambedkar Marathwada University, Maharashtra(2004). He has worked with several universities-Swinburne University of Technology (AUS), Misurata University with joint program from Nottingham Trent University, Skyline College University (Dubai) for with joint program from American National University. He has worked with Software developments firms-Keane Inc (Canada and India center), TechMahindra etc. His areas of research interest are Real Time Distributed Database, Cloud Computing, Digital Signal processing, Expert Systems etc. He has published several papers in International and National Journal and Conferences. He is presently executing AICTE sponsored Research Project and consultancy works.