Virtual Machine Consolidation for Data Center with SLA and Energy Efficiency in Cloud Computing

S.R JeyaPraveena
PG Scholar
Department of Computer Science and Engineering
Kongu Engineering College,
TamilNadu, India.

Dr.P.Keerthika
Assistant Professor(Senior Grade)
Department of Computer Science and Engineering
Kongu Engineering College,
TamilNadu, India.

Dr.P.Suresh
Assistant Professor(Senior Grade)
Department of Information Techology
Kongu Engineering College,
TamilNadu, India.

The cloud computing infrastructure provides severalon-demand services. It is a pay-per-use model. A key issue for cloud providers is to maximize their profits by minimizing power consumption along with Service Level Agreement. Dynamic Virtual Machine consolidation is an approach for reducing energy consumption by dynamically adjusting the number of active machines to matched resources in data centers. The Service Level Agreement and energy-efficient dynamic virtual machine consolidation is proposed to meet quality of service expectations and service level agreements requirements. A multi-criteria decision is also proposed to make heuristics for both the determination of under-loaded hosts and migration of virtual machine. Inorder to provide the solution to the consolidation of virtual machines the enhanced optimization policy is introduced. It aims to provide an efficient resource management procedure towards the cloud data center for reducing energy consumption. The heuristics is simulated using cloudsim3.0 simulator. The result is compared with LR and MMT policies and shows a reduction in energy consumption, SLA violation and number of VM migration.

Keywords:**cloud computing, virtual machine consolidation, energy consumption, service level agreement, SLA violation, VM migration, Linear Regression, Minimum Migration Time.**

I. INTRODUCTION

Cloud computing is an internet based computing that provides shared processing resources and data to the computer and other devices on demand. The cloud service providers uses "pay as you go" model, so that is easy for the user to access their required resources anywhere at anytime. The users need to pay only according to their usage of resources. The several cloud services are as follows- Platform as a Service(PaaS), Software as a Service(SaaS), Infrastructure as a Service (IaaS). There are also several types of cloud computing such as Public cloud, Private cloud and Hybrid cloud provides many benefits to the user which leads to enormous increase in population of cloud users.The increased number of computing facilities leads to high energy consumption cost in datacenters.

Virtualization is an enabling technology used to reduce energy consumption of datacenters [5].By separating the virtual machines(VMs) from under-loaded physical machine(PM) of datacenter and allowing to place it on PM where energy consumption improves. This approach is called as Virtual Machine Consolidation. The VMs are consolidated into limited subset of physical resources. So the remaining idle nodes are switched to low power consumption modes or turned-off which reduces the energy consumption [7].

The Virtualization of cloud computing allows multiple VMs on single physical machine as well as migration. VM placement is an NP hard and workload is unstable and it makes dynamic VM consolidation more complicated. The dynamic virtual machine consolidation in cloud data centers are divided into four sections i) determination of under-loaded hosts ii) determination of overloaded hosts iii) selection of VMs that should be migrated from an overloaded hosts iv)finding a new placement of selected VMs for migration from the overloaded and under-loaded hosts[4]. This paper focused on first and fourth section.

The rest of the paper is coordinates as follows. The related work in section II, problem statement in section III, system model in section IV, proposed methodology in section V, simulation results and analysis in section VI and conclusion in section VII.

II. RELATED WORKS

The power management technique for large scale virtualized technique for the first time in addition to VM consolidation and hardware scaling is proposed in [13]. The soft resource scaling is a new power management method proposed for virtualized system. Also that resource management problem is divided into local and global levels. The power management of guest osare monitored using algorithms in local level whereas global policies coordinates multiple PMs. In [2] Abbas Horri et.al has proposed the efficient algorithms for consolidation for reducing energy consumption and SLA violation. Authors in this paper have proposed the efficient SLA aware resource allocation algorithm that considers efficient deal between energy consumption and performance. The proposed resource allocation algorithm considers both host utilization and correlation between resources of VMs present on hosts.

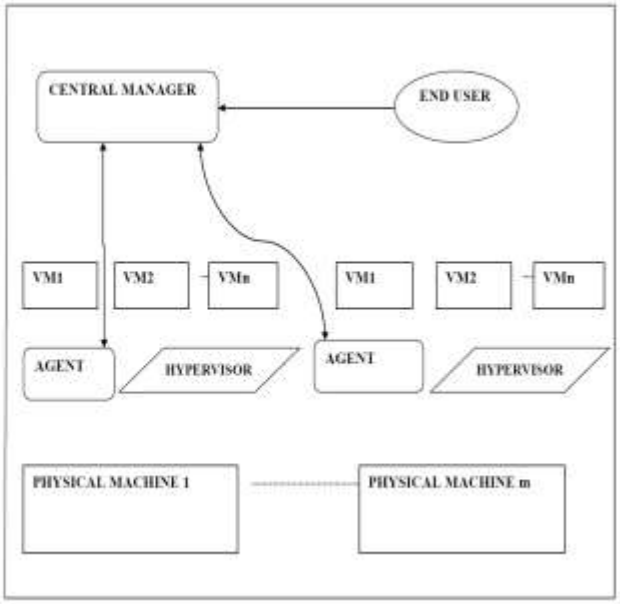
The novel algorithm for determination of under-loaded PMs in cloud datacenters considers both CPU utilization of hosts and number of VMs on the hosts. An architectural framework and principle designed in [3] for energy efficient cloud computing while meeting QoS requirements defined by SLA. Authors divided the VM allocation problem into two parts: 1)admission of new request for VM provisioning and VM Placement on hosts. 2)the optimization of current VM allocation. The authors have modeled MBFD policy for solving bin packaging problem where VMs are sorted in descending order. The authors in [4] have analyzed online deterministic algorithm for single VM migration and dynamic VM consolidation. The dynamic VM consolidation problem is divided into four sections: i) determination of overloaded hosts ii) determination of under-loaded hosts iii) Selection of VMs that should be migrated from the overloaded hosts iv) finding new placement of VMs selected for migration from overloaded and under-loaded hosts. The problem of dynamic VM consolidation has proposed by Jung et.al in [9] is meeting SLA requirements it is difficult to run multi-tier web applications to boost global utility function. The algorithm proposed by Khanna G et.al in [10] has triggered CPU utilization and memory availability to improve the VM placement. In this paper the cost of control is a work accounts for time delays and opportunity cost incurred while switching on/off the hosts and VMs.

III. PROBLEM STATEMENT

In cloud computing environment the major impact is carbon dioxide (CO_2) emission and wastage of energy in cloud datacenters. The hardware infrastructure including servers, storage and network devices are the main portion for energy wastage in cloud datacenters. Forrester Research in [15] states that nearly 30% of hosts consumes power while sitting idle 70% of time. Virtual machine consolidation is an effective approach for better energy consumption so that the remaining idle hosts is switched to low power consumption modes and turned off. Not only an energy wastage problem the cloud environment also faced the VM migration and SLA violation problems due to VM consolidation. Even though there are several benefits VM migration also faces some problems n cloud environment such as migration cost, communication cost and service disruptions. The service level agreement (SLA) is defined as a commitment between service providers and customers. The SLA must also contain penalty when service providers fails to deliver pre-agreed services to the end users at the right time. This is known as SLA violation.

IV. SYSTEM MODEL

The System model contains cloud datacenters with heterogeneous resources for various users.



The System model contains cloud datacenters with heterogeneous resources for various users. The system model [1] in figure 1 has two important parts: i) central manager ii) agent. The central manager is the resource manager which allocates virtual machines to available hosts in cloud datacenters and also it resizes the virtual machines according to their needs on resources. The central manager decides when and which VMs should be migrated from PMs. The agents are connected to central manager through network interface. Agents monitor PMs and send gathered information to the central manager. Hypervisor performs actual resizing and VM migration. To provide Fault Tolerance (FT) and High Availability (HA) the central manager runs on any VMs instead of PMs.

A. POWER AND ENERGY MODEL

The power consumption with CPU utilization by server can be approximated using linear relationship. This approximation comes from the idea that most of the power consumption in cloud datacenter is by CPU. The energy consumption is modeled by Anton Beloglazov in [4] as energy consumed is formulated as follows in equation. That is summation of power consumed during period of time.

$$E(t) = \sum_{t} P(t)dt \qquad (1)$$

B. SLA VIOLATION METRICS

The QoS requirements are formulated in the form of SLAs in cloud datacenters. The SLA violation metrics is defined as the multiplication of two metrics as formulated in below equation 2,

$$SLAV = SLATAH \times PDM \qquad (2)$$

Where SLATAH is SLAV Time per Active hosts and PDM is defined as Performance Degradation due to Migration.

V. PROPOSED METHODOLOGY

The dynamic virtualization is an efficient way to reduce energy consumption in cloud datacenter. The important aspects of this method are optimization of resource utilization and performance on energy consumption. The Enhanced Optimization (EO) is the proposed resource allocation policy used here. EO suggests gathering all the VMs to be migrated from either overloaded or under loaded PMs in the VMs migration list and solving the dynamic resource allocation problem.

A. DETERMINATION OF UNDER-LOADED HOSTS

The under-loaded hosts are determined using multicriteria policy (MCP). The multicriteria policy is a decision making process considers three criteria such as Available Capacity (AC), Number of VMs (NV) and Migration Delay (MD) as shown in below Table I. This policy takes advantage of Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). This policy evaluates the score for all candidate hosts and selects the hosts with highest score as under-loaded.

TABLE I: Criteria considered in MCP

NO	NOTATION	PARAMETE R	BENEFIT/ COST

1	AC	Available Capacity	Cost
2	NV	Number of VMs	Cost
3	MD	Migration Delay	Cost

The under-loaded hosts are selected by Multicriteria policy if the following constraint exists: i) the selected hosts should have minimum available capacity, ii) the selected hosts should have minimum number of virtual machines, iii) the selected hosts should have minimum number of migration delay for all VMs[1]. The pseudo code for TACND policy is shown in below algorithm I.

INPUT: Candidate PMs to be under-loaded
OUTPUT: Under-loaded hosts to be found
For all candidate PMs
Decision Matrix= Criteria of all PMs
Divide each entries with maximum values of each column for Normalize decision matrix
Determine PM^+ and PM^-

$$Score^{PM^j}_{criterion}= \frac{\sqrt{\left(PM^j_{criterion}-PM^-_{criterion}\right)^2}}{\sqrt{\left(PM^j_{criterion}-PM^-_{criterion}\right)^2}+\sqrt{\left(PM^j_{criterion}-PM^+_{criterion}\right)^2}}$$

$$Score(PM^j)= \sum_{Criterion=1}^{\#Criterion} Weight_{criterion} * Score^{PM^j}_{criterion}$$

Rank the PMs
If Rank(PM==1)
Under-loaded PM= This PM
END
RETURN Under-loaded PM

ALGORITHM I: Multicriteria policy (MCP)

B. PLACEMENT OF VMs

The VM placement after evaluation of an under-loaded hosts are determined using Power and SLA Aware(PSLA) resource allocation policy which takes advantage of TOPSIS method. The criteria considered for this decision process [6] are depicted in Table II. This policy computes the score for all candidate VMs and selects the hosts with highest score as destination hosts for VM placement.

TABLE II: Criteria considered in PSLA policy

NO	NOTATION	PARAMETER	BENEFIT/COST
1	PI	Power Increase	Cost
2	AC	Available Capacity	Benefit
3	NV	Number of VMs	Cost
4	RC	Resource Correlation	Cost
5	MD	Migration Delay	Cost

PSLA computes score hosts based on the following constraint: i)the selected hosts should have minimum power increase, ii) the selected hosts should have maximum available capacity, iii) the selected hosts should have minimum number of VMs, iv) the selected hosts should have minimum number of resource correlated VMs. v) the selected hosts should have minimum number of migration delay.

In PSLA method the chosen destination has minimum distance from ideal positive point (PM^+) and maximum distance from ideal negative point(PM^-). The distance between each of these points are measured using the Euclidean distance.

VI. SIMULATION RESULTS

The system model is general cloud computing environment on virtualized cloud datacenter infrastructure. The cloudsim 3.0.3 toolkit is used for simulating and measuring the performance of large-scale virtualized datacenter. The infrastructure has datacenters with 800 installed heterogeneous hosts and five types of VMs. The characteristics of VM and 10 days planetlab workload in [1] are taken as a dataset for simulation.

Performance Metrics

ESM is the metrics used to measure simultaneous improvement in energy consumption, SLA violation and Number of VM migration as expressed in equation (3) below.

$$ESM = Energy\ consumption * SLA\ violation * VM\ migration \quad (3)$$

A. Energy Consumptions

The energy consumption between LR\MMT and EO policy is shown in figure 2. The proposed EO policy gives better performance over energy consumption.

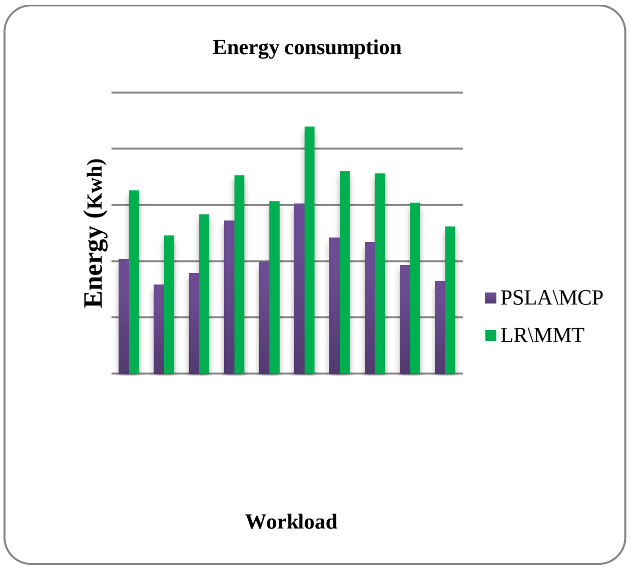


Figure 2 Energy consumption of resource allocation policy

B. Number of VM Migrations

The number of VM migration between LR/MMT policy and EO policy is shown in figure 3. The proposed EO policy gives better performance over VM migration.

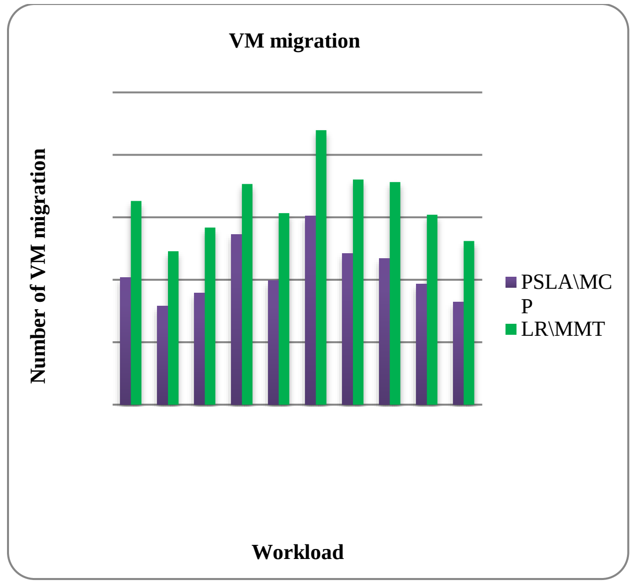


Figure 3 Number of VM migration and resource allocation policy

C. SLA Violation

The SLA violation comparison between LR/MMT policy and EO policy is shown in figure 3. The proposed policies give better performance over SLA violation.

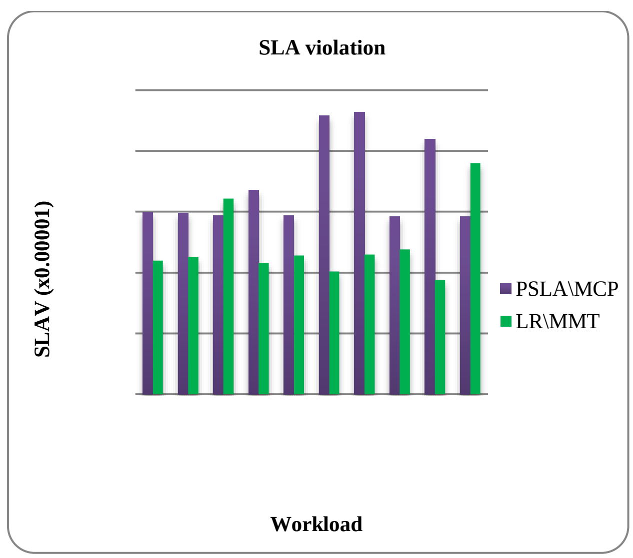


Figure 4 SLA violation of resource allocation policy

The above are the several performance metrics used and the experimental resultant graph drawn for an output obtained when comparing the proposed EO policy and with traditional LR\MMT policy.

VII. CONCLUSION

The huge datacenters all around the world leads to enormous energy consumption and carbon dioxide emission. It is more essential to reduce energy consumption with SLA violation in data centers. The dynamic VM consolidation is the resource management strategy used to reduce energy consumption and SLA violation and also switched off the hosts into sleep mode. The proposed PSLA and MCP policies provide maximum user satisfaction with reduced energy consumption, SLA violation and number of VM Migration. The PSLA policy for resource allocation and MCP for determination of under-loaded hosts results in best performance shows overall 99% reduction in ESM metrics.

REFERENCES

[1] EhsanArianyan, Hassan Taheri, SaeedSharifian. "Novel energy and SLA efficient resource management heuristics for consolidation of virtual machines in cloud data centers".ComputElectrEng 2015.

[2] Abbas Horri, Mohammad SadeghMozafari(2014), 'Novel resource allocation algorithms to performance and energy efficiency in cloud computing'. J SupercomputVol. 69, pp. 1445-61.

[3] Anton Beloglazov, Abawajy J, Buyya R.(2012), 'Energy aware resource allocation heuristics for efficient management of data centers for cloud computing'. FutureGenerComputSyst ;28:755–68.

[4] Beloglazov A, Buyya R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in Cloud data centers. ConcurrComput: PractExper 2012;24:1397–420.

[5] Buyya R, Ranjan R, Calheiros RN. Modeling and simulation of scalable cloud computing environments and the CloudSim toolkit: challenges and opportunities. In: High performance

[6] Chen C-T. "Extensions of the TOPSIS for group decision making under fuzzy environment". Fuzzy Sets Syst 2000;114:1–9.

[7] Gao Y, Guan H, Qi Z, Song T, Huan F, Liu L. Service level agreement based energy-efficient resource management in cloud data centers. ComputElectrEng 2013.

[8] Jeyarani R, Nagaveni N, Vasanth Ram R. Design and implementation of adaptive power-aware virtual machine provisioner (APA-VMP) using swarm intelligence. Future GenerComputSyst 2012;28:811–21.

[9] Jung .G, M. A. Hiltunen, K. R. Joshi, R. D. Schlichting, and C. Pu, "Mistral: Dynamically managing power, performance, and adaptation cost in Cloud infrastructures," in Proc. of the 30th Intl. Conf. on Distributed Computing Systems (ICDCS), 2010, pp. 62–73.

[10] Khanna, G., Beaty, K., Kar, G., Kochut,(2006), 'Application performance management in virtualized server environments'. In: Proc.of the IEEE Network Ops. and Mgmt. Symp., pp. 373–381.

[11] Luo J-P, Li X, Chen M-R. Hybrid shuffled frog leaping algorithm for energy-efficient dynamic consolidation of virtual machines in cloud data centers. Expert SystAppl 2014;41:5804–16.

[12] Maurer M, Brandic I, Sakellariou R. Adaptive resource configuration for cloud infrastructure management. Future GenerComputSyst 2013;29:472–87.

[13] Nathuji R, Schwan K., (2007), 'VirtualPower: coordinated power management in virtualized enterprise systems'. In ACM SIGOPS Operating Syst Rev, Vol. 41, pp. 265-78.

[14] Park K, Pai VS. CoMon: a mostly-scalable monitoring system for PlanetLab. ACM SIGOPS OperSyst Rev 2006;40:65–74.

[15] Poess M, Nambiar RO. Energy cost, the key challenge of today's data centers: a power consumption analysis of TPC-C results. Proc VLDB Endowment 2008;1:1229–40.