Abstract: In the present scenario, cloud computing environment grants all the resources in scalable manner to every users in pay-per-use processing model over the Internet through various data centers. An energy consumption of these resources have to be addressed in many issues in the cloud. A key strategy of virtual machine (VM) management is a live VM migration in data center networks. One of the significant problems of cloud provider is the energy cost. VM migration and placement has been shown as an efficient approach for energy saving. In this paper, we are proposing an algorithm, Modified Energy Conscious Greeny Cloud Dynamic Algorithm (MECGCD), goes for preventing unnecessary traffic in a datacenter network, and excessive energy consumption (EC) started from wrong routing management and improper VM allocation. In this paper, we observe at the issue of how to choose the host for VM placement and to migrate VMs from abnormal loaded hosts such as under loaded or over loaded to another and switching off the idle host machine into sleep mode. VM placement be determined the host machines by shortest distance, minimum EC and maximum bandwidth usage in the cloud environment. The evaluation of experiments confirmed that the proposed algorithm minimizes EC and network traffic in a cloud data center in a quotable manner than other existing algorithms.

Keywords: Cloud Computing, Haversine, Data Center, Live VM Migration, Energy Consumption.

1. INTRODUCTION

The distributed computing such as Cloud computing provides the virtualized resources, for example, software applications, storage and computing power to the clients on the internet in whole year. By utilizing service level agreements (SLA), the cloud infrastructure provider provides the Quality of service (QoS), and the resource pool is provisioned as high availability. Platform as a Service (PaaS), Software as a Service (SaaS) and Infrastructure as a Service (IaaS) [1] are the services utilized in the cloud.

In resource provisioning, the active host count is to be moderated by the best possible VM allocation which diminishes the EC in cloud data centre [2]. To diminish the EC on a decreased set of physical servers to sleep mode or switch off idle servers [3] the dynamic VM consolidation is utilized. A suitable VM position stimulates superior resource utilization along with lower cost. On the whole, considering the quantity of VMs tenants present their resource prerequisites and without any regarding of network bandwidth constraints [4], cloud data centres instantaneously assign these resources.

In any case, the higher part of enterprises, cloud server farms utilize little few number of their accessible resources when an extraordinary amount of their energy usage is vanished as a result of the two ineffective resources and over-provisioned [5]. The infrastructure and energy expenses allowed about 76%, as well as the IT, allowed about 27% of expense in the data centre in 2016 [6-7].

Green cloud is a buzzword that project to attain not only the proficient processing along with utilization of computing resources, but also diminish the EC. It refers to the eventual environmental benefits, that Information Technology (IT) services dealt out over the web which can be offered to the society. The term connects the two different terms – green computing which means strong revolt towards a more environmentally widen computing or IT, by means of varied studies, practices as well as scrutinise of proficient and eco-friendly computing commencements – and cloud computing which mentions the vital sign for the Internet and the defined title for a category of on demand self service deployment model. Besides, cloud resources are allocated not only to moderate energy usage, but also support QoS necessities required by clients by way of Service Level Agreement (SLA) [8].

Virtualization is the foremost innovation in cloud computing which operates as a chief tool for eliminating energy inefficiency by placing several Virtual Machines (VM) in a single Physical Machine (PM) through live VM migration strategies. VM consolidation is a technique to reduce the number of working PMs by migrating and consolidating the VMs into reduced set of PMs [9]. VM consolidation refers VM placement (the course of selecting the proper host for the given VM) and VM migration (is the task of shifting a VM from one unhealthy physical hardware environment to another).

With the intension of deciding VM migration is firmly detected while a server is overloaded or under loaded [10]. To select whether a host is either overloaded or under loaded is the primary hazardous because of the different categories of client applications and the instability of the VM workloads with time and especially in a cloud server farm with a significant amount of heterogeneous machines [11]. In this situation, a small number of VM solidification plans to take the current benefit of a single resource (i.e., CPU) into account when selecting whether a physical server is over-burden or underutilized.

A.C. Subhajini, Assistant Professor, Department of Computer Applications, Noorul Islam Centre for Higher Education, Noorul Islam University, Kumaracoil, Tamilnadu, India.

A.V. Sajitha, Research Scholar, Department of Computer Applications, Noorul Islam Centre for Higher Education, Noorul Islam University, Kumaracoil, Tamilnadu, India.

Manuscript received January 25, 2019.

Published By: Blue Eyes Intelligence Engineering & Sciences Publication

Retrieval Number: ES2013017519/19©BEIESP
Network-Conscious VM Placement for Energy Efficiency in Green Data Centres through Dynamic VM Consolidation

The implementation of the facilitated applications are not just impacted by the migration choices still and moreover that of the server farm in common. For power decrease, server switch off and VM migrations are essential in this manner and to avert large movements and limit power state switches [12] is not easy.

To minimize the EC and shrink the number of migrations we advised a shortest distance with congestion free path for host discovery algorithm in this paper. The given information are the VM consolidation dilemma has been classified into four sub problems and directed in queries.

- Acquire the overloaded host and migrate few VMs from the obtained data center to other data centers.
- Acquire the under loaded host and migrate all the VMs to other optimal host and put the under loaded host in sleeping mode.
- Determine the VM(s) which are need to be migrated (selection of VM).
- Position the opted VMs on best host (minimum EC, shortest distance host, maximum bandwidth).

From the different categories of hosts, our focus is on choosing the most excellent host for VM placement in this paper.

II. RELATED WORK

By unravelling different migrations from a particular PM within a time duration where the traffic due to network acquiring a leading place, that offers a testing and acute predicament, to such a degree that the cloud framework won’t have adequate resources openly available to meet the expansion in resources required. Consequently, it will set aside a more extended time for a variety of movements to be done, which thus prompts execution abseasments for VMs. Thus, a PM is reserved as an under-utilized or over-utilized status for a required situation, persuading its Service Level Agreement(SLA), that is an agreement provided by the cloud supplier and user depicting the reaction time as well as the throughput guarantees [13-14].

Farahnakian,F. et al. [15] had produced a dynamic Virtual Machine (VM) Consolidation with the utilization prediction model algorithm which decreases SLA violations and ejected worthless VM migrations. From physical hosts, this proposition expected to move few VMs which were over-burden or expected to complete plainly over-burden in future. According to present and future resource prerequisites, additionally, it assigned a VM to a host.

Paya,A. et al. [16] have been presented an energy-aware operation display which was exercised for stack modifying and application, scaling on a cloud. Characterized an optional energy operation administration and attempted to raise the number of servers functioning in this framework was the underlying logic of their approach. The servers like inactive and lightly loaded were transferred to any one of the rest states to protect energy.

Farahnakian,F.et al. [17] have been suggested an algorithm based on Ant Colony System named Ant Colony System VM Consolidation (ACS-VMC). By utilizing the minimization of host count, the EC was decreased through this methodology while maintaining the QoS burdens. They described a multi-objective work that considers both the quantity of stagnant PMs and the migration count.

Gai,K. et al. [18] have been focused a Dynamic Energy-aware Cloudlet-based mobile cloud computing Model (DECM) for blocking the excess energy consumption by using dynamic cloudlets (DCL). Through a renovation of sensible condition they formulated in their model and provided strict effects to the evaluations. Twofold was the principle commitments of this paper. The original enquiry of this work was to take care of energy consumption problems inside the dynamic systems administration condition.

Rao,K.S. et al. [19] have been focused on investigating heuristics, in which residual resource fragmentation could reduce to make other assets highly precious and diminish vitality usage and value carry to the cloud data center.

Li,H.,Zhu,G et al. [20] have been designed for the server loading status threshold through multi-resource utilization, a procedure to encourage the VM migration. In conventional heuristic calculations, the Modified Particle Swarm Optimization method was initiated within the combination of VMs to avert dropping into local optima was usually defect.

Kansal,N.J et al. [21] have been suggested for distributed computing, an energy-oriented VM migration strategy on the basis of the firefly algorithm. During holding on the implementation and power competence of the data centres, the suggested practice had been moved the supremely full virtual machine to the minimum loaded dynamic.

Son,J et al. [22] have been suggested dynamic overbooking algorithm for VM and traffic combination that jointly utilize the virtualization abilities and SDN. The process allotted new accurate measure of assets through the dynamically developing workload to VMs and traffics. In a host and architecture, this process could increase overbooking during inactive state also; adequate assets offer to restrict SLA infringement. From the online enquiry of the host and system utilization through no pre-learning of workloads their approach calculated asset assignment proportion considering the essential observing data.

Khosravi,A et al. [23] have been evaluated the unique energy and carbon-aware dynamic VM position algorithm. In geologically dispersed cloud data centres for the underlying arrangements of VMs is a new tactic which dealt with the amount of projected energy, servers’ energy as well as CO₂ footprint. Moreover, the stable power source was enhanced by the suggested VM arrangement strategy and utilize it at every datum concentrate to control the aggregate cost. The productive two-phase VM position approaches were shown by them which answer to dynamic PUEs. For a distributed computing environment, in controlling the energy consumption and carbon cost, they displayed a set of techniques that evaluated the influences of many parameters.

Duggan,M. et al. [24] have been suggested an autonomous network-aware VM migration policy which observed the existing appeal stage of a system as well as executed appropriate actions in consideration of what it would be encountered.

Published By: 
Blue Eyes Intelligence Engineering 
& Sciences Publication
The Artificial Intelligence process was considered as Reinforcement Learning, and it was going an option of the emotionally supportive network, authorizing an operator to take the perfect time to plan a virtual machine movement depending on the current network traffic appeal.

III. PROPOSED WORK

3.1. The system architecture

There are \( m \) number of hosts are presented in the cloud data center and the hosts can be described as Host \( H = \{h_1, h_2,..., h_m\} \) which is represented in Fig 1. The data center is isolated into \( R \) regions, and every region has Count Hosts (total number of hosts, in all regions). Therefore, \( R = m \times \text{Count} \). We express the path of network amongst two different hosts as the network hops connecting them, which is illustrated in the below equation 1 as:

\[
l(h_p, h_q) = \begin{cases} 1, & h_p - h_q > 0 \\ l(h_p, h_q) + 1, & h_p - h_q \leq 0 \end{cases}
\]

The set of regions is defined as \( O = \{o_1, o_2,..., o_1, o_M\} \). The set of hosts in a region \( O_x \) be represented by \( H_x = \{h^x_1, h^x_2,..., h^x_{\text{Count}}\} \) and \( VM^x_i = \{vm^x_1, vm^x_2, ..., vm^x_m\} \) represents the number of VMs in a host \( h^x_i \) that belongs to \( O_x \). Each host \( i \) is defined by the CPU utilization which is measured in the amount of memory, the bandwidth (BW) and Million Instructions Per Second (MIPS).

![Fig 1: System model](image)

In the beginning, the VMs are placed in a host arbitrarily whereas in descending order VMs are arranged with their resource utilization. Occasionally, owing to the dynamic work-loads, the resource usages of VMs could be different. Through a VM combination algorithm, the starting effective allotment approach should be prolonged in this process. The suggested framework show is depicted in Fig 1, which is consisting of two varieties of agents: such as the local and global agent. By noticing the present resource usages of the host, the host status discovering sub-issue can be solved through a Local Agent (LA) which stays in a host. By utilizing the suggested algorithm, the Global Agent (GA) functions as an administrator and develops the VM placement.

3.2 Energy model

The required steps can be taken for the EC and it is measured in Joule. We utilize few standard in EC that is known as the model. By a linear relationship of CPU utilization, through physical machines, the power usages can be accurately portrayed. They showed that a free host utilizes 80% of its EC while it is used as a whole. Therefore, as a CPU usage task, we can characterize power utilization. In equation 2, the function is displayed as:

\[
pwr(u) = z.pwr_{\text{max}} + (1 - z).pwr_{\text{max}} u \quad (2)
\]

\( pwr_{\text{max}} \) is the ultimate power of a host in the operating state. \( z \) is the proportion of power devoured by an inactive host. The utilization of CPU is represented by \( u \). For the reason that, the utilization of a CPU varies in due course, we characterize it as a function \( u(t) \) of time. Thus, aggregate EC can be attained by equation 3.

\[
\text{Energy} = \int_{t} Pwr(u(t))dt \quad (3)
\]

Based on this function, the EC of a host is decided by the utilization of CPU.
3.3 Methodology

VM consolidation is to be formulated for more competent use of computing resources and avoid servers as well as storage units from being underutilized. The VM consolidation helps to lessen the amount of servers in the cloud datacentre organization. It aims to reduce the wastage of a large space by optimal resource utilization as well as higher energy efficiency in the data centre.

(i) VM Selection

The major sub problem of VM consolidation is tried to solve in this segment. The issue of this dilemma is to make a decision whether a host is treated as overloaded, prone to be overloaded or under loaded and when the VMs to be migrated from the diseased host(s). As afore mentioned, there are several tactics have been suggested to select the occasion of migration of the VMs to avoid a probable SLA violation. For healthy host utilization, the most commonly used methods is to put upper as well as lower usage thresholds. Then all the VMs of these hosts must be kept the aggregate utilization among these thresholds. If the CPU usage surpasses the thresholds, it will call upon the VM selection and VM placement policies. These host categorization and threshold value calculation is done in our previous algorithm ECGCD[25].

Algorithm 1: VM consolidation approach

Input: No. of host H; Interface with cloud data
VM is designed and allocated to hosts;
Workload is assigned to VM;
For each period t;
Identify overloaded and under loaded host by threshold;
If host
\[ \geq \text{bandwidth} \& \& \leq \text{energy consumption} \& \& \leq \text{dis tan ce} \]
then
VM is placed in available datacenters;

(ii) Best host selection for VM placement

To find the best host we have to find the below objectives:
- Shortest distance host
- Minimum EC
- Bandwidth utilization of each host

Shortest distance host

In the topology, from every source host to each other host we will compute the shortest path in this division. On or after the source host to target host for discovering the shortest distance, by utilizing the haversine formula we measure the distance between every pair of coordinates, connecting hosts and target locations.

\[ d = 2r \arcsin \left( \sqrt{\sin^2 \left( \frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left( \frac{\lambda_2 - \lambda_1}{2} \right)} \right) \]

(4)

By selecting accessible connections of data centre nodes, the latency between two points source host \( (I_i) \) and destination host \( (I_j) \) is assessed which minimize the response time. ‘d’ is the distance between longitude and latitude points. The latency between the two locations is assessed utilizing the following formula:

\[ \text{latency}(I_1, I_2) = \sum \frac{d_{\text{min}}}{d_i} \]

(5)

Where, \( d_i \) is the overall distance of a pair \( i \) between the source location and chosen data centre nodes respectively. The smallest distance among the accessible pairs is denoted as \( d_{\text{min}} \) and the measured latency between the respective locations \( I_1 \) and \( I_2 \).

Energy consumption

The anticipated cost of energy is computed by algorithm 2 as a result of the candidate allotment while the time interval t. An energy model ought to be utilized for calculating power utilization which is specified above.

Algorithm 2: obtain total energy cost

Input: Estimated energy cost (z, t)
Output: Total energy cost of running host;
Initialization;
Energy consumed per host[s] = 0;
Total energy consumed=0;
Total energy cost=0;
All host utilizations [] =Estimated utilization (z);
For each host in the assignment z do
Energy consumed per host [host] =host. getPowerModel() * get power (CPU utilization [host]);
Total energy consumed += Energy consumed per host [host];
End
Total energy cost=total energy consumed*energy cost per second * t;
Return total energy cost;

In the stated time period t, the cost of the total energy utilized is the value return from this algorithm because of the provision of resources in the cloud data centre.

Bandwidth utilization rate for every host

The whole VMs determine the hosts’ utilization rate of the bandwidth itself. The below equation obtains the bandwidth utilization rate of \( h_i \) at a particular point,

\[ L_i = \frac{\sum b_{w_i}}{B_{W_i}} \]

(6)

From that, \( B_{W_i} \) defines the total bandwidth of \( h_i \). \( n \) is the set of VMs in each host. Bandwidth used by \( v_{m_j} \) is represented as \( b_{w_j} \). Every host have different number of VMs and also the executing applications in the VMs are varying. So there is variation among the different hosts in the utilization of bandwidth rate. The bandwidth utilization rate \( h_i^* \) which belongs to \( O_3 \) can be described in below equation,

\[ L_i^* = \frac{\sum b_{w_j}^*}{B_{W_i}^*} \]

(7)
So, the load of average, bandwidth, utilization rate of the hosts in $O_4$:

$$\text{Load} = \frac{\sum_{j=1}^{\text{Count}} L_i^x}{\text{Count}} \quad (8)$$

If the host satisfying the above mentioned objectives then we select the host is the best host. Then the particular VM which is elected for migration will be placed on the selected host.

IV. IMPLEMENTATION

A. Workload data

To accomplish our goals for this movement, we are having concentrated on BitBrains data set containing workload traces for 1750 VMs from its distributed data centres corresponding to host and business-critical enterprise applications. The trace is comprising of one file for each VM, portraying generally the VM’s dynamic workload, sampling in every 5 minutes. (2nd August, 2013 preferred for this work). These data comprise the VM’s both utilized CPU and main memory at the assured time period, along with disk as well as network I/O throughput values. Each data centre traffic differs hour wise, day wise, week wise, and month wise. Traffic behaviour of a server could be enabled us to find patterns of transformations that can be abused for more proficient resource allocations. The network traffic data utilized in this work is obtained from the website of University of Wisconsin-Madison’s University data centre - UNI2 data centre [26] traces are selected for the data centre traffic experimental simulation. The starting timestamp of the data utilized in this work is 00:00:00 and the end timestamp is 23:59:59 of whole day.

**Table 1: Electricity consumption of selected servers for dissimilar workloads (in Watt)**

<table>
<thead>
<tr>
<th>Server</th>
<th>0%</th>
<th>1%</th>
<th>2%</th>
<th>3%</th>
<th>4%</th>
<th>5%</th>
<th>6%</th>
<th>7%</th>
<th>8%</th>
<th>9%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP Proliant G4</td>
<td>8</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>HP Proliant G5</td>
<td>9</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>13</td>
</tr>
</tbody>
</table>

**Table 2: Host Parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Server</th>
<th>HP ProliantG4</th>
<th>HP ProliantG5</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of host</td>
<td>400</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>No of Cores</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MIPS</td>
<td>1860</td>
<td>2660</td>
<td></td>
</tr>
<tr>
<td>RAM</td>
<td>4096</td>
<td>4096</td>
<td></td>
</tr>
<tr>
<td>BW</td>
<td>1GB</td>
<td>1GB</td>
<td></td>
</tr>
<tr>
<td>Storage</td>
<td>1.5GB</td>
<td>2GB</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3: VM Parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>High-CPU Medium Instance</th>
<th>Extra Large Instance</th>
<th>Small Instance</th>
<th>Micro Instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Cores</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MIPS</td>
<td>2500</td>
<td>2000</td>
<td>1000</td>
<td>500</td>
</tr>
<tr>
<td>RAM</td>
<td>870</td>
<td>1740</td>
<td>1740</td>
<td>613</td>
</tr>
<tr>
<td>BW</td>
<td>1MB</td>
<td>1MB</td>
<td>1MB</td>
<td>1MB</td>
</tr>
<tr>
<td>Storage</td>
<td>3.85GB</td>
<td>2GB</td>
<td>1.75GB</td>
<td>613MB</td>
</tr>
</tbody>
</table>

B. Experimental setup

We have intended a large-scale IaaS environment which provides remarkable computing resources for its clients. For the evaluation the efficacy of our experimentation, executions have been carrying out on the CloudSim simulation toolkit, which is becoming more and more established because of their supporting for elastic, scalable, proficient and iterated assessment of resource provisioning practices for a variety of applications[27]. It simulates the workload of host machines after VM deployment. Besides, it can also calculate the EC. We have been created te simulation of a cloud computing architecture consisting of a data centre through 800 heterogeneous PMs equipped with 2 different categories. Features of these devices are illustrated in Table 1, host parameters are depicted in Table 2 and VM parameters are portrayed in Table 3.

C. Performance evaluation metrics

Energy consumption

We are considering the aggregate EC by the physical machines of a data centre grounded by load of works assigned to it. The EC of hosts relies upon the CPU utilization, main memory, storage and network routing. The CPU devours more electricity usage than other devices namely, hard-disk storage, network I/O and main memory. So the utilization of host’s resources is generally corresponded by its CPU utilization.

**Fig. 2: Analysis of energy consumption**
Migration time

It defines the total time to do all VM migrations based on the migration plan.

Figure 2 to 5 represents the performance evaluation of proposed algorithm MECGCD comparing with existing algorithms LUPDATE [28] RL- Agent Actions [29], and ECGCD [30].

LUPDATE strived to select a suitable path from all potential shortest paths for each new flow. The worst situation is that if a new flow cannot be assigned to any candidate path effectively, that flow should wait for an infinite time on the ingress switch until any candidate path becomes ready for traffic of the VM Migration because of the completion of various active flows. It raises the chance of total migration time increase and creates a network traffic congestion because the flow can be wait for an infinite time to get a feasible path. As a consequence, it is generating a high energy consumption.

RL- Agent Actions, formulated an independent learning agent with the potentialities to fix on a suitable time period to schedule the migration of VMs as of an overutilized host by monitoring the present bandwidth accessibility in cloud data centres. While considering the over-utilized hosts with static threshold value of CPU utilization, they were completely ignoring the lion portion occupier of energy consumption even at their idle state such as under-utilized hosts. They were opting the VMs on the basis of its RAM utilization using the method MMT. They were migrating the VMs only when the network traffic is low, it can be generated the congestion as well as a raise of network traffic at peak times.

ECGCD, our previous algorithm was giving concentration of both under and over utilized hosts with dynamic threshold value of load balancing techniques while considering 3 factors such as CPU, RAM and bandwidth in the cogent basement of least power consumption of hosts. Even though, it was not giving importance to shortest distance or minimum bandwidth. It was giving emphasis to least power consumed hosts. So it may be the cause of generating a heavy traffic in network at peak times and as a reason for increase of total migration time.

In this proposed system – MECGCD, providing equal importance to least power consumption of the host, shortest path with minimum latency and finest bandwidth of the proposed hosts. So, it achieves better performance than the existing algorithms.

V. CONCLUSION

In the aforementioned work, we have been proposing a network-aware energy-efficient dynamic VM consolidation approach, MECGCD that formulates a VM consolidation as a multi-objective problem. The utilization of resources are assessed and migrated VMs are chosen using threshold concept. The multi objective criteria - minimum power consumption, shortest distance and highest bandwidth - obtains the optimal host for VM placement. The optimal host is detected from the host list. Bitbrains data set is used for the simulation, and the performances are appraised and compared with the existing algorithms.
As a future endeavour, we will extend our approach with additional enhancement of SLA violation minimization as well as improvement of scalability of this approach.

ACKNOWLEDGEMENT

The authors would like to acknowledge and thank all people who help them in different ways to fulfill their research work.

REFERENCES