Cluster’s Node Emigration using Cloud’s Adaptive Power Panel

Nour M. Azmy
BMB Solutions
Cairo, Egypt

Islam A.M. El-Maddah
Ain Shams University
Cairo, Egypt

Hoda K. Mohamed
Ain Shams University
Cairo, Egypt

(9580, islam.elmadah, hoda.korashy)@eng.asu.edu.eg

Abstract- Cloud computing creates new business opportunity for small, medium or even large size companies, as it eliminates the need of investing in a huge datacenter in order to operate the environment. The contract between the cloud service provider (CSP) and the lessee is called service level agreement (SLA), which guarantee the quality of service (QoS) that should be fulfilled from the CSP to the lessee. The problem is that an over utilized node may result to SLA violation; However, virtual machine (VM) migration from utilized node to unutilized node or a node in power saving mode can decrease the SLA violation percentage, but potentially increase the cloud’s power consumption. Our proposed solution is to automatically modify the number of nodes in each cluster, to find the optimum cluster size based on the current SLA violation percentage, and to be an extension to our previously proposed adaptive power panel, in which the cloud’s datacenter is divided into clusters, where different VM migration policies are applied.

Keywords- Cloud computing, Green computing, SLA, VM migration, Power Panel, Cloudsim, Fuzzy Logic, Node Migration.

I. INTRODUCTION

Cloud computing is the future of the system and network models; Moreover, the current direction is to move the current working environment of both companies and even the individuals to the cloud. Cloud has three deployment models [1], which are private, public, and hybrid, which is connecting on premises private datacenter with the leased resources from the CSP. The world is aiming for cloud maturity, as from the cloud’s user perspective, it eliminates the need of purchasing expensive infrastructure, a location for the private datacenter, troubleshooting physical problems, and the mandatory upgrade for conveying the rapid market change. Scalability empowers the cloud, simply by adding more nodes to the cloud’s datacenter, in which CSP can accommodate more lessees.

Cloud’s main engine is the virtualization [2], where the operating systems are not directly installed on the physical node. But a hypervisor is the first platform on the node, then other operating systems can be installed on that platform. Fig. 1, [3] describes how the hypervisor acts as a broker that handle the CPU scheduling, memory management, the packets queues in network cards, and disk management between the physical resources of each node and the VMs. The virtualization approach is beneficial for better resource utilization, where each particular VM can use allocated resources, or even can share resources with other VMs.

Figure (1): Hypervisor handle the communication between Node’s Resources and VMs

Single node can handle many VMs, which may result in SLA violation. Since VMs can share resources of one node, some VMs may face resource starvation. Lessee from the CSP can accept a specific agreed percentage of SLA violation based on a SLA contract between the actors, but any further violation may result in penalties.

Key factor for decreasing the SLA Violation is to Migrate VMs from an over utilized node to other node. That task may require extra nodes to be powered on, which will potentially increase the power consumption. So tradeoffs between SLA violation and power consumption is studied farther in this paper. CSP aims for optimum solution in which to minimize the cloud’s power consumption as well as the SLA violation percentage.

In this paper we survey the VM migration policies that is done by VM allocation policies, then VM selection policies, and finally VM placement policies, and we survey also node migration policies that will be added to the Adaptive Power Panel that we discussed in [4]. Related work is briefly mentioned in Section II. In Section III different Policies are discussed. Section IV introduces our contribution in the automatic cluster sizing of the cloud nodes. Then our Evaluation and Results are represented in Section V. And Finally Section VI shows our conclusion and the future work.

II. RELATED WORK

In [4] we had proposed an adaptive power panel. The idea was that applying a single Migration policy on all of the VMs of the Cloud’s datacenter will decrease the flexibility with the customers’ need. So the adaptive power panel first step is to divide the datacenter into many clusters, where an
independent VM migration policy is applied to each cluster not the datacenter. That had introduced three modes for each cluster. Power aware mode, balanced mode, and high performance mode, where a tradeoff between power consumption and SLA violation is adjusted for each cluster to accommodate more customers’ requirements, which will enhance the cloud flexibility. But the policy that role a node motion from a cluster to another cluster wasn’t discussed.

In [5], Ziming Zhang et al. introduced an adaptive power management framework, in which the main engine is a power estimation model that estimates the power consumption from the VM correlation with the physical node, and the VM configuration. Then apply new configuration VM configuration to reduce the power consumption with a minimal SLA violation as possible. But without VM migration, that framework didn’t solve case where a host is highly utilized, so even the new VM configuration will not reduce the power consumption, nor the SLA violation.

In [6], S.R. Hussein et al. survey the building blocks of VM migration policies i.e. VM allocation, VM selection, and VM placement Policies. Also introduced new VM allocation, selection, and placement fuzzy logic policies. Again applying only one policy to the cloud’s datacenter, which is not sufficient for conveying wide variety of customer requirements. Regardless that some of the results are not accurate, But the SLA violation had increased by around 35%, which is not applicable for any of today’s cloud providers.

In [7], Seokho Son et al. create a framework that divides the SLA operation into two phases. Phase 1 is SLA Establishment (SLA-E), that handle the agreement between the CSP, and the Lessee. While phase 2 is SLA Management (SLA-M) that manage the load balancing, and actual available resources between the customers and the CSP. But managing the resources of the cloud’s datacenter as one block is not a power aware solution. As the main and only concern is to satisfy the SLA regardless the power consumption which is not an optimum solution.

### III. FOUNDATION

#### Performance Tradeoffs

- **Service Level Agreement (SLA):**
  
  SLA is a quality forcing mechanism [8], which should be a part of the service design of any service provider not only the CSP. The SLA is assuring the availability, security, response time, many other performance attributes, and many other logical attributes i.e. scalability.

  SLA should be a written contract, which should be clear and available for both the provider and the consumer. No change in any of the contract points of the SLA, unless all of the contract actors agree on those changes.

  For a successful SLA in cloud computing, SLA is divided into two phases [6]. SLA-E is the first phase, in which the SLA contract is signed, then after the transit SLA-E phase, the SLA-M phase is to ensure and manage the SLA, by allocating resources and keep maintaining it along the contract period. Good SLA-M is the main result of SLA renewal. Usually SLA in multi-tier as the CSP can rent some resource to a broker, and then the broker rent those resources to end user; In that case, even the communication is not directly between the end user and the CSP directly, as no direct SLA contract between the two parties.

  - **Power Consumption:**
    
    The fact that the cloud is a set of hundreds, may be thousands of nodes operating in a datacenter [9] that consumes a lot of power in order to be operated. Moreover, the nodes need a very critical cooling requirements [10], in which another power consumption factor that can’t be neglected which is air conditions. The cloud’s power consumption is further translated into heat pollution and electric bill.

    Power consumption is a challenge that highly affect the CSP. Regarding expansion, the power consumption should be studied as more nodes means more power and heat.

  - **Migration Policies**

    The virtualization feature had added flexibility to each VM, so that VMs are independent from the hosting physical node. VM migration can be used to utilize the node resources more efficiently, so VM is migrated from over utilized host to less utilized host, in which SLA violation is eschewed. Successful VM migration needs three layered consequential function as shown in Figure (2)

    ![Figure (2): VM Migration layer model](image)

    - **VM Allocation Policies**

      The main function of the allocation policies is to find and mark an over utilized node. The output of each policy is an upper CPU threshold, so that if any node has CPU utilization more than the upper CPU threshold, then mark that host to be over utilized.

      Equation 1 [11] shows the returned value of the dynamic allocation policy function, which is used to calculate the upper CPU threshold.

      \[
      T_u = 1 - (S \times X)
      \] (1)
Where $T_u$ is the CPU upper threshold, $X$ is the output of the dynamic VM allocation policy, and $S$ is an adjustable parameter called Safety parameter. Even in the dynamic allocation policies, the cloud administrator need to adjust the safety parameter. Relatively high safety parameter is used to guarantee small SLA violation percentage, as it results in lowering the upper threshold, so a host with normal CPU utilization is marked as over utilized. That means that the workload of the cloud will be distributed over a large number of nodes, which relatively increase the power consumption, and vice versa.

- **Static Threshold**
  The cloud administrator manually adjust the threshold. That policy is the fastest allocation policy that can be applied, as no calculations is done before applying the policy. The static threshold policy is not suitable if the workload on the nodes is changing frequently.

- **Median Absolute Deviation (MAD)**
  MAD is a dynamic allocation policy based on the median of the CPU utilization of all nodes. That policy is characterized by its stability, as the nature of the median is to neglect the overshooting small or big values.

- **Interquartile Range (IQR)**
  Another dynamic allocation policy based on median, its idea is to arrange the CPU utilization of all hosts, then divide these values into two equal sets, after that calculate the median of each set separately, and finally to subtract the lower median from the upper median. Equation 2 shows the IQR algorithm, where $Q3$ is the median of the upper set, and $Q1$ is the median of the lower set.

\[
IQR = Q_3 - Q_1
\]

- **Robust Local Regression (LRR)**
  Enhanced Local Regression technique, based on linear approximation, where a dynamic numerical analysis is used to fit a curve using given CPU utilization values. LRR is more robust than LR, which is not vulnerable to outliers.

- **Other AI Policies**
  Neural networks can be used to divide the nodes into clusters, where the utilization is the cluster criteria, so nodes that are in the high utilized cluster will be marked as utilized.

- **VM Selection Policies**
  The second step in the VM Migration layer approach is the VM Selection policy, in which after an over utilized host is marked, one or more VMs need to be migrated to another host to avoid high SLA violation.

- **Maximum Correlation (MC)**
  As the CPU has the biggest power share among different components that consume power in each node, the MC VM with the CPU is most probably the most VM that results in consuming power. Migrating the MC VM will result to minimize the number of VM migration.

- **Minimum Migration Time (MTT)**
  The fact that the VM migration itself result in SLA violation, MTT is a good choice to SLA aware CSP, as the MMT is to find the fastest VM that can be migrated.

- **Random Selection (RS)**
  A VM to be migrated is selected randomly without any calculations. The RS policy is used when a fast decision is needed.

- **Other AI Policies**
  Recalling the main goal of the VM selection policies, which is to find and mark the most suitable VM to be migrated. The selection of the VM is based on it’s resource utilization i.e. CPU, RAM, HDD, Network activity, and others less important attributes, so AI techniques e.g. genetic algorithm, neural networks, self-organizing-maps, and K-means [12] can be used to select VM to be migrated based on an input criteria.

- **VM Placement Policies**
  VM placement is the Final layer in the VM migration layered approach, which aims to find the node that best fits to place the marked VM from the previous step.

- **Power Aware Best Fitting Decreasing (PABFD)**
  PABFD [13] is the modified power aware version of the original BFD. The BFD idea was the same of the pin packaging problem [14], whose idea was first to arrange the VMs from the VM selection phase in a descending order, then start to place them to the currently least utilized host.

### Adaptive Power Panel

- **Definition**
  Applying only one policy for each of VM migration layers will not be sufficient for all of the customers’ requirements [15]. In other words, some customers are SLA aware, so they need low SLA violation percentage regardless the price. Other customers are cost aware, and can accept relatively high SLA violation percentage. But applying only one policy will be suitable for just one setup for whole of the CSP datacenter.

  The adaptive power panel [4] is to divide the CSP datacenter into different clusters, then apply different policies on those clusters. That approach will allow the CSP to accommodate different requests from lessee, and will be able to satisfy the need of all of them. The Adaptive power panel is introduced with three different
modes, which are listed below. The CSP can change the cluster mode by manipulating the safety parameter.

- **Modes**
  - **Power Aware Mode**
    The power aware mode can be adjusted by choosing a small safety parameter, so the threshold of the allocation policy will be big, so only small amount of nodes can handle workload of the cluster, so the rest of the nodes in that cluster will be moved to sleep, which will save their power consumption. But in the other hand the SLA violation will be high as some VM can face resource starvation.
  - **Balanced Mode**
    This mode tries to balance between the power aware mode and the high performance mode, that mode is the default mode for the clusters.
  - **High Performance Mode**
    The high performance mode targets the least SLA violation percentage among other modes, regardless the power consumption. This mode can be activated by configuring relatively high safety parameter, which will results that the threshold of the VM allocation policies to be small, so nodes will be marked as utilized before its actual highly utilized. The fact that the high performance mode will consume the most power is that the workload will be distributed on a large number of nodes.

**IV. PROPOSED CLUSTER SIZING**

Although the adaptive power panel had introduced the clustering to the cloud computing, but the size of cluster should be variable depending on the current workload on the CSP datacenter. The dynamic sizing is to move a node from a cluster to another cluster. Our proposed cluster sizing consists of two phases. Phase I is the cluster migration test, and phase II is applying node migration policy.

- **Cluster Migration Test**
  Before applying node migration policy, a proposed test should be applied to each cluster. Equation 3 shows the input I of the cluster migration test (CMT), which is the actual average SLA violation percentage over the expected average SLA violation percentage. Figure 3 shows our proposed cluster migration test, which got one input I, the output will be to mark a cluster as over utilized, average utilized, or under utilized which will be discussed later in this paper.

$$I = \frac{\text{Average SLA violation}}{\text{Average SLA violation expected}}$$

The CMT is simply a function whose formula is shown in Equation 4. Where I is the input from Equation 3, and W is a positive weight parameter, which will be manually adjusted by the CSP administrator. The weight parameter is directly proportional with the CMT.

$$\text{CMT} = W \times I$$

**Figure (3): Node Migration Test input and outputs**

- **Cluster Condition**
  The output of the proposed CMT will specify the next action on that cluster. The effectiveness of the input I on the CMT output can be manipulated by the weight parameter. We propose that the CMT to be a fuzzy equation, to eliminate the cluster’s instability as shown in Figures 4 to 7.

  - **Under Utilized Cluster**
    In case of under utilized cluster, the cluster can contribute with one or more of it’s nodes to support other over utilized clusters. The cluster is under utilized when the CMT value is small.

**Figure (4): Fuzzy Equation with Weight and SLA-Ratio as inputs and CMT-output as the output**
Figure (5): Fuzzy Rules of Proposed CMT Fuzzy Equation

- **Average Utilized Cluster**
  The cluster is marked as average utilized if the CMT value is around the value of one, where the current average SLA violation % is almost equal to the expected average SLA violation %.

- **Over Utilized Cluster**
  The characteristics of the over utilized cluster is the cluster whose current SLA violation is higher than the expected SLA violation, so more nodes are needed to handle the current VMs workload. Those nodes are typically from the nodes of the under utilized clusters.

- **Node Migration Policy**
  Phase II is to apply node migration policy, which is to create a list of under utilized nodes from the under utilized clusters, thus to be migrated to the over utilized clusters. The proposed idea of the policies is finding a threshold, in which if the node’s utilization is less than that threshold, then that node is marked to be in the migration stand by list. The policy is only applied to the under utilized clusters.

- **Static Threshold**
  This policy is to manually set a static threshold for the resource utilization. This policy eliminates the computational step of calculating the threshold, and can be useful if the CSP want a fast action on a particular situation. All of the hosts that are under that threshold will migrate their VMs to another nodes, and will be added to the migration standby list in an ascending order.

![Figure (5): Fuzzy Rules of Proposed CMT Fuzzy Equation](image)

![Figure (6): CMT output sample after Fuzzification and Defuzzification](image)

- **Minimum VM counts**
  The node that handle least number of VMs, will be most likely the fastest node to migrate its VMs to another nodes, then join the migration standby list.

V. EVALUATION AND RESULTS

Cloudsim [16] [17] toolkit and Matlab are chosen to hold our testing. Cloudsim is a java based cloud simulation tool. We had used a real VMs’ workload from a project called CoDeeN [18] as an extension to the Cloudsim, which monitor the workload of a huge datacenter across the world called PlanetLab.

Our setup as shown in Table 1, contains two clusters, where one is facing resource starvation and the other has more nodes. Cluster 1 is in balanced mode contains 50 node and 100 VMs is over utilized, where cluster 2 in power aware mode contains 50 node and 70 VMs is under utilized, so we had applied our proposed solution. Table 2, and 3 shows the Nodes and VMs characteristics used in our test simulation. Our values are chosen to prove the cluster sizing concept.

![Figure (6): CMT output sample after Fuzzification and Defuzzification](image)

![Figure (7): Full Fuzzy Surface output with all reasonable values of SLA-Ratio and Weight](image)

<table>
<thead>
<tr>
<th>Table 1. Cluster Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster ID</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>
Table 2. Nodes Characteristics

<table>
<thead>
<tr>
<th>Host Types</th>
<th>PES</th>
<th>RAM</th>
<th>MIPS</th>
<th>BW</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP ProLiant ML110 G4</td>
<td>2</td>
<td>4096 MB</td>
<td>1860</td>
<td>1 Gbps</td>
<td>1 TB</td>
</tr>
<tr>
<td>HP ProLiant ML110 G5</td>
<td>2</td>
<td>4096 MB</td>
<td>2660</td>
<td>1 Gbps</td>
<td>1 TB</td>
</tr>
</tbody>
</table>

Table 3. VMs Characteristics

<table>
<thead>
<tr>
<th>VM Types</th>
<th>PES</th>
<th>RAM</th>
<th>MIPS</th>
<th>BW</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>870 MB</td>
<td>2500</td>
<td>100 Mbps</td>
<td>2.5 GB</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1740 MB</td>
<td>2000</td>
<td>100 Mbps</td>
<td>2.5 GB</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1740 MB</td>
<td>1000</td>
<td>100 Mbps</td>
<td>2.5 GB</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>613 MB</td>
<td>500</td>
<td>100 Mbps</td>
<td>2.5 GB</td>
</tr>
</tbody>
</table>

First Applying the CMT on both cluster 1, and 2. The inputs and output are shown in Table 4. CMT value is computed by a MATLAB project.

Table 4. Cluster’s Condition

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Expected SLA Violation %</th>
<th>Current SLA Violation %</th>
<th>Input (l)</th>
<th>Weight</th>
<th>CMT Value</th>
<th>Cluster Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.13%</td>
<td>0.15%</td>
<td>115.38%</td>
<td>3.2</td>
<td>1.39</td>
<td>Over Utilized</td>
</tr>
<tr>
<td>2</td>
<td>0.17%</td>
<td>0.07%</td>
<td>41.18%</td>
<td>0.75</td>
<td>0.31</td>
<td>Under Utilized</td>
</tr>
</tbody>
</table>

Second applying node migration policy on cluster 2 to find the node migration list, and finally moving the marked nodes to the under utilized cluster (Cluster 1). Table 5 shows our final results when setting the threshold to 0.4, which is 6 nodes.

Table 5. Node Migration

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Expected SLA Violation %</th>
<th>Current SLA Violation %</th>
<th>Input (l)</th>
<th>Weight</th>
<th>Nodes Moved</th>
<th>CMT Value</th>
<th>Cluster Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.13%</td>
<td>0.12%</td>
<td>92.31%</td>
<td>1.2</td>
<td>6</td>
<td>1.1</td>
<td>Average Utilized</td>
</tr>
<tr>
<td>2</td>
<td>0.17%</td>
<td>0.21%</td>
<td>123.53%</td>
<td>0.75</td>
<td>6</td>
<td>0.93</td>
<td>Average Utilized</td>
</tr>
</tbody>
</table>

VI. CONCLUSION AND FUTURE WORK

VM migration is the main principle of the adaption of the SLA violation percentage and the power consumption tradeoffs. The safety parameter is the datacenter’s controller, as it is directly proportional with the power consumption; Moreover, inversely proportional with the SLA violation percentage. Applying only one policy to the CSP’s datacenter can’t accommodate different lessee requirements.

The Adaptive power panel had solved this problem by dividing the CSP’s datacenter into clusters, then apply different policies independently on each cluster; Moreover, as an enhancement to the adaptive power panel, the proposed cluster sizing is introducing node migrating from its home cluster to another cluster, thus solves the cluster starvation for more nodes.

Our future work is to add the time dimension in the adaptive power panel, so the mode of a particular cluster is time dependent.

VII. REFERENCES

[3] Website: https://media.lucidn.com/mpr/mpr/shrinknkp_400_400/AAEAAQAADA AAAAABAAAAAAAADC00NMTk0OTMyWQx0NWMNGZtYi04YjB3LT VlMgynNTkzJ1Mw.png