EFFICIENT ALLOCATION OF VIRTUAL MACHINE IN CLOUD COMPUTING ENVIRONMENT

SONAM RATHORE
Faculty of Engineering and Technology, MITS, Lakshmangarh, India
e-mail:sonamrathore18@gmail.com

Abstract: Cloud computing is a latest new computing paradigm where applications, data and IT services are provided over the Internet. Computing service is provided on demand as a utility as other utilities. Cloud computing provides dynamic provisioning of computing services. Resource scheduling is a key process for clouds such as Infrastructure as a Service cloud. To make the most efficient use of the resources, an algorithm must be used which improves the utilization of system resources. In this work an algorithm is used which uses all the combination of allocation sequence and chooses the allocation sequence on the basis of strength of allocation. Experimental shows that the proposed algorithm gives the better utilization of resources.

Keywords: Cloud Computing; Virtual Machine; Allocation

1. INTRODUCTION

Cloud computing is a paradigm that is rising in this world of technologies. Cloud computing provides shared pool of resources on-demand over network on pay-per-use. Cloud computing ensures access to virtualized IT resources that are present at the data center and are shared by others. Cloud computing can be broadly classified into three services: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS). These services are provided over a network and accessible across computing technologies, operations and business models.

For the Infrastructure as a Service (one of the layers of cloud stack), one of the significant issues is the scheduling of virtual resources and virtual machines (VMs). It has been widely accepted that virtual machines can be employed as computing resources for high performance computing. Thus, efficient virtual machine allocation is essential in cloud computing environment for increasing resource utilization and efficient deployment of applications in virtual machine.

There are some popular open-source cloud systems, such as Eucalyptus, OpenNebula, and Nimbus, to decide the allocation of resources. To deal with the problem of allocating VM instance request to available computing nodes, Eucalyptus uses Greedy (first fit) or Round robin algorithm, with GREEDY, the first node which can meet the IRs will be chosen. The ROWNROBIN query all the nodes in circular order, until find the fitted node [1] [2]. The OpenNebula default scheduler provides a rank scheduling policy that places VMs on physical resources according to a ranking algorithm that is highly configurable by the administrator, and relies on real-time data from both the running VMs and available physical resources [3][7]. Nimbus uses some customizable tools like PBS and SGE. PBS is a queuing system and SGE uses Job Scheduling Hierarchically (JOSH) [4][5]. However, all of these algorithms fail to achieve higher VMs utilization rate. Therefore an algorithm must be used which is efficient in allocating VMs instance request and hence increases the resource utilization.

2. MOTIVATION AND PROBLEM DEFINITION

IaaS layer of cloud computing serves as a foundation for the other two layers (i.e PaaS and SaaS), for their execution therefore we have focused on the IaaS. IaaS deliver computer infrastructure typically a platform virtualization environment - as a service. Efficient scheduling of VMs instance request which meet user’s requirements and improve the resource utilization increases the overall performance of the cloud computing environment. VM instance scheduling in IaaS is the one of the crucial cloud computing questions to address.

Suppose, „M” physical machines are available and their resource capacities given along memory, CPU and hard disk dimensions. There are „N” virtual machines to be placed. The requirements of these virtual machines are given along the dimensions of memory, CPU and hard disk. We have to find allocations of VMs on available physical machine that satisfies the VMs” resource requirements and increases overall the resource utilization. The objective of this research work is to introduce an algorithm for efficient VM scheduling in cloud computing in terms of resource utilization rate. The proposed algorithm is compared with other existing algorithm for VM instance allocation.
3. DESIGN MODEL OF VM ALLOCATION IN CLOUD COMPUTING ENVIRONMENT

Figure 1 gives an architecture of VM allocation in cloud computing. Cloud users or the Cloud consumers request for the VMs and ask for VM requirement to the Cloud broker. The Cloud broker or data center broker acts on behalf of the broker, looks for the Cloud service providers who can make VM request fulfilled by querying the Cloud Information Service (CIS). Every new resource should be register to the CIS. Once the Cloud service provider has been chosen by Broker, it submits the VMs list to the Datacenter. Datacenter holds the physical computing servers i.e hosts and VMs place on host on the basis of allocation policies decided by the Cloud service provider.

Where, each $x_{ij}$ indicates whether VM instance $i$ is placed on the computing node $j$.

$x_{ij}$ will be 1 if VM $i$ is placed on computing node $j$, otherwise 0.

In (1) if $k=1$ it represents CPU Core, if $k=2$ it represents memory and $k=3$ represents hard disk capacity. Value of $P_k$ depends on value of $vm_k$/node$_k$ , so when value of $P_k$ is $a$ then placement of instance on node is right fit and we get a maximum usage of resources. When value of $P_k$ is $b$ then it means that the placement only get a suboptimal solution, it cannot get a maximum utilization of the resource. In both cases we assign a positive value to $P_k$, what’s more, $a$ bigger than $b$, so that if the placement is right fit, the solution will be encouraged. While the value of $vm_k$/node$_k$ is bigger than one which means the placement absolutely not fit, we assign a negative value to $P_k$.

3.1 Proposed Model

Let us consider a set of VM requests, a set of interconnected computing nodes connected by LANs. The computing nodes are different kinds of ordinary PCs, servers, and even high performance clusters. And cloud provides all kinds of machines it possesses in forms of virtual machine that clients can visit it through Internet as a service. In this work, we take the number of CPU cores, Memory capacity and Hard Drive capacity in consideration, which most of the existing IaaS cloud systems do.

Assumed there are $n$ VM instance requests (IRs) and $m$ idle computing nodes available in the cloud. Now the problem is to find the allocation sequence which makes the utilization rate of the resource achieve maximum. The problem can formulized as

$$\text{Strength of allocation} = \sum_{j=1}^{n} \sum_{k=1}^{3} P_k A_{ij}$$

where,

$$P_k = \begin{cases} a, & \text{if } vm_k / node_k = 1 \\ b, & \text{if } vm_k / node_k < 1 \\ c, & \text{if } vm_k / node_k > 1 \end{cases}$$

Where X>> 1000

In Strength of allocation any big number is added to $P_k$ to make it positive.

So, the allocation sequence with largest Strength of allocation is considered for VM allocation on available hosts in datacenter.

3.2 Flowcharts

Figure 2 and Figure 3 gives the flow of FCFS and proposed algorithms for VM allocation respectively.

- FCFS VM allocation policy works as follows:
  - Cloud user requests for VMs.
  - Cloud Broker which acts on behalf of the Cloud user submits the VMs request list.
  - An event has been generated for the submission of the VMs request list to the data center.
  - One by one VM in the series in the list has been passing to create VM on the host If not found host with enough of PEs required by VM then VM cannot be created.
  - If found first host with less number of PEs in use and fulfils VM PE requirement:
    - If host has enough RAM to fulfil the VM’s RAM requirement then place VM on host.
    - Else, VM is not created on the host. This VM is again sent to create a new list of VMs.

- VM allocation using proposed algorithm works as follows:
  - Cloud user requests for VMs.
  - Cloud Broker which acts on behalf of the Cloud user submits the VMs request list.
  - An event has been generated for the submission of the VMs request list to the data center.
The VM list is passed for the placement on the available host.
Permutations of placement order of VMs on available hosts are considered.
Strength of allocation is calculated for all the allocation sequence taken into consideration the VM’s parameters.
Allocation sequence with the highest strength of allocation value is considered for the VM allocation.

4. EXPERIMENTAL RESULTS

The proposed model is implemented in JAVA using Netbeans IDE. The implemented algorithm is then integrated with CloudSim package [8] for simulation. The algorithm is tested for different sets of VM instance request and computing nodes. Table I and Table II gives the example of parameters (CPU cores, Memory and Hard disk) for VM instance request and computing nodes.

Table I. VM instance request parameters

<table>
<thead>
<tr>
<th>VM id</th>
<th>CPU cores</th>
<th>Memory</th>
<th>Hard disk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>512</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>512</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1024</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2048</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>512</td>
<td>10</td>
</tr>
</tbody>
</table>

Table II. Computing Nodes Parameters

<table>
<thead>
<tr>
<th>Host id</th>
<th>CPU cores</th>
<th>Memory</th>
<th>Hard disk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>512</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1024</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1024</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2048</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1024</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>2048</td>
<td>40</td>
</tr>
</tbody>
</table>

The figure 4 depicts the VM allocation using proposed algorithm for the VM instance request parameter and Host parameter given in table I and II. The horizontal axis represents the hosts with the host ids. Number in the line graph represents the VM_id created on the host with respective host id.

4.1 Simulation result analysis

The figure 5 below depicts the comparison of resource utilization rate using proposed and FCFS VM allocation policy. The horizontal axis is the number of VMs requests and vertical axis represents the average resource utilization rate.
5. CONCLUSION

IaaS provides provisioning of processing, storage, networks, and other fundamental computing resources over a network. The VM allocation is a major issue in IaaS service of the cloud computing as the placement of these VMs can impact application performance because the IaaS providers are unaware of the hosted application’s requirements. Therefore the efficient VM allocation policy must be used. In this work the experimental results shows that proposed algorithm can improve resource utilization by efficient VM allocation.

6. ACKNOWLEDGMENT

We would like to thank Mr. P. K. Bishnoi of MITS, Lakshmangarh (Rajasthan) for his support and guidance.

REFERENCES