Aggregation of Resources Using HABA Algorithm for Efficient Service Provisioning in the Cloud Environment

P.Durgadevi¹ M.Sridhar² T.Chindrella Priyadharshini³ M.Swarnalatha⁴

¹²³Assistant Professor ²Sr.Software Engineer

R.M.K College Of Engineering And Technology, Puduvoor ³Gavis InfoTech, Chennai

Abstract— Cloud computing is that the outstanding technology that imparts computing resources as a service. In cloud, procedure resources area unit set in terms of Virtual Machines (VMs) and/or Virtual Clusters (VCs) that is outlined as a conglomeration of VM instances providing same service, front-ended by a network load balancer . During this surroundings, varicoloured cloud users will demand variety of cloud services contemporaneously. Service provisioning ought to be performed in an exceedingly deft and refined manner, in order that all the resources area unit created possible for a user to request in Associate in Nursing decisive manner to assuage their wants. In this paper , Hierarchical agglomerate bunch algorithm(HABA) for bunch of resources supported its attributes patterns that created simple of allocating resources to requesting users supported their demands, in and of itself, this bunch algorithmic rule include 2 steps. In opening, build a tree-based ranked taxonomy mistreatment info entropy of resource attributes. Second step, clusters area unit obtained by cutting the dendogram at a desired level. every connected part forms a cluster . The planned algorithmic rule is evaluated with a number of the fresh bestowed strategies and comparison with HABA algorithmic rule shows that the planned work furnish a typical solution for the matter of sharing the resources in cloud surroundings.

Key words: Clusters, Virtualization, Agglomerative, Dendogram, Heuristics

I. INTRODUCTION

Cloud computing are rapidly growing technology in popularity. Virtualization transforms the physical resources into virtual resources, the cloud operator manages the resource by allowing the user to access resource as pay per use basis, with ideal intensity of computing resources without nerve-racking of location and attributes of resources. The attractiveness of cloud leads to the increase of network speed, and utilization of virtualization technologies. The enterprises will hasten their migration from building and owning their own systems to renting cloud computing services because cloud computing services are simple to utilize and can reduce both costs and workloads[11].

With the status of cloud computing, the task scheduling becomes more and more important. How should we schedule a task to cloud resources effectively and efficiently based on user demands becomes a problem. Cloud computing discussion in general begins with virtualization. Virtualization is significant tool to cloud computing , it makes the service provisioning as simple provided that a platform for utilizing multifarious IT resources in a increase manner, which leads to cloud computing concept are cost effective[12].

Fig. 1: Virtualization concept in cloud environment

Virtualization has three important characteristics with relate to Cloud:

- Partitioning of single physical system into VMs
- Isolation of VMs with each other
- Encapsulating details of VMs from other VMs

Even though adapting VM concepts, still there is a problem of degradation in performance of cloud resource provisioning. A novel technique based on behavior and importance of VMs executing simultaneously is needed , to manage virtual resources such as CPU, Memory, Disks, Data etc and afford appropriate sharing of resources to cloud users. This problem is extensively deliberate in [2] and [4].

In this paper, we introduce Hierarchical agglomerate bunch algorithm (HABA) resource provisioning algorithm The main objective of this work is to organize the huge heterogeneous the information entropy based on attributes of resources from different cloud center into clusters according to resource. This will yield fast retrieval of resources and effective provisioning to the user with their requirements from huge cloud.

Our paper is organized as follows, section II discuss the related work, section III describes the proposed Algorithm and experimental results and comparison with other mining algorithms are discussed in section IV. Section V gives the conclusion with future work.

II. RELATED WORK

Mohammad Ghasemi Gol et al [6] have projected a stratified methodology and additionally they extended the stratified agglomeration algorithmic rule to cluster fuzzy knowledge for the primary time. Finally that approach has been compared with a number of the fresh bestowed strategies within the literature, the main advantage of the algorithmic rule was its fault tolerance against creaky samples.

Hui bureau t al [7] have planned associate new clustered graded clump rule was conferred and enforced by Map cut back framework. The approach splits the initial documents’ vectors’ set into partitions with the strategy of initial classification, hence it circulates the partitions to
Aggregation of Resources Using HABA Algorithm for Efficient Service Provisioning in the Cloud Environment

Consider, we have \( N \) number of VMs is represented by \( \text{VMs} = \{ \text{VM1}, \text{VM2}, \text{VM3}, \ldots \ldots, \text{VMn} \} \) and each VM is treated as resource is associated with attribute patterns.

\[ \text{i.e., } i^{th} \text{ resource can be expressed as } \]
\[ \text{VM}_i = \{ \text{VM}_{\text{COMP}}, \text{VM}_{\text{BW}}, \text{VM}_{\text{SC}} \} \]

These attributes are considering for aggregation of resources found in cloud environment for efficient allocation of resources to the users. In this paper, we propose an algorithm named as Hierarchial agglomerate bunch algorithm (HABA) for clustering of the resources in form of Hierarchical tree like structure called dendogram. Dendogram is defined as visualization of Hierarchical clustering where the distance of split or merge takes place.

**HABA Algorithm**

**Input:** Virtual machines set

**Output:** Dendogram and Clusters

```
For \( i = 1 \) to \( n \)
\[ \text{C}_i = \{ \text{VM}_i \} \]
End For
\[ \text{C} = \{ \text{C}_1, \text{C}_2, \ldots, \text{C}_n \} \]
L = \( n \)

While \( \text{C.current} > 1 \) do
    \( (\text{C}_{\text{min}}, \text{C}_{\text{max}}) = \min \text{dist}(\text{C}_i, \text{C}_j) \) for all \( \text{C}_i, \text{C}_j \)
    Remove \( \text{C}_{\text{min}}, \text{C}_{\text{max}} \) from \( \text{C} \)
    Add \( \{\text{C}_{\text{min}}, \text{C}_{\text{max}}\} \) to \( \text{C} \) and \( L = \text{L}-1 \)
End While
```

For each pattern of resource, the above algorithm is applied which forms a tree like structure called dendogram. Each object is treated as separate cluster. Then, these clusters are merged successively, which forms a new clusters are added to \( C \). Whenever a new cluster is formed, the base clusters are removed from \( C \). These procedure is repeated until there will be one cluster in set of VMs. Sample Dendogram is shown in figure 2 for set of 10 VMs.

**Fig. 2: Dendogram shows clustering of VMS**

**III. PROPOSED WORK**

In cloud, each VM, is characterized by two parameters 1) \( P(i) \) depicts its priority among VMs in cloud 2) \( F(i) \) depicts its predicted future demands. Apart from these parameters, each VM, has set of attributes such as memory, CPU, Storage, Capacity of VMs, Number of cores etc. In proposed work, VMs are analyzed based on three important parameters.

- Storage Capacity
- Bandwidth
- Computing power

**IV. EXPERIMENTAL RESULTS AND DISCUSSIONS**

The proposed algorithm is compared with Agglomerative Clustering Algorithm \([1]\) by measuring the performance based on throughput, cost and service time. Random datasets, data’s, processes have chosen for generate a more practical workload. In order to generate a more realistic workload, we randomly choose the datasets, data sizes, and number of processes. The result shows the improvements in the performance of virtualized environment are achieved in proposed algorithm.
A. Experimental Setup
Hierarchical agglomerate bunch algorithm (HABA) is executed on variety of file types such as .rtf, .bmp and .doc. Cloudsim is the simulated tool for cloud computing used here for simulating virtualized environment. The software programming language java is used in simulator.

B. Performance Evaluation
Hierarchical agglomerate bunch algorithm implemented for aggregating the resources into clusters and scheduling the tasks to appropriate VMs in clusters. When the task arrives, the scheduler proceeds the tasks to clusters. Based on requirements and cloudlet characteristics, task is scheduled with appropriate resource found in the cluster. Hence, Scheduling of task to suitable resource in cluster with least CPU utilization time is evaluated on HABA and ACA and the results are shown in below Table 1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Utilization</th>
<th>Cluster ID</th>
<th>Task Name</th>
<th>Utilization</th>
<th>Cluster ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>rtf</td>
<td>102.54</td>
<td>4</td>
<td>rtf</td>
<td>500</td>
<td>2</td>
</tr>
<tr>
<td>bmp</td>
<td>51.45</td>
<td>1</td>
<td>bmp</td>
<td>42.56</td>
<td>5</td>
</tr>
<tr>
<td>doc</td>
<td>54.87</td>
<td>6</td>
<td>doc</td>
<td>260.00</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: CPU Utilization Comparison

The throughput is defined as the number of tasks finished in a certain period of time. The normalized throughput is measured for scheduling methods. OPHAC algorithm achieves better throughput compared with HAC algorithm is shown in below Table 2.

<table>
<thead>
<tr>
<th>Employed Algorithm</th>
<th>Achieved Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>HABA</td>
<td>0.86</td>
</tr>
<tr>
<td>ACA</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Table 2: Throughput of HABA and ACA

The cost is estimated for the cloud based on time utilization in the cloud. By using minimum servicing time, cloud cost can be controlled. Based on the usage of resource, cost was calculated using INR 1 per second. Hence the proposed algorithm takes lesser time and lower cost while compared with ACA and the overall cost depends on the cloud providers.

<table>
<thead>
<tr>
<th>Employed Algorithm</th>
<th>Time (Sec)</th>
<th>Cost (Rs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HABA</td>
<td>13</td>
<td>13.00</td>
</tr>
<tr>
<td>ACA</td>
<td>38</td>
<td>38.00</td>
</tr>
</tbody>
</table>

Table 3: Cost and Time Comparison

The Throughput, service cost and execution time taken by the task are evaluated and the performance improvement is shown in Fig.3 & 4.

![Fig. 3: Comparison of throughput](image)

![Fig. 4: Time and Cost Comparisons](image)

Fig. 4: Time and Cost Comparisons

V. CONCLUSION
Aggregation of the resources is done based on attribute patterns of cloud resources such as Speed, Bandwidth and Storage Capacity. The proposed algorithm was implemented and statistically analyzed and also compared with ACA. Simulation results show that HABA algorithm aggregates the resources one by one, which means that the scalability problem becomes smaller, so the result of clustering is more effective and efficient. Our future work explore on scheduling algorithms to further improve the performance in virtualized environment.

REFERENCES


