Improved Virtual Machine Allocation Strategy using Particle Swarm Optimization Algorithm

B. Thirumala Rao¹*, K. Nandavardhini², K. Navya³, G. Krishna Venkata Sunil⁴

Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India – 522502
*Email: drbtrao@kluniversity.in

Abstract

Virtual machine position (VMP) is a critical issue in choosing most appropriate arrangement of physical machines (PMs) for an arrangement of virtual machines (VMs) in distributed computing condition. These days information concentrated applications for handling huge information are being facilitated in the cloud. Since the cloud condition gives virtualized assets to calculation, and information concentrated applications require correspondence between the registering hubs, the situation of Virtual Machines (VMs) and area of information influence the general calculation time. The essential target is to decrease cross system activity and transmission capacity use, by setting required number of VMs and information in Physical Machines (PMs) which are physically nearer. This paper exhibits and assesses by a meta-heuristic calculation in view of Parallel Computing and Optimization (PCO) which select an arrangement of adjoining PMs for setting information and VMs. In the wake of choosing the PMs, the information are duplicated to the capacity gadgets of the PMs and the required number of VMs are began on the PMs based on their VM allotment limits. Recreation comes about demonstrate that this determination diminishes the whole of separations amongst VMs and henceforth lessens the activity fruition time.

Keywords: ACO; Physical Machine; PSO; Virtual Machine;

1. Introduction

The services provided by the cloud computing are many ways. This is done by the service providers those are providing these services. There is no need of software and hardware requirements to deploy the application in cloud. Applications like accounting packages, industrial packages and may other applications can deploy in cloud irrespective of the hardware and software requirement. Quality of service can be provided by the many of the many service providers. The services are based on the servers and virtual machines, physical machines. In cloud computing the compatibility plays the major role in providing the resource to the users on demand. Cost is also one of the criteria to provide the services on demand.

Many companies offering high cost and some companies are offering the low cost. In this paper, the utilization and provisioning of the Virtual machines and physical machines can be done. It is still a big question that how the resources in cloud computing setup may be managed in cost effective manner.

In cloud computing when the end user accesses some services in the cloud there is no need of knowing about the service and platform used for the service.

Cloud computing platforms, such as those provided by Google, IBM Microsoft, Amazon, etc., the developers deploy applications by computers hosted by a central server. There are no of advantages using cloud computing for low cost. Providing better resources like resource provisioning and service accessibility no need of spending high expenditure to the cloud computing for better infrastructure like physical infrastructure for low better infrastructure from cloud computing. The bending nature of cloud computing provides fast services to the cloud users.

Some of the features provided by the cloud are:

- Hardware, operating system and platform is totally hidden from the client.
- All the infrastructure and platform licensing and maintaining is taken by the vendor.
- The client will avail the service using subscription and work with the software over web.
The organization of the paper is as follows. Section II discusses about Literature review. Section III discusses various Resource Provisioning types. Section IV gives conclusion & future research directions.

2. Related work

This section provides the information about the resource provisioning. The development of efficient service provisioning policies is among the major issues in cloud research. The issue here is to provide better quality of service in IaaS by provisioning the resources to the users or applications via load balancing mechanism, high scalability, high availability, and soft real-time context, even though various methods in Teng and Magoules [1] permits North American nation to select up a prime to bottom fact-finding comprehension of the service provisioning issue. Earlier scientific theory has been effectively connected to differing problems, for instance, net evaluating, and preventative management, routing, and networking. Resource provisioning can encompass 3 measurements consistent with Sotomayor [2] hardware resources, the package accessible on those resources, and therefore the time throughout that those resources should be ensured to be accessible.

A complete resource provisioning model should allow resource purchasing to choose times across these 3 dimensions, and also sets resource provider to proficiently fulfill those conditions. The service provisioning methodology as indicated by Hill and Varaiya [3] is taking into consideration a solution of the problem of allotting knowledge transmission and supports to satisfy a couple of variances of service appeals, separated by limits on the conventional rate and burstiness of the message and on the end-to-end delay. Here, the shoppers opt for the resources they have and also the system arranges their selections by means of that of plus evaluating to upgrade a general live of system execution perform multi-level component resource provisioning. Specifically, info level scaling isn’t thought of. Also, the authors don’t follow any approach toward dynamic resource management on clouds. Unused of resource provisioning prompts either over provisioning or below provisioning issue. Vijayakumar [4] propose a quick cloud resource provisioning (RCRP) calculation to reduce the combination cloud provisioning expense. Differing kinds of instability square measure thought-about within the formula. Dailey [5] propose a way for characteristic and retrieving over provisioned resources in multltier cloud-facilitated net applications. They exhibit the attainable usefulness of methodology in a very check assessment with a test bed EUCALYPTUS primarily based cloud and a factory-made employment. Anyhow, the problem is that they only address scaling of the online server level and a read-just information level. Specifically, they do not address programming configuration management.

Buuya [6] illustrates the various difficulties in tend to the problem of empowering SLA-situated plus distribution in server farms to meet competitive applications interest for registering administrations. Specifically, the shopper applications are becoming to be additionally an unpredictable and wish varied administrations to execute instead of one service. Specifically, work on cloud management in Cunningham and Holmes [7] and Armbrust and Fox [8] has focused on the provisioning and scaling of services within infrastructure clouds. Among these problems, the square measures confronted by the purchasers, execution and virtualization of problems. Square measure the foremost steady and customary problems owing a part attributable to the method that purchasers have zero ability to visualize into the cloud and square measure consequently compelled to counsel the cloud directors for facilitate. The reviews demonstrate that supply additional powerful bolster, cloud ought to produce devices to mechanize administrator’s enterprise. Vijayakumar [4] have thought-about the problem of costliest cloud provisioning for versatile data spilling applications is virtualized or cloud environmental. This technique increasingly accomplishes the shopper specified truth level by adjusting a flexible parameter at runtime. Chaisiri [9] have planned a perfect cloud resource provisioning (OCRIP) calculation to procure assets offered by various cloud suppliers. The best arrangement from OCRP is found by shaping and try random range programming. The OCRP calculation may be utilized as AN plus provisioning instrument for the developing distributed computing market within which the equipment will be success spare the combination price.

3. ACO (Anti Colony Optimization)

ACO represents an Ant Colony Optimization which is used to find the food for the ants based on the shortest path. For every ant there consists of chemical to send messages between them and this is called as a pheromone. Randomly ants search their food. Once they found the food they leave pheromone on the path to communicate with the other ants. In this paper, the existing system ACO is utilized to get the optimized solution to find the best virtual machine to the user. The disadvantages of the ACO Algorithm is as follows.

1. It takes more time to find the better solution.
2. The mobility support is not good within the VMs.

3.1. Algorithm

ACO algorithm Input: List of Cloudlet (Tasks) and List of VMs
Output: The best solution for tasks allocation on VMs Steps:

ACO Algorithm:
1. Initialize the no of tasks.
2. Initialize the list of VMs.
3. Start loop t=1.
4. Let m as VMs and start m randomly.
5. Update pheromones for all the VMs whose search for the paths.
6. The iteration continues until the cost of the task scheduling ends.
7. Based on the time t, the cost analysis is done for every VMs.
8:End

3.2. Results of ACO

Figure 3.2.1 Performance of ACO Algorithm with No of Servers and maintaining the bandwidth at all the servers.

4. Proposed System

The better dynamic adaptive global search-based optimization technique is called as Particle Swarm Optimization (PSO). In PSO, every server is considered as particle. Based on the behaviour of the server which is calculated by the loading and cost of the server which is based on the usage of the users. Here the distance is also a important parameter to get the correct output and to load the servers on time. All the servers located in various places,
work as a group based on their velocity. The best performance is calculated from all the other servers.

4.1. Algorithm
Scheduling heuristic.

1: Average cost of all tasks.
2: Average cost based on size between resources.
3: The wtn= acc (average computation cost)
4: The applications are started by the server.
5: All the servers runs according to the tasks loading.
6: If the server is in sleep or over headed the task assign for another available server.
7: Dynamic update of all the tasks.
8: Compute PSO based on the bandwidth, cost analysis and status of the server.

5. Experimental setup

In this paper, the experimental software is Java as the programming language and NETBEANS 8.0.2 & JDK 1.8 with HTML for static web pages. The hardware for the system is 4 GB ram and 500 GB hard disk for the storage of software and other project setup. The following steps are to execute the project.
1. Open project in NetBeans 8.0.2.
2. Start 8 or 10 VM’s
3. Start building website for the VM’s and select the number of user to access the website in the VM’s
4. The parameters such as time and cost of the VM’s processing is started to calculate the best feasible solution.
5. The existing system over provisioning is implemented without any features to stop the over provisioning.
6. The figure-3 and figure-4 shows the performance of the ACO and PSO.

6. Experimental results

Figure 6.1: Performance of PSO Algorithm with No of Servers maintaining bandwidth along with reducing the cost at all the servers.

7. Performance

The Figure-7.1 & Figure-7.2 explains the performance of the over provisioning, ACO and PSO. The graph represents the two parameters Time and Mobility (MCC) mobile cloud computing. For all the three algorithms the graph represents the time for recommending a best optimized virtual machine for the user in terms seconds. For the over provisioning and ACO the time taken for recommending a best virtual machine for the user compared with PSO. And the MCC represents the mobility between the users. Similarly, the graph represents the PSO-MCC as high compared with OP & ACO.

Among all the 3 algorithms PSO shows the better performance and it supports the mobility between the users.

8. Conclusion

This paper mainly focuses on scheduling of tasks by utilizing the Particle Swarm Optimization (PSO). The PSO maintains the load balancing between the servers and reduce the total cost of the users. The servers will communicate with the applications of the users. The results are clear that the PSO acquires the “Best Resource Selection” (BRS) for the better server map selection based on the load and distance of the servers.

References


Asia regional computer confederation conference, vol. 6, no. 4; 2011. p. 67–76.
