Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 115 (2017) 322–329

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Advances in Computing & Communications, ICACC-2017, 22-24 August 2017, Cochin, India

Dynamic load balancing algorithm for balancing the workload among virtual machine in cloud computing

Mohit Kumar[a,*], S.C.Sharma[b]

[a]*Research Scholar, IIT Roorkee, India*

[b]*Professor, IIT Roorkee, India*

Abstract

Performance of the cloud infrastructure is highly depends upon the task scheduling and load balancing. Therefore number of load balancing algorithms and technique are proposed by researchers throughout the world whose aim is to distribute the workload fairly among all the virtual machine while attaining the objective. We developed a load balancing algorithm that minimize the makespan time and enhance the utilization ratio of cloud resources. Computational results (Fig. 3 to Fig. 5) shows that develop algorithm decrease the makespan time and enhance the utilization of resource compare to min-min algorithm, fcfs and shortest job first in all condition.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

Keywords Virtual machine; task scheduling; workload; makespan time; load balancing ;

1. Introduction

Cloud computing is emerging technology and next big step in the evolution of virtual computing in the field of IT within last few years. It provides the scalability, flexibility, on demand services and virtualization type of services over the internet to the user. As the number of user is increasing day-by-day in cloud environment, load balancing has been become a challenging problem for cloud service provider. To overcome this type of problem, many

* Corresponding author. Tel. +91-9759950380
*E-mail address:*mohit05cs33@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications
10.1016/j.procs.2017.09.141

algorithms have been proposed by researcher, but all the algorithms have their advantages and limitation. The main aim of load balancing is to utilize the cloud resource (combination of hardware and software) in such way so that throughput and resource utilization type of parameter should be maximum[1]. Load balancing helps to execute the task in minimum time and increasing the performance of the system i.e., user can get back response in shortest period of time [2]. Other challenges also exist in cloud environment like security [3], data loss, heterogeneity and high communication delay.

Load balancing is achieved in cloud environment in two steps: first one is to distribute the task among the node, second one is to monitor the virtual machine and perform the load balancing operation using task migration or virtual machine migration approach. The aim of task scheduling is to create a schedule and assigned each task to node (virtual machine) for specific time period so that all task are executed in minimum time span. Task scheduling is NP complete problem in the field of computer science because number of task and length of task change very rapidly in cloud environment. It is difficult to calculate all possible task-resource mapping in cloud environment and find an optimal mapping is not easy task. Therefore we need an efficient task scheduling algorithm that can distribute the task in effective way so that less number of virtual machine should be in overloaded or under loaded condition. After allocating the task to virtual machine, cloud task scheduler start to perform load balancing operation so that task can be transfer from overloaded virtual machine to under loaded virtual machine and all virtual machine should remain in balance condition. Three basic components are required for task scheduling in cloud environment as shown in Fig. 1.

User level phase First phase is called cloud user phase where users submit their jobs J_1, J_2.....Jn through the graphical user interface or web interface with service requirement in terms of quality of service (QoS), hardware, software etc.

Cloud task scheduler phase All the task scheduling and load balancing operation are performed in this phase. Job request handler forward the authentic request to task scheduler for further processing where matchmaker match all the tasks to corresponding virtual machine and scheduler assigned task/job to virtual machine. Task scheduler contains the information about all the virtual machine (idle or busy).

Cloud level phase Last phase of basic architecture of task scheduling in cloud environment is called cloud level phase. A datacenter contain the many host and each host contain the heterogeneous virtual machine as shown in Fig.1, number of virtual machine can increase and decrease at run time, its depend upon the capacity of host and number of upcoming user request. Cloud monitoring and discovering service (CMDS) is used to monitor the virtual machine (idle or busy state) and discover the resource information. The rest of the paper is organized as follows: Section 2 describes the related work in which we will discuss existing load balancing and task scheduling technique in cloud environment that is related to my research work, Section 3 describe the problem formulation, Section 4 describe proposed and develop load balancing algorithm, Section 5 describe the simulation tool and experimental results and last Section 6 conclusion and future work.

2 Related Work

Several static [4,5,6,7,8,9,10,12] and dynamic algorithm [11,13,14,15,16] has been proposed in last decade. Static algorithm needs advanced information about the number of task and information about available resource. There is no need to continuously monitor the resource when static algorithm is working**.** These types of algorithm gives better results only when there is low variation in upcoming workload. User-Priority Guided load balancing improve Min-Min Scheduling Algorithm was proposed by Huankai Chen et al.[5] in which n different length tasks(T_1, T_2, T_3, T_4……..T_n) want to assigned to m heterogeneous resource (R_1, R_2, R_3, R_4……R_m), choose the minimum size task T and assigned to that resource which can execute the task minimum time in this algorithm. Suresh and vijayakarthick. P proposed an improved backfill algorithm (IBA) using balanced spiral (BS) method to improve the processing time of task [6], IBA algorithm provides the guarantee of quality of service in cloud environment but this algorithm do not provide better processing time when task comes randomly to cloud, to remove the limitation and achieve better quality of service with high resource utilization an algorithm IBA with EASY has been proposed for scheduling the task in cloud environment [7]. When user sends the request for the service, it also adds some quality of services parameter like deadline, priority, cost etc. M.Kumar and S.C.Sharma proposed a new algorithm that find