ISSN: 2321-7782 (Online)
Volume 1, Issue 7, December 2013
International Journal of Advance Research in Computer Science and Management Studies
Research Paper
Available online at: www.ijarcsms.com

Genetic Algorithm Based Resource Scheduling Technique in Cloud Computing

Hitesh A. Ravani[1]
M.Tech in CE
School of Engineering
RK University
Rajkot – India

Hitesh A. Bheda[2]
Assistant Professor
School of Engineering
RK University
Rajkot – India

Vrunda J. Patel[3]
M.Tech in CE
School of Engineering
RK University
Rajkot – India

Abstract: ***Resource Scheduling is the process of mapping tasks to available resources on the basis of tasks' characteristics and requirements. The received tasks are grouped on the basis of data and requested resources by the task and prioritized. Resource selection is done on the basis of its cost and turnaround time both using greedy approach. Task selection on the basis of a priority formula. This way of resource selection and task selection gives better results over sequential scheduling. The available resources should be utilized efficiently without affecting the service parameters of cloud. Resource scheduling focus on the complete utilization of the available resources. Main aim of this paper is analyze the various scheduling algorithm and manage the resources which are precisely available at certain fixed times and for fixed intervals of time. Find the optimizes scheduling algorithm for resource so the cloud provider get benefits in term of efficient resource management which provide more resources to allocate without postponing or declining any user requests. Cloud users also get benefits in term of their monetary gains at each front.***

Keywords: Cloud Computing, Resource Scheduling, Genetic Algorithm, Virtual Machine, Load Complementation.

I. INTRODUCTION

According to the National Institute of Standards and Technology (NIST) cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing is Internet-based computing service which provide resources, software and information on-demand like electricity. Cloud Computing provide capabilities to fly without investing in new infrastructure, provide training to new person and purchase licensing new software. Simply we can say that cloud is a connected pool of computing resources which available to the user in utility-style infrastructure. Management of these resources in dynamic manner provides great scope to the researcher. Resources are available for fixed interval of time as well as at certain fixed times so scheduling of these resources became major part of resources management. To maintain separation between users of the resources an optimized scheduling of resources is required.

Some problem like fragmentation of resources, low utilization of the resources are easily cause because of the expansion of system scale and virtual machines' migrations, etc because of this high energy consumption within an Internet Datacenter. Pre-migration strategy is based on three load dimension: CPU utilization, network throughput, disk I/O rate. TO obtain an

approximately best solution we apply the hybrid genetic algorithm [7] merge with knapsack problem with multiple fitness and experiments are conducted which prove the efficiency of the algorithm. The result shows that the algorithm can effectively achieve the goal of raising resources utilization and lower energy consumption under cloud computing environment. Cloud computing provide ability to existing internet technologies and web cluster to fit the required business needs by accessing distributed computing resources. To control the next generation data centers' and empowered application service providers for deploying applications there is a need of tools and mechanisms for analyzing the performance of cloud system. This paper gives two scheduling policies with an analytical resource prediction model for each policy for private cloud system. To provide exact performance measurement of such a system queuing model used. This technique help the cloud operators to tune the resources accordingly match with the requirements.

The algorithm which discuss in this paper will save energy consumption as well as increase the utilization of the resources. The execution of this algorithm with the virtual machine's actual migration will save 30-40% of the total physical machine's occupation as well as smooth the utilization of the loads. By dividing overall load into specific load we get more precise result about the virtual machine's placement. Cloud computing provide the service to the user as "pay-per use "base. Cloud provides software as a service, platform as a service, infrastructure as a service, and application as a service. There are various cloud deployment model depend on the usage of the cloud services like public cloud, private cloud, hybrid cloud, community cloud. In public cloud resources which provided by the provider are available to the general public over the internet, charges of resources are based on a fine-grained utility computing .Example are Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google AppEngine and Windows Azure Services Platform. In private cloud service is provided to the limited number of people on internal networks. The organizations which need accurate control over their data will mostly prefer to use private cloud. Example HP Cloud Start and eBay. In Hybrid cloud concept of public and private cloud as well as local infrastructures are combine. Hybrid provides proper placement of workloads depending on cost and operational and compliance factor. Example HP, IBM, Oracle and VMware provide this kind of service. In some extent Community cloud overlaps with Grid. Several organizations in a private community share cloud infrastructure. Community cloud can build up cross-boundary structure.

Resource Scheduling became complicated task in cloud computing as required resources are limited and the number of users increase day by day so it is most important to manage these resources in such a way that resources are properly utilize and the waiting time for resources is decreases. For proper scheduling of resources many algorithm are available as well as methods in cloud computing. Optimum resource scheduling help both user as well as service provider cloud provider get benefits in term of efficient resource management which in turn provides more resources to allocate user without postponing or declining user request for resources, cloud user get benefit in term of money proper utilization of money is done. The rest of the paper is organized as follows. We review some related work in section 2. In section 3, we describe the algorithm in detail. Section 4 describes the conclusion of the paper. Finally, in section 5 references for this paper is given.

II. RELATED WORK

Resource management and scheduling in a cloud computing become a hot topic recently in both the academic and industrial field. The problem of optimally scheduling the resources in a cloud is essential and important task for a cloud to be deployed large-scale. A good resource scheduling policy will raise the utilization of resources and save energy cost significantly. Various methods used in scheduling of resources but all are vary from each other and each have distinct characteristics. Resource Scheduling became complicated task in cloud computing as required resources are limited and the number of users increase day by day so it is most important to manage these resources in such a way that resources are properly utilize and the waiting time for resources is decreases. For proper scheduling of resources many algorithm are available as well as methods in cloud computing. Optimum resource scheduling help both user as well as service provider cloud provider get benefits in term of efficient resource management which in turn provides more resources to allocate user without postponing or

declining user request for resources, cloud user get benefit in term of money like proper utilization of money is done. In [1], author gives idea about various resource scheduling algorithm and comparison between these algorithms. It is highly essential to effectively manage the resources and effectively scheduling the job for these resources. For this purpose we need to find the most effective algorithm which schedules the resources in such a way that optimize use of the resource is done. Several algorithms are available for the scheduling of the resources but each one has its own limitation. These algorithms give optimal or least optimal solution of the problem. Currently there is a need of more accurate algorithm for resource scheduling and resource allocation which is biggest research challenge. In case of ANT Colony Optimization when more VM's are involved algorithm create colonies so an ant follow less likely pheromone trail from another colony. In case of Particle Swarm Optimization the solution space or search space is very large.

Cloud computing resource scheduling became complicated task because there are many alternative computers available with varying capacity. So in this paper we propose a model for job-oriented resource scheduling in cloud computing environment. In resource allocation task resources can be allocated to the process according to the rank of the job. This paper discus the analysis of resource scheduling algorithm. The time parameters of Round Robin, Pre-emptive Priority and Shortest Remaining Time First consider from this we can computed that SRTF has the lowest time parameters in all respects and most efficient algorithm for scheduling of resources. SRTF is the lowest time parameters and it is the most efficient algorithm for scheduling of the resources [2].For serving demanding jobs in enterprises and business market cloud computing is used. No of users, services and types of services increased because of this the performance of resources scheduling is important in cloud computing. Many factors affecting the resource scheduling like CPU Speed, memory, bandwidth etc that's why it is multi criteria decision making problem. In this paper we discuss an efficient Qos based resource scheduling algorithm using potentially all pair-wise rankings of all possible alternatives (PAPRIKA).First tasks are arranged based on the Qos parameters and then resources are allocated to the appropriate task using PARIKA method and user satisfaction. This algorithm reduces task completion time and improve resource utility rate. Algorithm simulated using Cloud Sim tool package [10]. Factor which effecting resource scheduling are CPU Speed(C) , Bandwidth (B), Stability (S), Length (L).PAPRIKA optimize the allocation of task to appropriate resources in minimum time as well as improved resource utility rate [3].

Cloud computing provide ability to existing internet technologies and web cluster to fit the required business needs by accessing distributed computing resources. To control the next generation data centers and empowered application service providers for deploying applications there is a need of tools and mechanisms for analyzing the performance of cloud system. This paper gives two scheduling policies with an analytical resource prediction model for each policy for private cloud system. To provide exact performance measurement of such a system queuing model used. This technique help the cloud operators to tune the resources accordingly match with the requirements. To obtain the steady-state system length distribution we developed a recursive method using the birth-death process. To provide performance metrics of private cloud computing system Finite source finite buffer multiple server queuing models is used. Various performance factor are consider like utilization of the system, utilization of server, the expected number of client requests in the system, the expected number of client requests in the queue, mean number of idle servers, mean waiting time, response time and blocking probability. The proposed model improves the system performance and optimizes the organizational resources in cloud computing system [4]. Cloud computing provide a necessity for utility computing service. Each provider has different service portfolio with a range of resources. Resource provisioning in cloud computing environment is important issue for resource allocation model. Every resource consider computational as well as network resources to accurately represent practical needs. In this paper we discuss new model to tackle the resource allocation problem for a group of cloud user requests. The objective of this model is minimizing the average tardiness of connection requests. This paper introduces four combined scheduling algorithms and used to schedule virtual machine on data center servers. Of the four methods, the method combining Resource Based Distribution technique and Duration Priority technique have shown the best performance getting the minimum tardiness while complying with the problem constraints. The model which we discuss in this paper is minimizing the average tardiness of connection requests and

minimizing the connection blocking percentage. The method in which we combine Resource Based Distribution technique and Duration Priority Technique (RB-DP) give best performance in term of request blocking percentage while complying with the problem constraint. The method in which Equal Time Distribution Technique and Duration Priority Technique (ED-DP) have shown the best performance in term of average tardiness per request while complying with the problem constraints [5].

In [6], author gives idea about Genetic algorithm based resource scheduling for that there is a fitness function which divided into three sub-fitness function and then linear combination of these sub-fitness value is done for getting the fitness value. Some problem like fragmentation of resources, low utilization of the resources are easily cause because of the expansion of system scale and virtual machines' migrations, etc because of this high energy consumption within an Internet Datacenter. The strategy which discuss in this paper is pre-migration strategy which is based on three load dimension: CPU utilization, network throughput, disk I/O rate. TO obtain an approximately best solution we apply the hybrid genetic algorithm merge with knapsack problem with multiple fitness and experiments are conducted which prove the efficiency of the algorithm. The result shows that the algorithm can effectively achieve the goal of raising resources utilization and lower energy consumption under cloud computing environment. The algorithm will save energy consumption as well as increase the utilization of the resources. The execution of this algorithm with the virtual machine's actual migration will save 30-40% of the total physical machine's occupation as well as smooth the utilization of the loads. By dividing overall load into specific load we get more precise result about the virtual machine's placement.

III. ALGORITHM

To solve the problem of resource optimization within the cloud I proposed a cloud computing resource scheduling policy based on genetic algorithm with multiple fitness. To validate the efficiency of the algorithm I proposed a typical cloud computing model which is same as other cloud computing environment. For management of all the physical resources and storage systems there is a cloud manager as the cloud management entry point. On every physical machine a virtualization layer called hypervisor is installed and all the virtual machine is created above this. Some user interfaces are provided and user can access the underlying resources via the interfaces. The scheduling policy which explains below is executed regularly and the result contains a reasonable virtual machine's migration schemes, this result used as the administrator's references or it can be directly used to migrate the virtual machines automatically [9]. The conceptual topology is showed as fig.1 where PM indicates the physical machine and VM indicate the virtual machine.

As shown in fig1 under cloud manager layer several physical machines and every physical machine carries several virtual machines. The cloud manager is containing several modules such as network module, image module, and certification module and so on. The HGA module which lies parallel to this module is combined with them to provide services. In HGA not all physical machines and virtual machines are marked out. The punctuation '…'indicates more unidentified physical machines or virtual machines. In GA there is fitness value which calculated from various fitness functions based on this individual selected for the next stage like for selection function, mutation function, and cross-over function.

The flow diagram of the resource scheduling policy is shown in fig.2. As Shown in flow chart first the system should fetch the load information and original placement information of all the virtual machines for the next stage. Then HGA module is executed the result of this module should be compared with the original placement information to decide whether the result is satisfied enough. The criteria for judging could be a threshold of the number of physical machine that will be saved or if the utilization is raised up to a certain level. As per the emphasis the threshold value could be different. At the end another judging conducted which help to judge whether the migration cost is higher than the benefits of the scheduling. If migration cost is low then the migration can be actually conducted.

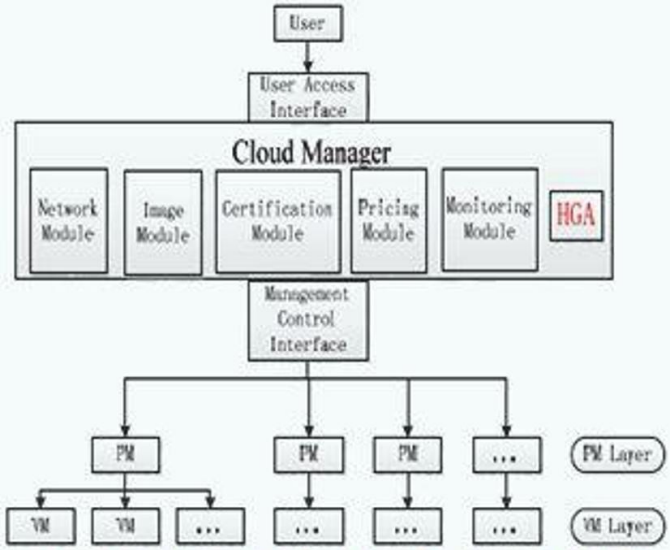
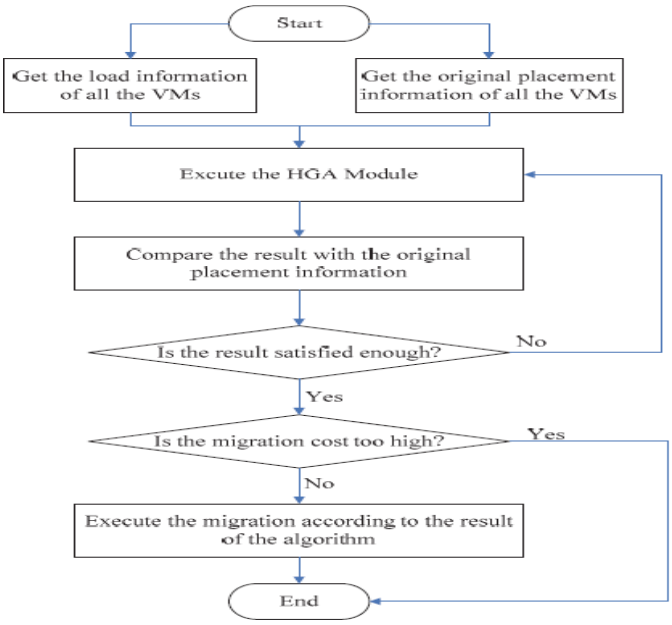


Figure 1. A conceptual topology graph of the typical cloud computing environment



The base of this algorithm is that the three load dimensions: CPU load, network throughput and disk I/O load of all the virtual machines carried on one specific physical machine, can be matched and calculated to get the optimal migration advice. So it is necessary that the system must monitor the three dimensions, which fetched from the monitoring-database when the algorithm is executed. In this algorithm three fitness functions are adopted Sub fitness function 1 represents the virtual machine's load completion. Sub fitness function 2 represents the amount of physical machines which is minimum after migration. Sub fitness function 3 represents the amount of virtual machines that need to be migrated is minimum. The combination of first two function give maximum resources utilization and the last sub fitness function consider the migration's cost.

The process of the algorithm

1. Generate the initial population P
2. Repeat:
3. For each individual i in P
4. Calculate sub fitness value-1

5. Calculate sub fitness value-2

6. Calculate sub fitness value-3

7. Calculate the overall fitness value of i

8. Keep the current best individual directly into the offspring population

9. Conduct selection function to P according to the fitness value

10 Conduct cross function to P

11 Modify invalid chromosomes after cross

12 Conduct mutation function to P

13 Modify invalid chromosomes after mutation

14 Update the population P to new generation

15 If the maximum evolution generation is reached, go to 16, else, go to 2

16. Output the current best individual as the optimal result

IV. CONCLUSION

From the review of these research papers I can conclude that in cloud computing resource scheduling is an important task for the resource management. These resources are available for fixed intervals of time and at fixed time. So if we want to find the optimized scheduling of resources then it required to maintain separation between users of the resources. In cloud computing there are number of computer with varying capacities so to meet the users requirements and improve the resource utilization there is a need of efficient resource scheduling. The cloud provider receives a number of computing requests with different requirements and preferences from users. Some tasks fulfilled at a lower cost and less computing resources while some required more bandwidth and computing resources and higher computing resources. Proper resource scheduling will reduce cost and time needed for processing. The proposed algorithm will save energy consumption as well as increase the utilization of the resources. The execution of this algorithm with the virtual machine's actual migration will save 30-40% of the total physical machine's occupation as well as smooth the utilization of the loads. By dividing overall load into specific load we get more precise result about the virtual machine's placement. The algorithm achieves the goal of improving resource utilization and saving energy cost by runtime resource scheduling. With the combination of the algorithm described above and live migration, the resource utilization can reach an ideal situation and thus some extra energy can be saved.

Acknowledgement

First we would like to thanks my guide Prof. Hitesh A. Bheda who supported me throughout my research with his patience and guidance. We would also like to thank Head of the department Prof. Amit Lathigara for helping me in sorting out the procedural work and his guidance.

References

1. Mayank Singh Rana, Sendhil Kumar KS, Jaisankar N,Comparison of Probabilistic Optimization Algorithms for resource scheduling in Cloud Computing Environment, International Journal of Engineering and Technology (IJET), Vol 5 No 2 Apr-May 2013, pp 1419 – 1427.
2. Vignesh V, Sendhil Kumar KS, Jaisankar N ,Resource Management and Scheduling in Cloud Environment, International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013.
3. Hilda Lawrance, Dr. Salaja Silas , Efficient Qos Based Resource Scheduling Using PAPRIKA Method for Cloud Computing, Hilda Lawrance et al. / International Journal of Engineering Science and Technology (IJEST).

4. Veena Goswami, Sudhansu Shekhar Patra, G.B.Mund ,Optimization of QOS parameters through flexible Resource Scheduling in Finite Population Cloud Environment, International Journal of Cloud Computing and Services Science (IJ-CLOSER) Vol.2, No.3, June 2013, pp. 170-183.
5. Mohamed Abu Sharkh, Abdelkader Ouda, and Abdallah Shami, A Resource Scheduling Model for Cloud Computing Data centers, 978-1-4673-2480-9/13/2013 IEEE.
6. Shi Chen, Jie Wu, Zhihui Lu , A Cloud Computing Resource Scheduling Policy Based on Genetic Algorithm with Multiple Fitness, 2012 IEEE 12th International Conference on Computer and Information Technology.
7. Ravindra K. Ahuja, James B. Orlin , Ashish Tiwari , A greedy genetic algorithm for the quadratic assignment problem, Computers & Operations Research 27 (2000) 917-934.
8. Sung Ho Jang, Tae Young Kim, Jae Kwon Kim and Jong Sik Lee ,The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing, International Journal of Control and Automation Vol. 5, No. 4, December, 2012.
9. Jing Tai Piao, Jun Yan, A Network-aware Virtual Machine Placement and Migration Approach in Cloud Computing, 2010 Ninth International Conference on Grid and Cloud Computing.
10. Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose, and Rajkumar Buyya ,Cloud Sim: A Toolkit for the Modeling and Simulation of Cloud Resource Management and Application Provisioning Techniques, Software: Practice and Experience, 41(1): 23-50, Wiley, January 2011.
11. Sung Ho Jang, Tae Young Kim, Jae Kwon Kim and Jong Sik Lee ,The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing, International Journal of Control and Automation Vol. 5, No. 4, December, 2012.
12. Michael Nelson, Beng-Hong Lim, and Greg Hutchins "Fast Transparent Migration for Virtual Machines" USENIX Annual Technical Conference, 2005

AUTHOR(S) PROFILE

Hitesh A. Ravani, completed his B.E degree in Computer Science from Faculty of Technology, M.S University, Vadodara, Gujarat, India in 2011. And Presently he is student of M.TECH (Computer Engineering) School of Engineering, RK University, Rajkot, Gujarat, India.

Hitesh A. Bheda, completed his B.E degree in Information Technology from V.V. P Engineering College, Saurast University, Rajkot, Gujarat, India in 2009 and M.E in Information Technology from L.D Engineering College, G.T.U, Ahmadabad, Gujarat, India in 2012.

Presently he is working as Assistance Professor in Computer Engineering Department in RK University, Rajkot, Gujarat, India.

Vrunda J. Patel, completed her B.E degree in Computer Science Engineering from Mahatma Gandhi Institute Of Technical Education And Research Center, Veer Nar mad South Gujarat University Surat, Gujarat, India in 2011. And Presently She is student of M.TECH (Computer Engineering) School of Engineering, RK University, Rajkot, Gujarat, India.

