Multi-Criteria Strategy for Job Scheduling and Resource Load Balancing in Cloud Computing Environment

Aminu Abdulkadir Mahmoud*, M. Zarina, Wan Nor Shuhadah Wan Nik and Fadhilah Ahmad
Department of Computer Science, Faculty of Informatics and Computing, University Sultan Zainal Abidin, Malaysia; si0453@putra.unisza.edu.my, zarina@unisza.edu.my, wnshuhadah@unisza.edu.my, fad@unisza.edu.my

Abstract
Cloud computing is growing rapidly over the years and it faces challenges especially in resource management. Resource management in cloud computing is necessary due to its distributed nature with different user demands. Quality of Service (QoS), load balancing and throughput are identified as some of the benefits of proper resource management. This research focuses on job scheduling and resource load balancing in cloud environment. We proposed an efficient algorithm based on multi-criteria strategy. The algorithm consists of two main phases. In the first phase the shortest job completion time is measured based on the completion time of three techniques i.e. min-min, max-min and suffrage. Meanwhile in the second phase genetic algorithm is implemented for resource load balancing. Cloud Sim simulator is used to measure the performance and efficiency of the proposed algorithm. The proposed algorithm enhances jobs scheduling and resource load balancing by ensuring an efficient utilization of the available resources.

Keywords: Cloud Computing, Genetic Algorithm, Job Scheduling, Load Balancing, Virtual Machine

1. Introduction
Cloud Computing is a new area of computing and interested field of research nowadays\(^1\). Cloud Computing is a term referring to applications, development platforms, infrastructures, storages, communications and collaborations over the internet on-demand via services offered by cloud providers\(^2\text{-}^4\). However Cloud computing has become an important pillar in the field of IT. Cloud computing in modern days computing environment can promise number of benefits by providing easy and affordable cost of IT equipment in terms of software and hardware. Furthermore, it can also reduce the cost of outsourcing, data backup and scarce IT talent by providing steady access for cloud users\(^5\).

Despite of the benefits and expectations Cloud computing can offer, it has some grey areas that need to be improved. This is especially true when cloud computing environment is heterogeneous in nature and its resources are geographically distributed.

Cloud computing brings together large amount of data which needs to be efficiently managed in order to provide on demand services. Job scheduling is a part of data management in cloud environment which is associated with numerous challenges due to the nature of cloud environment. Cloud environment consists of huge amount of data, distributed storage and demand of good services; which can provide high reliability and availability fairness of services and better access performance over the internet.

Cloud computing resources need to be managed and utilized properly to achieve desired results. This research focused on job scheduling as part of cloud resource management. We consider the distributed nature of cloud environment where different datacenters may be located at different geographical locations with their hosts and Virtual Machines (VMs). In addition to the characteristics
of datacenters, hosts and VMs, an efficient algorithm needs to be proposed that can enhance job scheduling and resource load balancing. The proposed algorithm implements three scheduling techniques i.e. min-min, max-min and suffrage to find the minimum makespan. Furthermore, the proposed algorithm will also take advantage on the benefits of genetic algorithm for resource load balancing. The idea on balancing the resource workload will indirectly improve job response time, resource availability, resource throughput and of course the performance of the cloud environment as whole. This can be achieved by avoiding extreme utilization (i.e. under utilization or over utilization) of some cloud resources.

2. Related Works

There are several publications and research papers on this new field of modern day computing with a lot of expectations on its deliverance. This is especially true when it aims to minimize the makespan of the job(s). Research proposed hosts scheduling by implementing the Suffrage algorithm coupled with Genetic Algorithm. The objective of the research is to minimize the make span of the job(s)\textsuperscript{10}.

Another research was conducted with the aim of addressing load balancing problems in VM resources scheduling\textsuperscript{11}. In this research, job scheduling strategy is proposed in order to load balance the VM resources based on genetic algorithm. In addition it also addressed the problems of high migration cost after jobs are scheduled on VMs. Meanwhile research, proposed an algorithm using two conventional task scheduling algorithm i.e. min-min and max-min algorithms\textsuperscript{12}. These algorithms uses certain heuristics to select between the two conventional algorithms so that the overall make span is minimized and the tasks are scheduled on machines in either space shared or time shared manner and it was compared with FCFS scheduling technique.

Furthermore, research based on job scheduling in cloud computing environment, where a generalized priority algorithm was proposed for efficient job execution\textsuperscript{13}.

Table 1. Summary of scheduling techniques in cloud computing

| Author/Year | Method and Issue | Strength | Weakness |
|-------------|------------------|----------|----------|
| Goyal T. and Agrawal A., 2013\textsuperscript{10} | The researcher proposed the use of Suffrage coupled with Genetic Algorithm. The objective of the research is to minimize the make span of the job(s) | The proposed algorithm is able to find minimum completion time for a given task(s). | The proposed algorithm does not concern about load balancing between the virtual machines. |
| Gu J., Hu J., Zhao T. and Sun G., 2012\textsuperscript{11} | The genetic algorithm was proposed in order to address the problems of load balancing and high migration cost. The aim of this research is to address the problems of load balancing in Virtual Machine (VM) resource scheduling. | This strategy is able to solve the problem of load imbalance and high migration cost | The strategy does not take into consideration the minimum completion time of a given task(s). |
| Katyal M. & Mishra A. 2014\textsuperscript{12} | An algorithm is proposed using two conventional task scheduling algorithm. This algorithm is based on min-min and max-min conventional scheduling techniques. The aim of the proposed technique is to minimize the overall makespan of the jobs | The algorithm in this research ensures that the jobs are executed with smaller delays to minimize the overall makespan | Not much improvement been made from FCFS. Some improvement is still needed to balance the load between the VMs and to minimize the execution time |
| Agarwal A. & Jain S. 2014\textsuperscript{13} | The proposed algorithm i.e. Generalized Priority Algorithm rank (prioritized) the virtual machines according to their Million Instructions Per Second (MIPS) and the cloudlets according to their sizes and it. The aim of the research was to propose algorithm for scheduling that would help to schedule task efficiently in Cloud Computing Environment. | The proposed Algorithm was found to be better in performance than FCFS and Round Robin | This Algorithm prioritized the job scheduling based on MIPS of VM and cloudlets size without considering the completion time and load balancing. |
The performance of the proposed strategy was compared with First Come First Serve (FCFS) and Round Robin Scheduling techniques. When tested in CloudSim toolkit, the result shows that it gives better performance than other traditional scheduling algorithms.

3. Proposed Multi-Criteria Strategy

Multi-Criteria algorithm (RMK) is a proposed algorithm in this research where it focuses on job scheduling and resource load balancing. The algorithm consist three main stages for job execution. In the first stage, a stack table is formed containing all the cloudlets information and their execution time in the available VMs as shown in Figure 1. Meanwhile, in the second stage a minimum makespan is found by using three scheduling techniques i.e. min-min, max-min and suffrage as shown in Figure 2a, 2b and 2c respectively. Furthermore in stage three, the result from stage two will be used as parameters to be encoded in to genetic algorithm for resource load balancing as shown in Figure 3.

This algorithm is believed can enhance job scheduling and resource load balancing in cloud computing environment. As a result, this algorithm was tested (simulated) using CloudSim simulator and found to be of greater impact in term of job scheduling and resource load balancing.

3.1 Flow Chart of the Proposed Algorithm

Figure 1. General format of the proposed algorithm (RMK).

Figure 2a. Min-Min algorithm.

Figure 2b. Max-Min algorithm.

Figure 2c. Suffrage algorithm.
4. Results and Discussion

The simulation and testing of the algorithms are presented. The objectives of this research as stated is accomplished as presented here, which is proposing an intelligent job allocation strategy to enhance job scheduling and resource load balancing using genetic algorithm. The performance of each individual technique involved is presented on the same graph and was compared with the RMK algorithm and the FCFS algorithm. The RMK algorithm outperformed each of the individual technique and the FCFS algorithm as presented in Figure 4.

The results show that in each of the three different parameters FCFS is outperformed by the other techniques which were also outperformed by the RMK algorithm as can be seen in Figure 4 also. The results show that in all the techniques with exception of the RMK, the VMs with smallest MIPS are over utilized and are having the highest execution time but in RMK the reverse is the case and this will increase the response time.

| Component Type | Parameter | Value |
|----------------|-----------|-------|
| Datacenter     | Number of Datacenter | 1     |
|                | Number of Host Scheduling Policy | 1     |
|                | Scheduling Policy | Space-Shared |
| Datacenter Broker | Number of Datacenter Broker | 1    |
| VMs            | Number of VMs | 4     |
|                | Number of Processing Element (PE) | 1     |
|                | MIPS of PE | 2000  |
|                | MIPS of VMS |       |
|                | VM 1      | 200   |
|                | VM 2      | 300   |
|                | VM 3      | 400   |
|                | VM 4      | 500   |
| VM RAM size    | VM Scheduling Policy Bandwidth | 2048MB Space-Shared |
| Cloudlet       | Number of Cloudlets | 10    |
|                | CL 1      | 5200  |
|                | CL 2      | 3000  |
|                | CL 3      | 2000  |
|                | CL 4      | 7000  |
|                | CL 5      | 6800  |
|                | CL 6      | 6600  |
|                | CL 7      | 5800  |
|                | CL 8      | 5400  |
|                | CL 9      | 5000  |
|                | CL 10     | 4800  |
| Cloudlet Scheduler Policy | Space-Shared Full Utilization of CPU |
| CPU Utilization Model | Full Utilization of RAM |
| RAM Utilization Model | Full Utilization of Bandwidth |
| Bandwidth Utilization Model |   |

Table 3. Parameters setting for CloudSim simulator
5. Conclusion

This research is conducted with the aim of finding an algorithm that is capable to enhance jobs scheduling and resource load balancing in cloud computing environment. Three scheduling techniques are integrated i.e. min-min, max-min and suffrage with GA. This integration brings about the proposed algorithm (RMK) which has three stages for execution. Stack table containing execution time of jobs mapped to the available VMs, finding minimum makespan and load balancing using GA are three stages of this algorithm.

The proposed algorithm can enhance jobs scheduling and resource load balancing as shown in Table 2 and Table 3 in the previous section.

6. References

1. Available from: www.google.com/trends [cited 2014 May 30].
2. Erkoc MF, Kert SB. Cloud Computing for distributed university campus: A prototype suggestion. International Conference the Future of Education.
3. Kalpeyeva ZB, Mustafina AK. IT-infrastructure of university based on cloud computing. International Journal of Computer Science Issues. 2013 Sept; 10(5):176–9.
4. Vouk MA. Cloud computing issues, research and implementations. Journal of Computing and IT. 2008; 4:235–6.
5. Kart R, Goldstein P, Yanosky R. Cloud computing in higher education.
6. Moothoor J, Bhatt V. A cloud computing solution for universities: Virtual computing lab. IBM Corporation; 2009.
7. White Paper. Introduction to Cloud Computing. Montreal, Quebec, Canada: Dialogic Corporation; 2010.
8. Huth A, Cebula J. The basics of cloud computing. Carnegie Mellon University produced for US-CERT, a Government Organization; 2011.
9. Bento A, Bento R. Cloud computing: A new phase in information technology management. Journal of Information Technology Management. 2011; 22(1):39–46.
10. Goyal T, Agrawal A. Host scheduling algorithm using genetic algorithm in cloud environment. International Journal of Research in Engineering and Technology. 2013 June; 1(1): 7–12.
11. Gu J, Hu J, Zhao T, Sun G. A new resource scheduling strategy based on genetic algorithm cloud computing environment. Journal of Computers. 2012 Jan; 7(1): 42–52.
12. Katyal M, Mishra A. Application of selective algorithm for effective resource provisioning in cloud computing environment. International Journal on Cloud Computing: Services and Architecture (IJCCSA). 2014 Feb; 4(1).
13. Agarwal A, Jain S. Efficient optimal algorithm of task scheduling in cloud computing environment. International Journal of Computer Trends and Technology (IJCTT). 2014 Mar; 9(7):344–9.
14. Sun H, Chen S, Jin C, Guo K. Research and simulation of task scheduling algorithm in cloud computing. TEL-KOMINA. 2013 Nov; 11(11):6664–72.