IJCSMS (International Journal of Computer Science & Management Studies) Vol. 15, Issue 06
Publishing Month: June 2015
An Indexed and Referred Journal with ISSN (Online): 2231–5268
www.ijcsms.com

Enhanced Job Scheduling Algorithm for Cloud Computing Using Shortest Remaining Job First (SRJF)

Mohammed Aboalama[1] and Adil Yousif[2]

[1]University of Garden City, Department of Computer Science, Khartoum, Sudan
mohammedribat@gmail.com

[2]University of Science &Technology, Department of Computer Science, Omdurman, Sudan
adiluofk@gmail.com

Abstract

Cloud computing environments facilitate applications by providing virtualized resources that can be provisioned dynamically. Computing resources are delivered by Virtual Machines (VMs). In such a scenario, resource scheduling algorithms play an important role where the aim is to schedule applications effectively so as to reduce the turn-around time and improve resource utilization. The problem of this paper is how to dynamically allocate the submitted jobs to the available resources in order to complete the tasks within a minimum turn-around time as well as utilizing cloud resources effectively. The objective of this paper is to propose new scheduling algorithm on cloud computing environment using Shortest Remaining Job First (SRJF) algorithm. The methodology of this paper depends on simulation using cloudsim. The results of this paper revealed that the proposed algorithm (SRJF) performed better than the default scheduling algorithm.

Keywords: *SRJF, turn-around time, Job Scheduling, Cloud Computing.*

1. Introduction

Cloud computing is an on demand service in which shared resources, information, software and other devices are provided according to the clients requirement at specific time[1]. It’s a term which is generally used in case of Internet. The whole Internet can be viewed as a cloud. Capital and operational costs can be cut using cloud computing. Cloud Computing is an entirely new computing service mode which is rose in the 21st century[2].Resource scheduling strategy in cloud environment is thus an important issue affecting not only the performance of cloud but also the turnaround time experienced by its clients and the prices paid by them. Scheduling is used to allocate particular resources for a certain tasks in particular time. Job scheduling problem is a core and challenging issue in cloud computing. The job execution time cannot be predicted in cloud computing. Hence the scheduler must be dynamic. The purpose of scheduling is to increase the utilization of resources[3].

The rest of the paper is organized as followed: section 2 Cloud Scheduling Algorithms, Section 3 Related Works, Section 4 Proposed Algorithm and Section 5 presents the Experiment and Results, The section 6 provide the Conclusion and Future works.

2. Cloud Scheduling Algorithm

Cloud service scheduling is categorized at user level and system level[4] .At user level scheduling deals with problems raised by service provision between providers and customers[5]. The system level scheduling handles resource management within datacenter[4]. Ant algorithms are one of the most popular examples of swarm intelligence systems[6].a Scheduling model based on minimum network delay using Suffrage Heuristic coupled with Genetic algorithms for scheduling sets of independent jobs algorithm is proposed, the objective is to minimize the make span. The Suffrage heuristic is based on the idea that better mappings can be generated by assigning a machine or resource to a task that would “suffer” most in terms of expected completion time if that particular machine is not assigned toit. Let the suffrage value of a task or job it be the difference between its second earliest completion time and its earliest completion time[7]. Genetic algorithm is based on biological concept of generation of the population, a rapid growing area of Artificial

intelligence[8]. GA's are inspired by Darwin's theory about Evolution. According to the Darwin "Survival of the fittest". It also a used as the method of scheduling in which the tasks are assigned resources according schedules in context of scheduling, which tells about which resource is to be assigned to which task[7]. Genetic Algorithm is based on the biological concept of population generation[8] .cloud task scheduling based on ACO, The basic idea of ACO is to simulate the foraging behavior of ant colonies[6]. When an ants group tries to search for the food, they use a special kind of chemical pheromone to communicate with each other .Task scheduling based ACO algorithm is used to decrease the computation time of tasks. In ACO, all ants are placed at the starting VMs randomly. The MACO algorithm inherits the basic ideas from ACO algorithm to decrease the computation time of tasks executing[9].
Resource Scheduling in Cloud Computing as shown in Figure1[10].

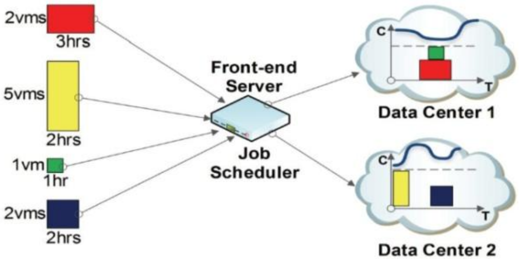


Figure 1: Resource Scheduling in Cloud Computing Environment[10].

3. Related Works

Cloud computing can be defined as "a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers" .Some examples of emerging Cloud computing infrastructures are Microsoft Azure ,Amazon EC2, Google App Engine, and Aneka The computing power in a Cloud computing environments is supplied by a collection of data centers, which are typically installed with hundreds to thousands of servers[11] .The layered architecture of a typical Cloud based data center . At the lowest layers there exist massive physical resources (storage servers and servers are transparently managed by the higher level Virtualization services and toolkits that allow sharing of their capacity among virtual instances of servers[9]. These virtual instances are isolated from each other, which aid in achieving fault tolerant behavior and isolated security context.
Emerging Cloud applications such as Social networking, gaming portals, business applications, content delivery, and scientific workflows operate at the highest layer of the architecture[11]. Actual usage patterns of many real-world applications vary with time, most of the time in unpredictable ways. These applications have different Quality of Service (QoS) requirements depending on time criticality and users' interaction patterns (online/offline)[12].

4. Proposed Algorithm

Shortest remaining time, also known as shortest remaining time first **(**SRTF**)**, is a scheduling method that is a preemptive version of shortest job next scheduling. In this scheduling algorithm, the process with the smallest amount of time remaining until completion is selected to execute. Since the currently executing process is the one with the shortest amount of time remaining by definition, and since that time should only reduce as execution progresses, processes will always run until they complete or a new process is added that requires a smaller amount of time[13].

4.1 Pseudo Code

Pseudo code of proposed algorithm as follows:

```
Begin
Initialize; executing the first coming
process to CPU;
If
        New coming process is shorter
        than the process that is
        currently process executing;
Then
        CPU will be de-allocated from
        currently executing process;
        CPU will be allocated to the
        process which is shortest;
End.
```

5. Experiment and Results

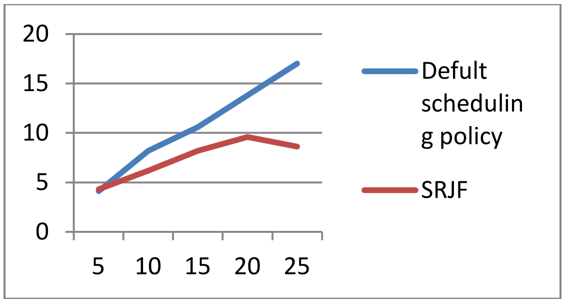
CloudSim is a simulation toolkit for Cloud computing. The default scheduling policy provided in CloudSim schedules sequentially between a list of virtual machines and a list of tasks[9].By extending cloudletSchedulerSpaceSharedclass, we can design the proposed scheduling policy based on SRJF[4].We use turnaround time (the total finish time of the tasks) to evaluate the performances of the default policy and the proposed policy.

5.1 The First Scenario

In this scenario we considered number of resource that are equal number of tasks.

Table 1: Average Turnaround Time in MIPS

	Number Of task and VM				
Algorithm	5	10	15	20	25
default scheduling policy	4.1	8.2	10.6	13.8	17
SRJF	4.3	6.2	8.2	9.6	8.7



The number of tasks

Figure 2: Comparison of Simulation Results in First Case.

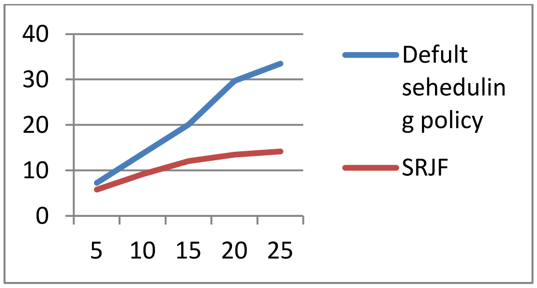
As described in table 1 and fig 2; the number of tasks equal number of resources. In this case comparison between default scheduling policy and SRJF algorithm; SRJF algorithm performed better the default scheduling policy based on turnaround time. Exception in case (5, 5) default scheduling policy performed better the SRJF algorithm.

5.2 The Second Scenario

In this scenario we considered number of resource that are not equal number of tasks.

Table 2: Average Turnaround Time in MIPS

	Number of task **Number of VM=3(const) to all task**				
Algorithm	5	10	15	20	25
default scheduling policy	7.3	13.7	20.1	29.7	33.5
SRJF	5.8	9.2	12.1	13.5	14.2



The number of tasks

Figure 3: Comparison of Simulation Results in Second Case.

As described in Table 2 and Figure 3: the number of tasks is not equal number of resources, Also SRJF algorithm performed better the scheduling policy based on turnaround time.

5.3 Discussion

The SJF algorithm can be either preemptive or non preemptive. The choice arises when a new process arrives at the ready queue while a previous process is still executing. The next CPU burst of the newly arrived process may be shorter than what is left of the currently executing process. Preemptive SJF algorithm will preempt the currently executing process, whereas a non preemptive SJF algorithm will allow the currently running process to finish its CPU burst. Preemptive SJF algorithm is called Shortest-Remaining-Job-First (SRJF) algorithm.

6. Conclusion and Future work

6.1 Conclusion

In this paper we propose a task-scheduling policy based on SRJF and evaluated it in a Cloud simulator with the number of tasks varying from 5 to 25. The results of above experiment demonstrate the effectiveness of the algorithm. To define your own policy, you have to extend either VmScheduler or Cloudlet Scheduler, create the methods for deciding sharing of PEs and pass the new class during construction of the relevant object.

6.2 Future Work

Cloud computing is a vast concept and scheduling a very important role in case of Cloud computing. There is a huge scope of improvement in this area. Paper have discussed only one proposed new scheduling algorithms that can be applied to cloud, but there are still other scheduling algorithms that can be applied in cloud computing environment. The performance of the given algorithms can also be increased by varying different parameters.

References

[1] Padhy, R.P., Load Balancing in cloud computing systems, 2011, National Institute of Technology, Rourkela.
[2] Computing, C., a practical Approach, Anthony T. Velte, Toby J. Velte, Robert Elsenpeter, McGraw Fill, 2010.
[3] Zhang, H., et al., A PSO-Based Hierarchical Resource Scheduling Strategy on Cloud Computing, in Trustworthy Computing and Services. 2013, Springer. p. 325-332.
[4] The CLOUDS (2012) Lab: CloudSim:. 2012.
[5] Ercan, T., Effective use of cloud computing in educational institutions. Procedia-Social and Behavioral Sciences, 2010. 2(2): p. 938-942.
[6] Mishra, R. and A. Jaiswal, Ant colony optimization: A solution of load balancing in cloud. International Journal of Web & Semantic Technology (IJWesT), 2012. 3(2): p. 33-50.
[7] GOYAL, T. and A. AGRAWAL, Host Scheduling Algorithm Using Genetic Algorithm In Cloud Computing Environment. International Journal of Research in Engineering & Technology (IJRET) Vol. 1.
[8] Pandey, S., et al. A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments. in Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on. 2010. IEEE.
[9] Calheiros, R.N., et al., CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software: Practice and Experience, 2011. 41(1): p. 23-50.
[10] Li, Q. and Y. Guo. Optimization of Resource Scheduling in Cloud Computing. in SYNASC. 2010.
[11] Calheiros, R.N., et al., Cloudsim: A novel framework for modeling and simulation of cloud computing infrastructures and services. arXiv preprint arXiv:0903.2525, 2009.
[12] Cardoso, A.J.S., Quality of service and semantic composition of workflows, 2002, University of Georgia Athens, Georgia.
[13] Silberschatz, A., et al., Operating system concepts. Vol. 4. 1998: Addison-Wesley Reading.