

Exploration of Task Scheduling Algorithms in Cloud Computing Environments

A. Saravanan, M. Murali

Abstract: Cloud computing is a model that provides computing environment where a shared pool of resources are provisioned to customers as a service via the internet. Task scheduling forms the essential part of cloud environment. The scheduling of task in cloud mainly focuses on reducing average waiting time, Makespan and maximizes resource utilization which in turn reduces the response time of the system. This paper study various scheduling algorithms in cloud computing environments.

Keywords: Cloud Computing, Resource Utilization.

I. INTRODUCTION

Cloud computing is a model of computing services that can be provisioned to users via the internet. The services provisioned to the users can be accessed through end user devices running different applications. The users are charged by the cloud service providers on pay per use basis. Examples of cloud service providers include Google, Amazon, and Microsoft.

Cloud services are provided to users in form of infrastructures as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS).

In infrastructure as a service the physical components of the cloud are provided to the user as virtual machines and virtual storage instances for program execution and data storage. The platform as a service model provide the environment for the development and deployment of user applications in the cloud and the software as a service model provide the user required applications accessible via web browsers as a service.

A. Cloud Architecture

Cloud architecture is designed to have many components that are loosely coupled. Cloud components can be divided into Front End and Back End. The Front End components comprises of clients, mobile device that provides the applications and interfaces to access the cloud platform. The Back End components include the Servers, Storage, Virtual Machines, applications, Datacenters that are provisioned to the users as a service [2].

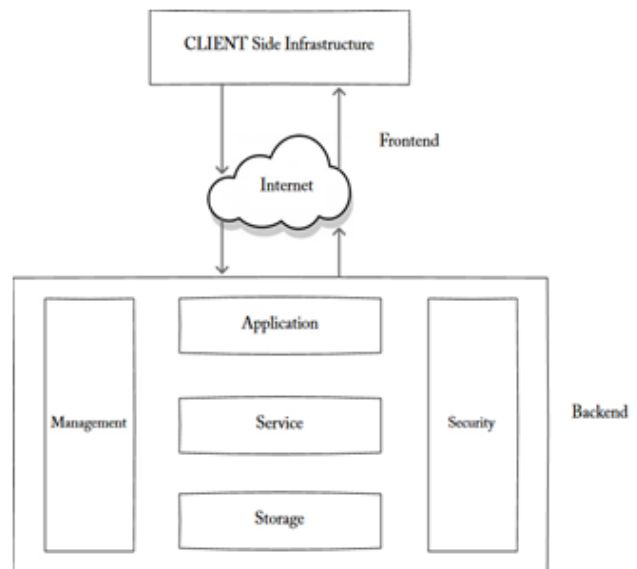


Figure 1: Cloud Architecture (Source [2])

B. Scheduling in Cloud Computing

Resource scheduling is the allocation of a resource to a user task at a particular period of time. The main goal of resource scheduling is to minimize waiting time, execution time and increase resource utilization which in turns reduces the response time of the overall system.

C. Types of Resource Scheduling

Resource scheduling is classified into static and dynamic scheduling. In static scheduling the scheduler is aware of the resources required by a task before its execution. In dynamic scheduling the scheduler is unaware of the resources required by the task.

II. LITERATURE SURVEY

Cloud computing consist of different scheduling algorithms. Major parameters are task length, execution time, type, deadline considered in the proposed approach. Some other parameters tempt the resources utilizations and scheduling of tasks.

A. Y. Gital. et al. [1,20] proposed a grouped task scheduling Algorithm to classify task into two groups (Urgent and Normal) based on task size, type of task, deadline. Tasks in each group are scheduled based on deadline and the two groups are scheduled based on resource need. Virtual Machines with capacity greater or equal to the average required capacity of each group are allocated to the task in that group.

Revised Manuscript Received on January 24, 2020.

A. Saravanan, Department Of Computer Science and Engineering, SRM Institute Of Science and Technology, Chennai, Tamil Nadu, India.

M. Murali, Associate Professor, Department of Computer Science and Engineering, SRM Institute Of Science and Technology, Chennai, Tamil Nadu, India.

Exploration of Task Scheduling Algorithms in Cloud Computing Environments

Y. Zhang et al. [3,15] proposed an infrastructure resource management approach in Infrastructure as a Service Cloud Model. These approach use resource consumption optimization loop to improve the utilization of resource through some management functions provided by the Middle ware layer of the Platform as a Service Model. A resource allocation loop provisioned an appropriate resource to the application system guaranteeing the performance of the resource. The resource allocation and resource consumption loop are merged to consecutively and repeatedly run to provide resource on demand by improving resource utilization and then allocate more resources when the need arise.

P. Zhang et al. [4] proposed a method that uses historical task information to determine the various types of virtual machines and the number of each type. The submitted Job Requests are classified and matched to the Virtual Machines [18] that suit the Job Request

Ajay Thomas et al. [5] proposed a scheduling method that takes into account the deadline of each task that arrives into the system. The tasks are sorted based on the deadline [15] so that each task has a deadline less than the subsequent

task. The required resource for each task is allocated to the task for execution before the deadline of the task expires. This method focuses on satisfying the client demands but fails to satisfy the cloud service provider demand.

Li et. al [6], proposed a Greedy Based Job Scheduling Algorithm that focus on improving the quality of service. The method aimed at reducing the completion time of user's task and to provide a quicker solution to the scheduling problems. The method classifies task based on the QoS and the assign an appropriate function to the task.

Patel [7], authors proposed an Improved Priority based Job Scheduling Algorithm using Iterative Method; the method gives priority to some jobs which are scheduled first for execution. The result of the proposed method was compared with the round robin algorithm and produces a better result.

Pankaj et al. [8] proposed a scheduling mechanism that takes into account the deadline of task. The proposed method allocates credit to respective tasks based on their lengths and deadlines. The mechanism finally schedule the task according to demand.

Algorithms	Objective	Scheduling Category	Scheduling Parameters	Future Work	Tool
Grouped Task Scheduling and Resource allocation[1]	To reduce Makespan, Waiting time and Increase Resource Utilization	Deadline Based Task Scheduling	Task Deadline	To consider more dynamic task scheduling	CloudSim
Delay Time Scheduling Algorithm[5]	To increase resource utilization and number of task executed	Deadline Based Task Scheduling	Task Deadline	To implement the method on real cloud environment	CloudSim
Scheduling based on two Stage Strategy[4]	To reduce Makespan, Average Waiting time and Increase Resource Utilization	Resource Demand	Task Marking to VM and Task Matching With VM	To deploy the proposed method on actual cloud environment	CloudSim
Greedy Based Job Scheduling Algorithm [6]	To improve QoS and minimize makespan	Greedy Algorithm Based Job Scheduling	Time , Bandwidth , Expectation Time, JEF Function and function result	To reduce completion time and to gain more fairness.	CloudSim
Improved Priority based Job Scheduling Algorithm using Iterative Method [7]	To reduce makespan	Priority Based Job Scheduling	Priority of Jobs	The proposed algorithm can be optimized further to reduce makespan.	CloudSim

Grouped Task Scheduling Algorithm [9]	To minimize Makespan, Waiting Time and Increase Resource Utilization	Execution Time Scheduling	user, task type, task size and task latency	To increment the number of attributes that apply QoS in the algorithm	CloudSim
Dynamic Virtual Machine Placement [10]	To reduce Resource Under utilization and Reduce SLA Violation	Dynamic VM Placement & Runtime Reallocation	CPU and Memory	considering other types of resources such as network bandwidth or storage and conducting more experiments using real workload traces	CloudSim
Resource Allocation Algorithm for Multiple VM's [11]	Maximize Cloud provider Profit	Resource Allocation Algorithm for Multiple VM's	Type of VM Required	To consider deadline and priority, consider improving utilization in an elastic environment	CloudSim
Autonomic and Energy-aware Resource Allocation [12]	To reduce power consumption	Energy aware resource allocation	Resource requirement and dynamic resource allocation	To consider resource fragmentation issue	iCloud
Task Scheduling Based On VM's Grouping [13]	Reduce Makespan and average waiting time, increase resource utilization	VM's Grouping	VM capacity	No Future Work	Cloudsim
Multi-objective Genetic Algorithm [14]	To increase resource utilization and minimize energy consumption	resource prediction and allocation	Resource requirement	to test the proposed GA with the Google data center traces to verify its prediction accuracy	MATLAB
Dynamic Cloud Task Scheduling Based on a Two-Stage Strategy [15]	To increase Resource utilization, minimize makespan and average waiting time	historical data scheduling	Task Deadline	to deploy the method on the actual cloud environment	CloudSim
Integrating Resource Consumption and Allocation for Infrastructure Resources on Demand [115]	To improve resource utilization and performance	Resource Consumption and resource Optimization Loop	Execution Time	Improving the Method to adapt more middleware and VM types and extend the optimization loop from middleware layers to Oses, and extend the allocation loop to include VM migration	SmartRod

III. TASK SCHEDULING METHODS

First Come First Serve (FCFS): each incoming task is attributed the any idle virtual mane.

Shortest Job First: the incoming task are sorted in the order of their lengths, and for the new sorted list, proceed as first come first serve.

Round Robin: all incoming task are attributed to available VM’s. The VM’s assigns quantum to each task to be executed. The execution of task on a virtual machine is interrupted when the quantum expires before the task is completely executed and the next task starts.

Min-Min: it computes the execution time matrix of each incoming task on each VM, execute the task on a VM with least execution time, and remove the task from the queue. Finally the execution time matrix is updated

Max-Min: it computes the execution time matrix of each incoming task on each VM, execute the task with maximum execution time first, and remove the task from the queue. Finally the execution time matrix is updated.

IV. RESULTS

Table 1: Comparison of Mean Execution Time

Number of Task	FCFS	SJF	Min-Min	Max-Min
20	0.9	0.97	0.84	0.91
40	0.95	0.95	0.87	0.9
60	1.02	0.98	0.93	0.95
80	1.1	1	0.95	0.97
100	1.11	1.11	1.06	1.07

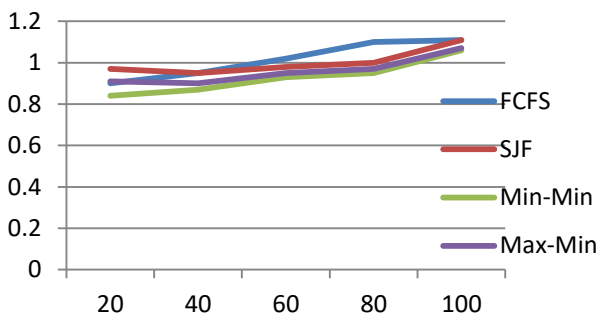


Figure 1: Comparison of Average Execution Time

V. CONCLUSION

FCFS, SJF, Min-Min and Max-Min have been used to schedule task in different small distributed systems. In this, paper an experiment was conducted using CloudSim and the methods were executed using the same dataset. The average execution times of the methods were calculated and the results were displayed under the result section of this paper.

REFERENCES

1. A. Y. Gital, Ismail Zahraddeen Yakubu, Ilya Musa, and S. Boukari, "Grouped task scheduling and resource allocation in cloud computing environments," *International Journal of Recent Technology and Engineering*, p. 12203-12206, 2019.
2. Wadhonkar and D. Theng, "A survey on different scheduling algorithms in cloud computing," *2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, 2016.
3. Y. Zhang, G. Huang, X. Liu, and H. Mei, "Integrating resource consumption and allocation for infrastructure resources on-demand," in *2010 IEEE 3rd International Conference on Cloud Computing*, pp. 75-82, IEEE, 2010.
4. P. Zhang and M. Zhou, "Dynamic cloud task scheduling based on a two-stage strategy," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 2, pp. 772-783, 2017.
5. Ajay Thomas and C. Santhiya et al., "Dynamic resource scheduling using delay time algorithm in cloud environment," in *2017 2nd International Conference on Computing and Communications Technologies (ICCT)*, pp. 55-58, IEEE, 2017.
6. Li, Ji, Longhua Feng, and Shenglong Fang. "A greedy based job scheduling algorithm in cloud computing." *Journal of Software* 9.4 (2014):921-925.
7. Patel, Swati J., and Upendra R. Bhoi. "Improved Priority Based Job Scheduling Algorithm in Cloud Computing Using Iterative Method." *Advances in Computing and Communications (ICACC)*, 2014 Fourth International Conference on. IEEE,2014.
8. P. K. Chauhan, P. Jaglan, and P. Dabas, "An intensify deadline aware credit based cloud task scheduling," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 1267-1270, IEEE, 2016.
9. H.G.E.D.H.Ali,I.A.Sarait,andA.M.Kotb,"Groupedtasksschedulingalgorithm based on qos in cloud computing network," *Egyptian informatics journal*, vol. 18, no. 1, pp. 11-19,2017.
10. A.MosaandR.Sakellariou,"Dynamicvirtualmachineplacementconsidering gcpu and memory resource requirements," in *2019 IEEE 12th International Conference on Cloud Computing(CLOUD)*, pp. 196-198, IEEE, 2019.
11. M. Li, Y.-E. Sun, H. Huang, J. Yuan, Y. Du, Y. Bao, and Y. Luo, "Profit maximizationresourceallocationincloudcomputingwithperformanceguarantee," in *2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC)*, pp. 1-2, IEEE,2017.
12. M. Shelar, S. Sane, V. Kharat, and R. Jadhav, "Autonomic and energy-aware resourceallocationforefficientmanagementofclouddatacentre," in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pp. 1-8, IEEE,2017.
13. N.Chitgar,H.Jazayeriy,andM.Rabiei,"Improvingcloudcomputingperformance using task scheduling method based on vms grouping," in *2019 27th Iranian Conference on Electrical Engineering (ICEE)*, pp. 2095-2099, IEEE,2019.
14. F.-H.Tseng,X.Wang,L.-D.Chou,H.-C.Chao,andV.C.Leung,"Dynamicresource predictionandallocationforclouddatacenterusingthemultiobjectivegenetical- gorithm," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1688-1699,2017.
15. Y. Zhang, G. Huang, X. Liu, and H. Mei, "Integrating resource consumption and allocation for infrastructure resources on-demand," in *2010 IEEE 3rd International Conference on Cloud Computing*, pp. 75-82, IEEE,2010.
16. Varsha Gupta, M. Murali, "Key -summative searchable encryption (KSSE) for faction information distribution via cloud storage" ,*International Journal Of Control theory and applications* ,Vol 9 ,13,PP.6389-6334,2016.
17. Ankita sadh, ,M.Murali ,,"wireless sensor data access through mobile cloud computing" *International Journal Of Control theory and application*, Vol 9 ,15 ,PP-7325-7331,2016
18. Shaik Saleem , M Murali , "Privacy Processing auditing public auditing for data integrity in cloud", *Journal Of physics*,1000(1),012164,2018.
19. Venisha A ,M Murali, "Discovering the trustworthy cloud service provider in collaborative cloud environment", *International journal of Engineering and advanced technology* ,Vol 8(22),360-365 ,2019.

20. Venisha A ,M Murali , “A Conception for identifying trust service providers in collaboration Clouds Environment “ , International journal of Recent technology and Engineering ,Vol 8(24),110-116 ,2019.

AUTHORS PROFILE



A. Saravanan, Pursuing M.Tech in Computer Science and Engineering SRM Institute of Science and Technology , Chennai , Tamil Nadu , India.



M. Murali, Associate Professor, Department of Computer Science and Engineering SRM Institute of Science and Technology , Chennai , Tamil Nadu , India.