Available Online through   *Review Article*
www.ijptonline.com

ANALYSIS OF TASK SCHEDULING FOR PRIVATE AND PUBLIC CLOUD IN REAL-TIME ENVIRONMENT

[1*]**Sushant Nair,**[2] **P. Akilandeswari**
[1]PG student, CSE department, SRM University, Kattankulathur.
[2]Assistant Professor, CSE department, SRM University, Kattankulathur.
Email:nsushant007@gmail.com

Abstract:

Task scheduling plays a key role in cloud computing systems. Scheduling of tasks cannot be done on the basis of single criteria but under a lot of rules and regulations that we can term as an agreement between users and providers of cloud. This agreement is nothing but the quality of service that the user wants from the providers. Providing good quality of services to the users according to the agreement is a decisive task for the providers as at the same time there are a large number of tasks running at the provider's side. The task scheduling problem can be viewed as the finding or searching an optimal mapping/assignment of set of subtasks of different tasks over the available set of resources (processors/computer machines) so that we can achieve the desired goals for tasks. In this paper we are performing comparative study of the different algorithms for their suitability, feasibility, adaptability in the context of cloud scenario, after that we try to propose the hybrid approach that can be adopted to enhance the existing platform further. So that it can facilitate cloud-providers to provide better quality of services.

Keywords: Virtualization, cloud computing, aws, ec2,Filezilla, cloud-sim.

1. Introduction

There is no exact definition of cloud but we can define cloud in various ways and by considering various means. Cloud computing is Internet-connected mode of supercomputing. It is a type of shared infrastructure, which simply puts the huge system pools together by using various means; distributed, virtualization etc. It gives users a variety of storage, networking and computing resources in the cloud computing environment via Internet, users put a lot of information and accesses a lot of computing power with the help of its own computer.

WHAT IS CLOUD?

Cloud is a parallel and distributed computing system which basically consist of a collection of inter-connected and virtualized computers that are provisioned dynamically and presented as one or more than one unified computing resources based on service-level agreement (SLA) established through negotiation between the service providers of cloud and users [1].

Cloud computing is a large-scale distributed computing model, which depends on the economic size of the operator of cloud that is abstract, virtualized and dynamic. The main content of cloud computing is to manage computing power, storage, various kind of platforms and services which assigned to the external users on demand through the internet. Cloud computing is a rapidly emerging computation paradigm with the goal of freeing up users of cloud from the management of hardware, software, networks and data resources and shifting these burdens to cloud service providers[2]. Clouds provide a very large number of resources, including platforms for computation, data centers, storages, Networks, firewalls and software in form of services. At the same time it also provides the ways of managing these resources such that users of cloud can access them without facing any kind of performance related problems. Cloud Computing Services are divided into three classes, according to the abstraction level and the service model of providers, namely:

(1) Infrastructure as a Service,

(2) Platform as a Service,

(3) Software as a Service.

Distribution, virtualization and elasticity are the basic

characteristics of cloud computing. Virtualization is one of the main features of cloud. Most of the software and hardware have provided support to virtualization. We can perform virtualization on many factors such as hardware, software, storage and operating system, and manage them in cloud platform.

Scheduling is the one of the most prominent activities that executes in the cloud computing environment. To increase the efficiency of the work load of cloud computing, scheduling is one of the tasks performed to get maximum profit. The main objective of the scheduling algorithms in cloud environment is to utilize the resources properly while managing the load between the resources so that to get the minimum execution time.

There are so many algorithms for scheduling in cloud

computing. The main advantage of scheduling algorithm is to obtain a high performance. The main examples of scheduling algorithms are FCFS, Round-Robin, Min-Min algorithm, Max-Min algorithm and meta- heuristic algorithms (ACO, GA, Simulated annealing, PSO, Tabu search and many more).

FCFS: First come First serve basis means that task that come first will be execute first.

Round-Robin algorithm (RRA): In this Scheduling algorithm time is to be given to resources in a time slice manner.

Min-Min Algorithm: Min-Min algorithm selects thesmaller tasks to be executed first.

Max-Min algorithm: Max-Min algorithm selects thebigger tasks to be executed first.

2. Literature survey

A. *Deadline and Budget Distribution based Cost-Time Optimization Algorithm [11]*

This approach considered two constraints: deadline of executing the tasks and budget. This paper proposed (DBD-CTO) scheduling algorithm to schedule tasks in cloud computing environment. This algorithm achieve its goal or target by finishing the execution of tasks before the given deadline as well as at the same time minimizes the cost of computation that occur during execution.

B. *Improved cost-based algorithm for task scheduling in Cloud Computing Environment [12]*

The author of this paper proposed the approach which is known as improved cost-based scheduling algorithm. The main objective of his work is to schedule groups of task in cloud computing platform, where resources are having different resource costs and different computation performance. When grouping of jobs is done, communication between jobs and resources optimizes computation/ communication ratio. This algorithm measured performance of computation and cost of resources. This also increased the execution of tasks / transfer of data between tasks ratio by combining various tasks during execution. The process of combining task is usually done by after analysing the capability of different available resource and its processing. CloudSim has been used for performing the simulation and the inputs of the simulation are: average MI of tasks, granularity size of tasks, total number of tasks and task overhead time. Result of his work shows that for this particular algorithm time taken to complete tasks after grouping of tasks is very less as compared to when grouping is not done.

C. *A PSO-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments [13]*

The paper gives the idea about particle swarm optimization based algorithm. Here scheduling of tasks considers both computation cost and data transmission cost. The author mainly focuses on minimizing the total cost of executing the

applications on resources that is provided by the cloud service providers. Work that is performed in this paper can be divided into two parts: 1. Formulation of a model for task-resource mapping to minimize the overall execution cost. 2. Designing a heuristics that uses PSO for solving task-resource mappings based on the above model. The proposed algorithm considers communications costs. All the tasks, including various dependencies between them.

So when a comparison is made in terms of the cost savings with the 'Best Resource Selection' (BRS) algorithm, this PSO based algorithm achieved better performance over BRS with three times savings of cost as well as good distribution of workload onto resources. The heuristics that is proposed in this paper is as generic that it can be used for any number of tasks by increasing the particle dimensions and the resources number.

D. *Multi-Objective Task Assignment in Cloud Computing by Particle Swarm Optimization [14]*

Here author of this paper performs scheduling of tasks by keeping in mind not only minimizing the total cost of transferring and executing, but also minimizing the total transferring time and the time of execution. In this paper the problem of task scheduling described as the assigning all the data of task to all resources in cloud computing environment makes the total cost and time of processing and communication to minimize. Here task scheduling can be regarded as the mapping which maps all the tasks to a Directed Acyclic Graph (DAG) G (V, E). Each node has a weight which mainly denotes the amount of data to be performed by a task on a special node. Each edge has a weight which denotes the amount of information from one task generating to another task to be dealt with. This algorithm objective is not the one but includes processing and transferring time, processing and transferring cost as well. Experiment result shows that this algorithm is efficient in decreasing all types of cost and time in cloud computing environment.

E. *Bi-Criteria Priority based Particle Swarm Optimization Workflow Scheduling Algorithm for Cloud (BPSO) [15]*

This paper develops an approach for scheduling workflow tasks over the available resources of cloud that minimized the execution time and execution cost under the given deadline and budget constraint. In, BPSO the workflow tasks are executed in order of their priority which is basically computed using bottom level which is same as that is defined in HEFT [20]. The assigned priority is then used to initialize the PSO. After assigning the priority the

tasks are sorted according to the descending order of bottom level. The tasks are then sending to different processors according to their order of execution for completing the workflow application. For performing the experiment author developed a simulation program in java which consists of a data center which includes six resources with different processing speed. Experiment results shows that this algorithm has a promising performance when compared with PSO.

F. *Independent tasks scheduling based on Genetic Algorithm [16]*

The paper author introduces GA in scheduling tasks to adapt to memory constraints and request of high performance in cloud environment. The author made various assumptions about tasks:

1. Tasks are a periodic.

2. Tasks are non-preemptive.

3. Each task has two ways of access of a process unit **i.e.** shared and exclusive.

This paper also assumes the centralized scheduling scheme the algorithm that had been developed in this paper synthetically considers both recourse and time utilization so the result that is obtained consists high satisfaction

G. *Genetic Simulated Annealing Algorithm for Tasks Scheduling [17]*

The paper considered the value of both Genetic Algorithm and Simulated Annealing and develops an algorithm. The algorithm considered the QOS requirements of different type of tasks, corresponding to the nature of tasks of user in cloud. QOS is nothing but a standard of user's satisfaction towards the services. This paper says that as the parameters' dimensions are different and orders of magnitude are very different so, in order to evaluate the implementation of tasks in cloud resources by these parameters, the various parameters must be dealt with

dimensionless. According to type of the tasks, different weight of parameters can be given to find resources that satisfy the QOS of the tasks and their expectations. Algorithm steps first execute steps of GA after then annealing comes which helps improving local search ability of GA. This algorithm efficiently completes the searching of resources and allocation process in cloud computing.

WHAT WE ARE DOING?

In this project it will be working on task Web API . Implementing each of the work on OPENSTACK, which is an open source or EC2 AWS (Elastic Cloud Computing of Amazon Web Service) which will be implemented on real time

- As the input , it will use online trace as an input(like text , audio , video)

- And after using various algorithm as PSO , DBD CTO and many more which will be implement one by one and would take the best outcome result of it .

- And will be computing with algorithm and check the result of it and note which algorithm giving the best outcome of its efficiency like makespan ,minimum time of execution (MapReduce), Energy, performance and speed.

HOW IT WILL BE DONE?

In this project, a person or a owner of a file will be running an algorithm (any algorithm) in an attribute based encryption and will be stored data in cloud storage and will be shown an uploading status (efficiency) of files by uploading time and will show the result in graph based result. And when that person willing to download the file it will request for an access and will have to provide a key that key will be uploaded and the user will decrypted and will be able to access the file.

Conclusion

We have discussed the existing methods to give minimum time taken to upload any kind of files in any format (.dll,.mp3,.mp4,.mkv,etc). with use of virtual machines and help of filezilla a channel to connect with cloud which makes easy for other users to work with those services of vm.

Reference

1. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, Cloud Computing and merging IT Platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems, 25:599-616, 2009.

2. M. R. Garey and D. S. Johnson, Computer and intractability: a guide to the theory of NP- Completeness. New York: W. H. Freeman, 1979.

3. A. Zomaya, C. Ward, and B. Macey, "Genetic scheduling for parallel processor systems: comparative studies and performance issues," Parallel and Distributed Systems, IEEE Transactions on, vol. 10, no. 8, pp. 795-812, aug 1999.

4. H. Cheng, "A high efficient task scheduling algorithm based on heterogeneous multi-core processor," in Database Technology and Applications (DBTA), 2010 2nd International Workshop on, Nov 2010, pp. 1–4.

5. A. Amini, T. Y. Wah, M. Saybani, and S. Yazdi, "A study of density-grid based clustering algorithms on data streams," in Fuzzy Systems and Knowledge Discovery (FSKD), 2011, Eighth International Conference on, vol. 3, july 2011, pp. 1652–1656.

6. A. Ebaid, R. Ammar, S. Rajasekaran, and T. Fergany, "Task clustering amp; scheduling with duplication using recursive critical path approach (rcpa)," in Signal Processing and Information Technology (ISSPIT), 2010 IEEE International Symposium on, dec 2010, pp. 34–41.

7. T. Chen, B. Zhang, X. Hao, and Y. Dai, "Task scheduling in grid based on particle swarm optimization," in Parallel and Distributed Computing, 2006. ISPDC '06. The Fifth   International Symposium on, july 2006, pp. 238–245.

8. M. Kashani and M. Jahanshahi, "Using simulated annealing for task scheduling in distributed systems," in Computational Intelligence, Modelling and Simulation, 2009. CSSim '09.   International Conference on, sept. 2009, pp. 265–269.

9. S. Song, K. Hwang, and Y.-K. Kwok, "Risk-resilient heuristics and genetic algorithms for security-assured grid job scheduling," Computers, IEEE Transactions on, vol. 55, no. 6, pp. 703–719, june 2006.

10. K. Li, G. Xu, G. Zhao, Y. Dong, and D. Wang, "Cloud task scheduling based on load balancing ant colony optimization," in Chinagrid Conference (ChinaGrid), 2011 Sixth Annual, aug. 2011, pp. 3–9.

11. A. Verma, S. Kaushal, "Deadline constraint heuristic based genetic algorithm for workflow scheduling in cloud," Forthcoming article in international journal of grid and utility computing.

12. S. Selvarani, G.S. Sadhasivam, "Improved cost-based algorithm for task scheduling in cloud computing,"computational intelligence and computing research, pp.1-5,2010.

13. S. Pandey, L. Wu, S. Mayura Guru, R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," 24th IEEE international conference on advanced information networking and applications, PP 400-407, 2010.

14. Lizheng Guo, Guojin Shao, Shuguang Zhao, "Multi-Objective Task Assignment in Cloud Computing by Particle Swarm Optimization", 978-1-61284-683-5/12/$31.00 2012 IEEE.

15. A. Verma, S. Kaushal, "Bi-Criteria Priority based Particle Swarm Optimization Workflow Scheduling Algorithm for Cloud" Proceedings of 2014 RAECS UIET Panjab University Chandigarh, 06-08 March, 2014.

16. Chenhong Zhao, Shanshan Zhang, Qingfeng Liu, Jian Xie, Jicheng Hu, "Independent tasks scheduling based on Genetic Algorithm in Cloud Computing", 978-1-4244-36934/09/$25.00 2009 IEEE.

17. GAN Guo-ning, HUANG Ting-lei, GAO Shuai, "Genetic Simulated Annealing Algorithm for task scheduling based on Cloud Computing Environment", 978-1-4244-6837- 9/10/$26.00 2010 IEEE.