IJCSN International Journal of Computer Science and Network, Volume 4, Issue 4, August 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

Proposal of Max-Min Algorithm for Scheduling of Workflows in Cloud Environment

[1] Sandeep Singh Brar, [2] Sanjeev Rao

[1] CSE Department, Chandigarh University, Gharuan
Mohali, Punjab, India

[2] CSE Department, Chandigarh University, Gharuan
Mohali, Punjab, India

Abstract - Cloud Computing gave a new direction for the betterment of IT industry. It provides services over Internet according to pay per services use The advent of Cloud Computing as a new model of service provisioning in distributed systems, encourages researchers to investigate its benefits and drawbacks in executing scientific applications such as workflows like Montage, Sipht and Cyber Shake. Most of the algorithms that are currently in use, like First Come First Serve, Round Robin etc., are ignoring the consideration of dependent and independent tasks that directly influence the overall execution time. We propose an approach based on Max-Min algorithm that will consider dependent and independent tasks separately and process the independent tasks simultaneously. It directly gives profit in minimizing computation time.

Keywords - *Cloud Computing, Cloud broker, Scheduling Algorithms, Workflow, Scheduling.*

1. Introduction

Internet is one of the most influential technology, that provides a platform for organizations and individuals to achieve their goal in a better way. New concept of Cloud Computing further exploits features of internet and make services and information available conveniently regardless of location and time. Cloud Computing is a term that provides services over the Internet and uses central remote servers to organize data. Blend of Cloud Computing and Internet is a boon to small scale as well as to large scale industry. The main target behind Cloud Computing is to reduce the cost of infrastructure, maintenance and users can access services what they want and pay for the services that they use. In Cloud, instead of storing data and running applications at individual desktop one can perform all these operations on Cloud. Anyone can access all the services and data any time regardless of location using internet. Citing to the past, in 1969, a leading scientist of the Advanced Research Projects Agency Network project, Leonard Kleinrock, who sowed the Internet, talked about the term utility computing [1]:"As of now, computer networks are still in their infancy, but as they grow up and become sophisticated, we will probably see the spread of, 'computer utilities', which, like present electric and telephone utilities, will service individual homes and offices across the country". Concluding above stated quote that reveals utility computing is dependent on a service purveying model that means services are available on demand and there is sharp alteration for the computing world at this time.

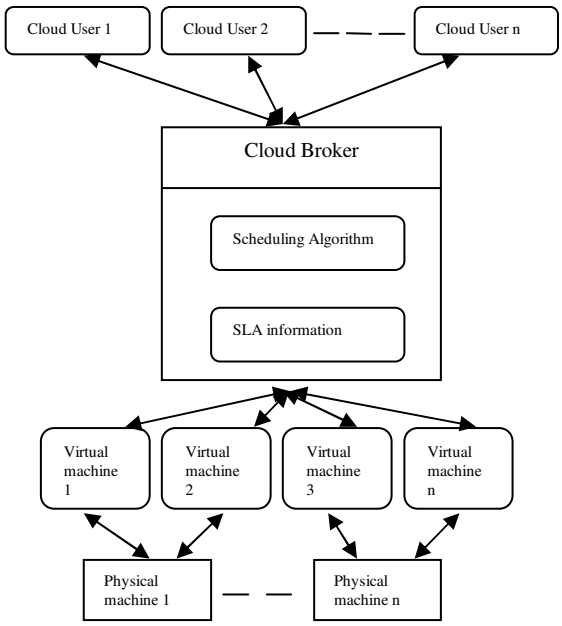

Fig. 1. Role of broker in Scheduling

When a computer is running more than one processes that competes with one another for the CPU resources simultaneously, then there is a need to make a decision by Operating System that which process to run next. The process of taking decision that which process will get resources is called Scheduling. The element of Operating System which makes a choice about the next process is called Scheduler. The algorithm that is used for performing scheduling is regarded as Scheduling algorithm.

As a vital part of Cloud Computing, task scheduling is a mechanism that helps the clients in selecting the suitable resources that can execute their tasks. Cloud Computing has many peculiar features like virtualization and flexibility. With the help of technology of virtualization, all the resources that are physically available can be made virtualized and transparent for users. Additionally, more than one virtual machine are able to operate on a single host computer so that the rate of employment of resources has been effectively enhanced. Supplying resources under Cloud Computing environment is flexible; we increase or reduce the supplying of resources based on users demand.
The primarily objective for scheduling algorithms is either to minimize execution cost or to minimize the data transmission cost or to minimize latency and many more.

2. Related Work

Rajkumar Buyya et al.[1] introduces the simulation toolkit named as CloudSim that helps in simulating and modeling of environment of Cloud. With the help of this toolkit, desired number of virtual machines, datacenters etc. can be created. It enables the user to understand the concept of federation and migration of virtual machines for reliability of applications.

K. Agrawal et al. [2] presented the concept of scheduling and mapping of workflows in Cloud. Author arise the trouble of scheduling that comes in case of linear workflows if mapping is given. The scheduling algorithm can be used to lessen the retro or potential or both. Two models are used in arrangement, one is one-port model and another is multi-port model. In one port model all processes are sequential and restricted communication capacities and computation overlap comes in multi-port model.

Chen et al. [3] presents a Particle swarm optimization (PSO) based approach that can be applied on the scheduling of projects where resources are constrained. The major issue in these projects is long duration of projects that is represented by the multidimensional particle. As a solution, two representations are offered that are priority-based representation and permutation-based representation. Then, frameworks are developed on the basis of two solution representations. In this paper, the outputs of these two representations are compared and a new option of developing a new mechanism that is useful for searching best particle from swarm is offered.

Huang Q.Y. et al. [4] explains the importance of Cloud Computing that is supported by the virtualization technology. This concept is advertized by the business section rather than academic one due to its concentration on users' applications. It arises the point that different clients have varied QoS prerequisites. This paper reviewed different scheduling models and compared them on the basis of given deadline and budget.

Baomin Xu et al. [5] proposed an algorithm that is purely based on Berger model. While implementing this algorithm to the scheduling process, it demonstrates the concept of fairness constraint. The basic constraint is to classify the tasks defined by users on the basis of QoS preferences and then develop a general expectation function fitting with classification of tasks.

V.Krishna Reddy et al. [6] gives the overview of emerging technique that has an enormous impact on the Information Technology (IT) within some past years. The author finds out the varied issues related with cloud computing and also search out interesting future research options. The different challenges of cloud computing are divided in four types such as security challenges, data challenges, performance challenges and design challenges.

Li Jian-feng et al. [7] described about the large number of users, tasks and data involved in cloud computing that arise a very important question that is how to schedule these tasks. A new model Double-Fitness Genetic Algorithm (DFGA) is proposed for programming framework of cloud computing environment. This algorithm is able to not only shorten total-task-completion time but also has shorter average-completion time. The result of DFGA is better than Adaptive Genetic Algorithm (AGA) when these two are compared through simulation experiment.

Jing Liu et al. [8] proposes a new task scheduling model that is mainly focused on cut down the power consumption of Cloud Computing and amending the benefits of service providers. This algorithm is basically designed with the help of multi-objective genetic algorithm (MO-GA). This research consists of encoding

rules, crossover operators, selection operators and the method of sorting Pareto solutions. With the help of CloudSim, this algorithm is implemented and compared with existing algorithms and it is successful in yielding better result for balancing the performance of multiple objects.

Tarun Goyal et al. [9] discusses the new paradigm of Cloud computing that helps the clients in employing the computation, storage, data and services throughout the globe in a commercial way. Among the most important issues of cloud computing, scheduling is still a striking research area. In this paper, a new scheduling model is offered that is centralized on the concept of Minimum Network Delay. In this model, Suffrage Heuristic is conjugate with genetic algorithms to schedule the independent tasks. This work is successful in reducing the make span.

Swachil Patel et al. [10] explained job scheduling priority as the biggest issue because some jobs need to schedule first than the other jobs that can wait for a long time. In this paper, a systematic review of various priority based job scheduling algorithms is presented. These algorithms have different perspectives and working principles.

Rohit O. Gupta et al. [11] discusses the different kinds of scheduling algorithms that have been implemented for executing workflows. These algorithms are compared with each other on the basis of tools, scheduling factors and so on. This paper concludes that an algorithm is still demanded that is related with improvement of reliability and availability in Cloud environment.

B. Anuradha et al. [12] described the demand of scheduling in cloud computing environment due to the heterogeneous pool of resources. The scheduling is simply executing large number of jobs under given constraint. The concept of fairness in scheduling is mandatory criterion which improves the efficiency and provides optimal resource allocation.

Jia ru et al. [13] introduced a new and effective scheduling algorithm that is focused to maximize the cloud resource utilization, improve the computation ratio, reduce make span and overheads and delay in cloud based systems. The author reviewed all existing scheduling algorithms that have been implemented in Cloud Computing environment.

3. Proposed Approach

In our proposed work, Max-Min algorithm is proposed for optimizing workflow scheduling. In this approach, different types of workflows are assigned as a input. Then, apply different scheduling algorithms and execute. The results are obtained and analyzed on the basis of computation time.

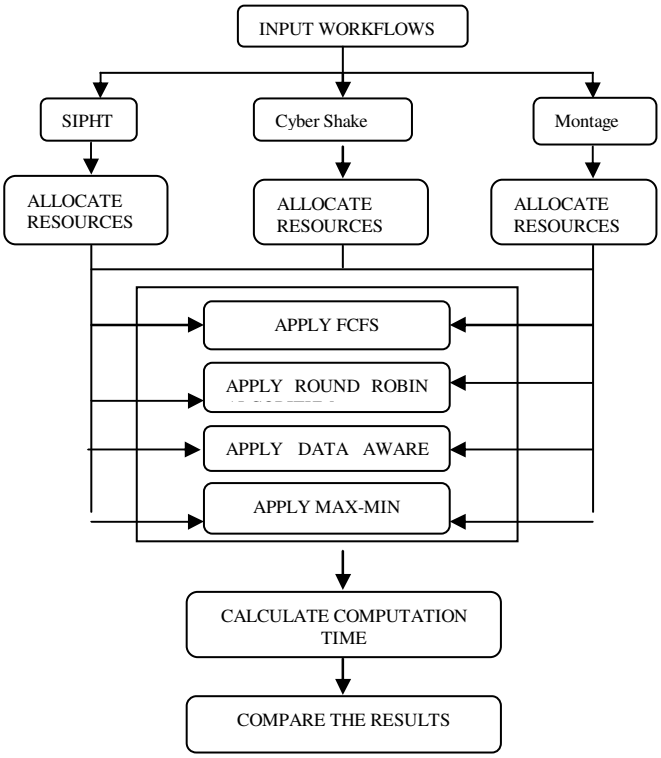

Fig. 2. Proposed Methodology

The proposed approach basically consists of :

1. **Input:** The Four wide used scientific workflows are involved as input: Montage, Cyber Shake, and SIPHT.

- Montage is an astronomy application that is used to build huge copy mosaics of the sky.
- Cyber Shake is a geophysical science application that computes Probabilistic unstable Hazard curves for geographic sites in the Southern California region.
- SIPHT workflow organize a good seek for minor untranslated RNAs (sRNAs) that controls numerous procedures like secretion or virulence in bacteria.

2. **Allocate Resources**: To all these workflows, basic resources are allocated .These resources include virtual machines, broker, cloudlets and datacenters.

3. **Apply Algorithms**

- First Come First Serve Algorithm: It is one of the most common algorithms that is employed in scheduling resources for different workflows. This algorithm works on a very simple principle that the request that comes first will be executed primarily. It does not consider the size of request neither account the dependency and independency of tasks that are requested.
- Round-Robin Algorithm: It is second most commonly used algorithm for assigning the resources, while scheduling, to the different requests submitted for execution. It is also implemented for utilizing Cloud Computing resources for different workflows. In this algorithm, the basic principle is that assign some predefined time slots to the submitted requests. It does not consider dependency and independency that exists between different tasks.
- Data-Aware Algorithm: It is well known algorithm for scheduling of Cloud Computing resources while executing different workflows. This algorithm works on a simple principle that can understand the size of submitted task and works like FCFS. It firstly executes the request that is firstly submitted to it.
- Max- Min Algorithm: It is one of the commonly known algorithms for artificial intelligence. Max–min heuristic sets the priority to the task that requires the longest execution time rather than the shortest execution time. In each iterative step, choosing longest executed process and expected to complete the task at the earliest time. Max–min attempts to minimize the total workflow execution time by assigning longer tasks to comparatively best resources. Therefore best resources provide to big resources and independent resources execute parallel.

1. Proposed Algorithm 1.

Input : Workflow of independent and dependent tasks
Output : Processed Workflows with Computation Time

```
For I=0 to Length (workflow tasks)
  Begin
    Separate tasks from according to flow of graph
    Provide resources (virtual machine) to every task.
  End
    Apply Max –Min algorithm on Workflow
For I=0 to Length (workflow tasks)
  Begin
    Separate tasks from according to flow of graph
    Provide resources according to Max-Min.
  End
Analyze cost and computation time
```

The proposed algorithm consists of another algorithm i.e. Max- Min algorithm. This algorithm works on a simple principle that schedules the largest tasks first and parallel execute small tasks that are independent w.r.t largest task.

2. Proposed Max- Min Algorithm

Input: Work flow graph
Output : Schedule list of tasks

```
While (Workflow task is not completed)
Begin
    Available task = task whose parent task is scheduled
    Schedule(available task)
End
    Procedure schedule(available task)
While available task not scheduled
Begin
        For all task find the resources
    Begin
        For all task computes ECT( t, resources )
    End
End
Put the priority on big ECT( t, resources )
Delete task from available list
```

Formula Used:ECT(t,r)=EET(t,r)+Max(EAT(t,r),FAT(t,r))
Estimated Execution Time (EET): execution time of t on r resources
Estimate Availability Time (EAT): resource available on task t
File Available Time(FAT): the earliest time by which all the files required by the task t will be available at the resource r

References

[1] Rajkumar Buyya, Rajiv Ranjan, Rodrigo N. Calheiros, "Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities," The International Conference on High Performance Computing and Simulation, HPCS2009, pp:1-11, Year 2009.

[2] K. Agrawal, A. Benoit, L. Magnan and Y. Robert, "Scheduling Algorithms for Linear Workflow optimization," IEEE, Year 2010.

[3] Chen, R. M., Wu, C. L., Wang, C. M. and Lo, S. T., "Particle swarm optimization scheme to solve resource-constrained scheduling problem," Expert systems with applications, Vol. 37, pp. 1899-1910, ISSN: 0957-4174 March, Year 2010.

[4] Huang Q.Y., Huang T.L. , "An Optimistic Job Scheduling Strategy based on QoS for Cloud Computing

," IEEE International Conference on Intelligent Computing and Integrated Systems, Guilin, pp. 673-675, Year 2010.

[5] Baomin Xu, Chunyan Zhao, Enzhao Hu, Bin Hu, "Job Scheduling algorithm using Berger model in Cloud Environment," Elsevier in Advances in Engineering Software, Vol. 42 , Issue No. 7, pp. 419-425, Year 2011.

[6] V.Krishna Reddy, B. Thirumala Rao , LSS Reddy, "Research issues in Cloud Computing," Global Journal Computer Science & Technology, Vol. 11, pp.70-76, June, Year 2011.

[7] Li Jian Feng, Peng Jian, "Task scheduling algorithm based on improved genetic algorithm in cloud computing environment," Journal of Computer Applications, pp 184-186 , Year 2011.

[8] Jing Liu , Xing-Guo Luo, Xing-Ming Zhang3, Fan Zhang and Bai-Nan Li, "Job Scheduling Model for Cloud Computing Based on Multi- Objective Genetic Algorithm," International Journal of Computer Science Issues, Vol. 10, Issue 1, ISSN : 1694-0784, January, Year 2013.

[9] Tarun goyal , Aakanksha Agrawal, "Host scheduling algorithm using genetic algorithm in cloud computing environment," International Journal of Research in Engineering and Technology, Vol. 1, Issue 1, pp. 7-12, June , Year 2013.

[10] Swachil Patel, Upendra Bhoi, "Priority Based Job Scheduling Techniques In Cloud Computing," International Journal of Scientific & Technology Research , Vol. 2, Issue 11, ISSN : 2277-861, November, Year 2013.

[11] Rohit O. Gupta, Tushar Champaneria, "A Survey of Proposed Job Scheduling Algorithms in Cloud Computing Environment," International Journal of Scientific & Technology Research, Volume 3, Issue 11, November, Year 2013.

[12] B. Anuradha, S. Rajasulochana, "Fairness As Justice Evaluator In Scheduling Cloud Resources: A Survey," International Journal of Computer Engineering & Science, ISSN: 22316590 , November, Year2013.

[13] Jia Ru , "An Investigation on scheduling policies for cloud based software-systems".

.