

# A Survey on Scheduling and the Attributes of Task Scheduling in the Cloud

Nallakumar. R<sup>1</sup>, Dr. N. Sengottaiyan<sup>2</sup>, Sruthi Priya K.S<sup>3</sup>

Assistant Professor, Department of CSE, Anna University Regional Centre, Coimbatore, India<sup>1</sup>

Professor, Department of CSE, Indira Institute of Engineering and Technology, Thiruvallur, Chennai, India<sup>2</sup>

PG Scholar, Department of CSE, Anna University Regional Centre, Coimbatore, India<sup>3</sup>

**Abstract:** Cloud Computing is the latest new trend that is growing rapidly in this era of technological advancements. The word “Cloud” is a conglomeration of computers and servers accessed via the Internet and this makes everyone in this world to access their documents and applications from anywhere and at any time. Scheduling is blooming as a vital research area in the field of Cloud Computing. It aims to map the appropriate tasks, jobs to the resources that are available by keeping in mind the QoS (Quality of Service) requirements of the user. The on-demand service in the Cloud is one of the guiding factor to discover new scheduling strategies. This paper surveys on Scheduling, the types of job scheduling algorithms, types of task Scheduling, and the attributes of Task scheduling.

**Keywords:** Scheduling, Scheduling procedure, Job Scheduling, Task Scheduling, Attributes of Task Scheduling.

## INTRODUCTION

Cloud Computing is emerging with a great popularity in recent years due to its unique trademark which bestows the users with high scalability, reliability, security, cost-effective mechanism, group collaboration and easy access to various applications. Many attributes of Cloud are actually part of Utility and Grid computing. Cloud services propound on the basis of “Pay for what you use” via public or private network. Software As a Service (SaaS), Platform As a Service (PaaS), Infrastructure As a Service (IaaS) marks the services of the Cloud.

Cloud serves in a wide number of users across different platforms. Any user who is acknowledged has the full rights to access their documents and applications from any computer situated in any place provided they have access to constant connection to the Internet. The technology or Infrastructure behind Cloud Computing is always not shown explicitly to the users. Private, Public, Hybrid and Community cloud are the various deployment models present in it. With the concept of Virtualization in Cloud, flexibility is highly achievable. Security and Scheduling are the two sub domains in the cloud and are the major research issues in recent times. There are many applications of Cloud among which “Google Cloud” is the one. The various characteristics of Cloud Computing from Google’s point of view [1] is as follows.

1. User-centric
2. Task-centric
3. Powerful
4. Accessible
5. Intelligent
6. Programmable.

Cloud comprises of Cloud provider, Cloud Consumer, Cloud Broker, Cloud auditor and cloud carrier as its main entities [2] whose interrelationship is depicted in figure 1 and explanation is given below the figure.

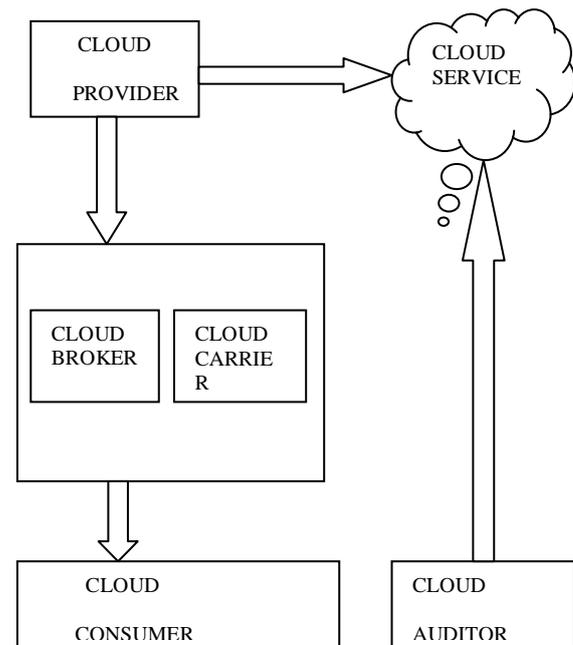


Fig.1. Main entities in Cloud computing with their Interrelationship

**Cloud Provider:** Offers various services to the cloud consumers who are the actual cloud users and it also manages the infrastructure to deliver the services to the consumer.

**Cloud Consumer:** Makes use of the cloud service from the cloud provider with the help of a cloud broker who act as an intermediary.

**Cloud Broker:** Guides the cloud provider and consumer in all the activities.

**Cloud Auditor:** Examines the cloud service based on various factors such as Service level agreement (SLA), reliability, availability, performance.

**Cloud Carrier:** Maintains the transport channel and connectivity between cloud provider and cloud consumer.

Scheduling in cloud computing involves a number of tasks or jobs to be executed with the resources that are present to achieve high performance, minimum time for response and better allocation as well as utilization of the resources [3]. The cloud makes use of a Virtual Machine for job and task allocation and there are many issues related to proper allocation and utilization of virtual resources using scheduling. Execution time and execution cost leads to a better scheduling strategy or algorithm.

The rest of the paper is structured as follows. Scheduling, its essential need and routine are described in section II followed by the Scheduling types in section III. Section IV explains the types of Job scheduling algorithms followed by the types of Task scheduling in section V. The attributes of Task scheduling are forecasted in section VI and section VII concludes the paper.

## II SCHEDULING

Scheduling is the process of finding out the appropriate resources available to allocate to the tasks or jobs. Scheduling is done effectively by taking into mind certain constraints such as budget constraints, deadline constraints, Quality of service of the user (QoS) and high throughput. Two main types of scheduling are job and task scheduling.

### A. Essential need of Scheduling

It is only with the help of various scheduling algorithms we can effectively manage the utilization of the resources with the emphasis on load balance. To schedule the tasks or jobs with various criteria, scheduling plays a key role. Another important fact is that only with the scheduling strategies we can schedule the appropriate Virtual Machines (VM) after the selection of the resources and also based on the task and VM utilization time. The performance level of Cloud Computing is very high which makes it a suitable platform for scheduling various tasks and jobs. To add on, Cloud offers a good environment for scheduling the application [4] when compared to other computing domains such as grid.

### B. Scheduling routine

The general scheduling routine [5] that is followed in all scheduling techniques is as follows

1 Resource spotting: The resources are identified based on user needs and are refined to find out the best.

2. Resource Preference: This is one of the most important stage where the resources are chosen on the basis of various task and resource parameters.

3. Task submission: The tasks are then submitted to the preferred resource.

## III SCHEDULING TYPES

The Figure 2 clearly depicts the two main types of scheduling namely Job and task scheduling and some of the major categories under the job and task scheduling.

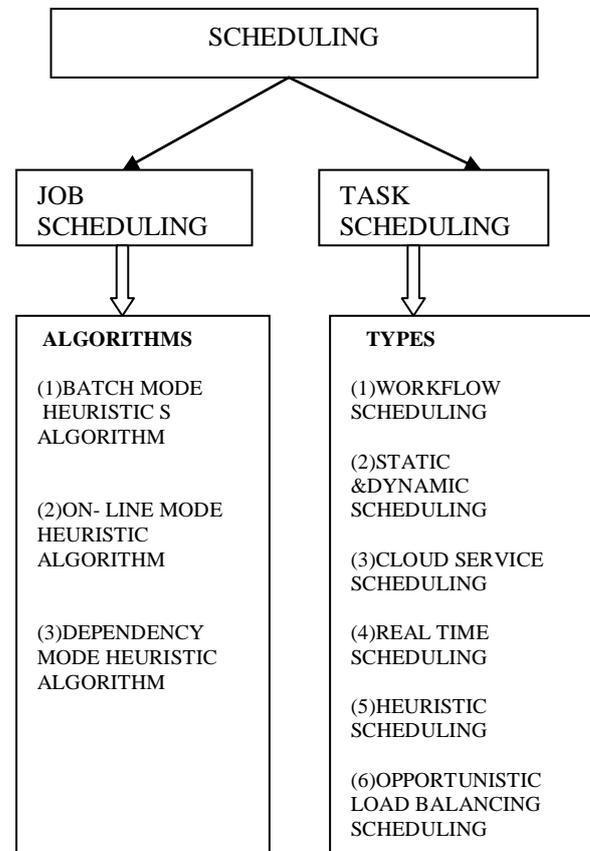


Fig.2.Scheduling Types

### A. Job Scheduling

Initially job scheduling was not carried out in the cloud environment, but the scalability of cloud computing lead to the movement of job scheduling technique from other environments into cloud computing. Job Scheduler plays the major role in this type of scheduling. It is the role of a job scheduler to find out the best optimized way for resource allocation to the jobs of the user to be executed [5]. Scheduling of jobs will be done by considering many constraints such as satisfying the QoS (quality of the Service) with maximum efficiency and cost minimization. Three main classifications of job scheduling algorithms are Batch mode heuristic algorithms, On-line Mode heuristic algorithms and Dependency mode heuristic algorithms.

### B. Task scheduling

Task scheduling technique sets out an example for NP-complete problem and it is one of the most commonly used

scheduling type in Cloud Computing. It aims to map the tasks of the user to the resources that are available. Virtualization concept that is present in Cloud is the backbone for efficient task scheduling. One of the drawbacks with this type of scheduling is that its efficiency will have an impact on the cloud environment's performance. There are many categories of task scheduling out of which Workflow scheduling, Static & dynamic scheduling, Cloud service scheduling, Real time scheduling, Heuristic scheduling and Opportunistic load balancing scheduling (OLB) are some of them.

#### IV JOB SCHEDULING ALGORITHM TYPES

The following paragraphs give a brief discussion on the most common types of Job Scheduling algorithms.

##### A. Batch Mode Heuristics Algorithm

In this type of job scheduling, queuing of jobs is the first step and then allocated into a set as they arrive into the system. Jobs are not scheduled immediately, but only after some time. Some of the most prominent examples of the Batch mode heuristic algorithm [6] and their features are tabulated in the table 1.

TABLE 1  
EXAMPLES OF BATCH MODE HEURISTICS ALGORITHM

Type	Scheduling procedure	Feature
First Come First Serve algorithm	Jobs are placed in a queue as soon as they arrive.	Straightforward and quick.
Max-Min algorithm	Bigger tasks are chosen first to execute.	Jobs that are large get executed soon than the tasks that are smaller, which will have to wait until all the bigger tasks gets executed.
Round Robin algorithm	All processes are allotted CPU time known as time-slice & it should be executed before time-slice expires.	Best suited for parallel computing due its good load balancing feature.
Random algorithm	Jobs that are randomly selected is assigned for execution into the virtual machines (VM).	It does not lead to heavy or low load as the VM position is not considered.

##### B. On-line Mode Heuristic Algorithm

Jobs are immediately scheduled once they arrive into the system. Most-fit task scheduling algorithm is the best example [7] for this type of scheduling algorithm. In Most-fit task scheduling algorithms (MFTF), the jobs are selected based on some parameters and are best fit into the queue which are the ones that are actually executed first.

##### C. Dependency Mode Heuristic Algorithm

It aims to map the tasks that are dependent on the resources which are homogeneous in a way of finding out the links that exist on the task graph [8]. The main goal is to complete the execution of the interdependent tasks as early as possible. The tasks that are assigned priority are ranked. The homogenous system and heterogeneous system have three categories of dependent task scheduling heuristics such as list heuristics, clustering heuristics, task duplication.

#### V TASK SCHEDULING TYPES

Some of the types of Task Scheduling are classified below with a short explanation on them.

##### A. Workflow Scheduling

Workflows generally consist of number of tasks that are interconnected in the form of a Directed Acyclic Graph (DAG). In a DAG, every node denotes the tasks and edges signify the dependencies among those tasks [9]. All tasks will generally communicate with all the other task present in the DAG. Workflow scheduling deals with the process of mapping these tasks to the correct resources and that workflow scheduling has to be done by considering certain parameters such as deadline, minimum budget and cost, Quality of service (QoS). Workflow management deals with the management of Workflow Scheduling. Efficient scheduling of workflows leads to high performance of the system.

##### B. Static and Dynamic Scheduling

Static Scheduling involves two steps, namely data pre-fetching and data pipelining to various stages of task execution. Having less overtime overload is one of the best feature of static scheduling. The execution time of the tasks can't be guessed earlier itself in dynamic scheduling due to the dynamic behaviour of the tasks. Scalable scheduling methods should be avoided in the Cloud environment in order to avoid greater overheads as described in [10].

##### C. Cloud Service Scheduling

User level and System level scheduling are the two types of Cloud Service Scheduling [11]. The problem of resolving the issues that exist between the cloud provider and users is dealt at user level scheduling. The problem of managing the resources present at the data centre is dealt at system level scheduling. The number of tasks that are obtained from huge number of users will be assigned to the appropriate physical machine that is located at the data centre. Certain requirements that are to be taken into account along with the task assigned to the physical machines is Service level Agreement (SLA), fault tolerance, reliability, Quality of Service (QoS).

##### D. Real Time Scheduling

Real Time Scheduling needs to concentrate on the increasing of throughput and minimizing of average response time. Non-preemptive scheduling is applied to schedule the real-time tasks in order to increase the total

utility. Any task is associated with two Time Utility Functions (TUF). One is profit time utility function and the other is penalty time utility function. Preemptive scheduling is also utilized for scheduling the real-time tasks as described in [12].

### **E. Heuristic Scheduling**

Heuristic algorithms are generally priority based and problem centric [13] NP-hard problems are generally solved by using heuristic methods. The general procedure followed in this type of enumeration method is that among all the solutions that are enumerated, the optimal solution is chosen. To find out the best optimal solution heuristic is generally used as a suboptimal algorithm. Workflow applications are generally heuristic or meta-heuristic in nature [14] Heuristics need human interaction as it is the role of the developer to assign priority to the tasks and resources based on his previous experiences. Meta-heuristics provides solutions without the human intervention.

### **F. Opportunistic Load balancing (OLB) Scheduling**

Every task is scheduled in some random manner to the available machine to be used next without the consideration of the tasks anticipated amount of execution time [15]. The main motive behind OLB (opportunistic Load balancing) scheduling is to make all machines to be busy. An important plus of OLB is its simpleness but it lacks in a good makespan as it does not consider the anticipated execution time of the tasks.

## **VI ATTRIBUTES OF TASK SCHEDULING**

Virtualization technology plays the most prominent role in scheduling the tasks in Cloud computing. Cloud providers make use of virtual machines to execute the users tasks "lucidly" and this is due to the flexible and dynamic nature of cloud computing. Task scheduling aims to focus on Quality of service (Qos), load balance, good running time and maximum throughput. The most important features of task scheduling [16] are as follows.

### **A. Task Scheduling provides scalability**

Cloud provider generally has the capability to provide a minimum amount of resources in the initial stage. If the resources are limited then the user demands will not be satisfied. When more computing resources are added, the size of the virtualized resources will expand and this becomes an important fact for the rise of user demands. Task scheduling should possess the capability to expand the amount of resources when it is needed (i.e. Scalability of resources) ,only then the throughput will be high.

### **B. Task scheduling serves to a Uniformed Resource Platform**

A common resource pool consisting of several hosts, computers, workstations is set up with the help of cloud virtualization. The resources are supplied to a wide range of distributed computers with the help of a data centre. A data centre usually has number of physical machines

connected in it. Virtual machines will be created above the physical machines with the virtualization technology.

### **C.Task Scheduling is dynamically adjusting**

According to the user requirements, the applications in the cloud environment may enlarge or shrink. This holds good even for Virtualized computing resources. The nature of the resources has been to keep changing often. When demands are very high there will be a need for the addition of the new resources and there will be failure of resources at times.

### **D. Every single Node in the cloud is self-reliant**

In the cloud there will be several nodes connected and the scheduling of all these nodes is separate and independent. The main role of the scheduler is to not interfere with the scheduling routine of all those nodes.

### **E. Task scheduling is of two parts**

First part: Used as Resource pool scheduling. The primary goal is to schedule the cloud applications and application programming interface.

Second part: Used for port resource scheduling. Examples for this includes Map reduce task scheduling.

"Scheduling and call back resource" are the two-way process in both parts whose short explanation follows below

Scheduling: To satisfy the requirements of the user, The scheduler will lease the resources for a particular time from the cloud provider.

Callback resource strategy: When the resources are used effectively the scheduler callback them.

## **CONCLUSION**

Scheduling is one of the most vital research issues in the Cloud computing environment. This paper surveys the two main types of scheduling (i.e. Job and Task Scheduling), some of the categories of Job Scheduling algorithms, types of Task Scheduling as well as the main attributes of task scheduling. At present Task Scheduling is on the main focus and more research is being carried out on it. Several algorithms and strategies already exist for Task Scheduling based on a particular parameter. As a future scope new scheduling strategies to optimize multiple parameters in a way to satisfy the user requirements can be found.

## **REFERENCES**

- [1] <http://googlesystem.blogspot.in/2008/04/Kai-fu-lee-on-cloud-computing.html>
- [2] <http://blog.simplilearn.com/it-service-management/cloud-entities>
- [3] Maheswari. R, S. Selvi "A Survey on Scheduling Algorithms in Cloud Computing", IJERT, Vol. 2, Issue 10, October – 2013.
- [4] Sujit Tilak, Prof. Dipti Patil, " A Survey of Various Scheduling Algorithms in Cloud Environment", IJEI, Vol 1, Issue 2, Sept, 2012.
- [5] Dr Ajay jangra, Tushar Saini, "Scheduling Optimization in Cloud Computing, IJARCSSE, Vol 3, Issue 4, April 2013,
- [6] Rohit O. Gupta, Tushar Champaneria, "A Survey of Proposed Job Scheduling Algorithms in Cloud Computing Environment", Vol 3, Issue 11, November 2013.
- [7] Shamsollah Ghanbari, Mohamed Othman "A Priority based Job Scheduling Algorithm in Cloud Computing", International Conference

- on Advances Science and Contemporary Engineering, ICASCE, 2012.
- [8] YAGOUBI Belabbas, KADRI Walid, "Assignment of DAGs Applications in Grid Computing Environment", PCC Vol. 2, Issue 2.
- [9] Simsy Xavier, S. P. Jenio Lovesum, "A Survey of Various Workflow Scheduling Algorithms in Cloud Environment", IJSRP, Vol 3, Issue 2, February 2013.
- [10] Yogita Chawla and Mansi Bhonsle, "A Study on Scheduling Methods in Cloud Computing", IJETTCS, Volume 1, Issue 3, September – October 2012.
- [11] Fei Teng, "Resource allocation and scheduling models for cloud computing", Paris, 2011.
- [12] Shuo Liu, Gang Quan, Shangping Ren, "On-line preemptive scheduling of real-time services with profit and penalty", IEEE, 2011.
- [13] Lovejit Singh and Sarbjeet Singh, "A Survey of Workflow Scheduling Algorithms and Research Issues", IJCA, Vol 74-No.15, July 2013.
- [14] Jia Yu, Raj Kumar Buyya and Kotagiri Ramamohanarao. "Workflow Scheduling Algorithm for Grid Computing. Meta-heuristics for Scheduling in Distributed Computing Environment", Vol. 146, 2008.
- [15] Gaurang Patel, Rutvik Mehta, "A survey on Various Task scheduling Algorithms in cloud Computing", IJARCCET, Volume 3 Issue 3, March 2014.
- [16] Hong Sun, Shi-ping Chen, Chen Jin, Kai Guo, "Research and Simulation of Task Scheduling Algorithm in Cloud Computing", TELKOMNIKA, Vol.11, No.11, November 2013, pp. 6664-6672

## BIOGRAPHY

**Mr. R. Nallakumar** received the Bachelor Degree in Computer Science and Engineering in 2009. He received the Master degree in Computer Science and Engineering in 2011. He also completed Master of Business Administration and currently he is pursuing his Ph.D. He is working as an Assistant professor at Anna University Regional Centre, Coimbatore, Tamilnadu, India. His area of interest is Cloud Computing.

**Dr. N. Sengottaiyan** received the Bachelor Degree in Electronics and Communication Engineering in 1986. He received the Master Degree in Computer Science and Engineering in 2004. He has received his Ph.D in 2011. He is currently pursuing P.D.F from California South University, California, U.S.A. He is working as a Principal at Indira Institute of Engineering and Technology Thiruvallur, Chennai, Tamilnadu, India. His area of interest are Cloud Computing and Wireless Sensor Networks.

**Ms. K.S. Sruthi Priya** received the Bachelor Degree in Computer Science and Engineering in 2013. She is currently pursuing her Master Degree in Software Engineering from Anna University Regional Centre, Coimbatore, Tamilnadu, India. Her area of interest is Cloud Computing.