A Brief Review of Scheduling Algorithms in Cloud Computing

Mohd Rahul
Research Scholar OPJS University,
Churu, Rajasthan, India
Email: mohd.rahul@gmail.com

Dr. Vaibhav Bansal
Associate Professor - CSE
OPJS University,
Churu, Rajasthan, India

Abstract—**Cloud computing comes in focus development of grid computing, virtualization and web technologies. The cloud computing is a mingle of technologies where a large number of systems are connected in private or public networks. This technology offers dynamically scalable infrastructure for data, file storage, and application. Scheduling is a main task in a cloud computing environment. In cloud computing environment datacenters take care of this task. The selection of a exacting scheduling algorithm depends upon various factors like the parameter to be optimized (cost or time), quality of service to be provided and information available regarding various aspects of job. Workflow applications are the applications which need various sub-tasks to be executed in a particular fashion in order to complete the whole task. Various scheduling algorithms surveyed in this paper. The goal of cloud task scheduling is to achieve high system throughput and to assign various computing resources to applications. The Complexness of scheduling trouble increases with the size of the task and becomes very difficult to solve effectively. Min-Min algorithm is used to lessen the make span of tasks by assuming the task length. Keeping this in mind, cloud providers should achieve user satisfaction**

Keywords: Cloud Computing, Scheduling in cloud computing, Main Entities in cloud computing, Types of Scheduling.

1. INTRODUCTION TO CLOUD COMPUTING:

Cloud computing is an extension of parallel computing, distributed computing and grid computing. It provides safe, quick, convenient data storage and computing power with the help of internet. Virtualization, distribution and dynamic extendibility are the vital characteristics of cloud computing [1]. Cloud Computing enables the procurement of huge quantity of computational resources on require by employing pay-per-use model. It delivers computational resources with the help virtualization technology.

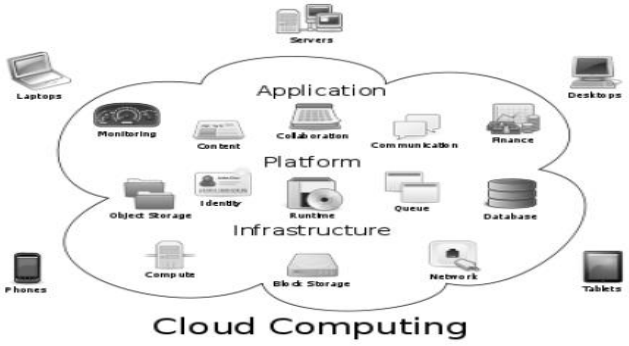

Fig 1 overview of cloud computing [1]

1.2 Main Entities in cloud computing:

In a Cloud, there are four major entities viz. Cloud User, Broker, Virtual Machines and Physical Machines. The cloud users are the real consumers of services and can submit their service requests from anywhere in the world. A cloud data center can be of physical machines. Using virtualization technology, virtual machines are created on the top of physical machines. Broker acts as an intermediator among cloud users and cloud datacenters. It is responsible for allocating cloud resources to user's workflow applications. It allocate virtual machines to user's workflow applications by making use a scheduling algorithm and (Service Level Agreements) SLA which is a written and agreed document between service provider and cloud user. Figure 1 demonstrates the role of a cloud broker in a cloud environment [2].

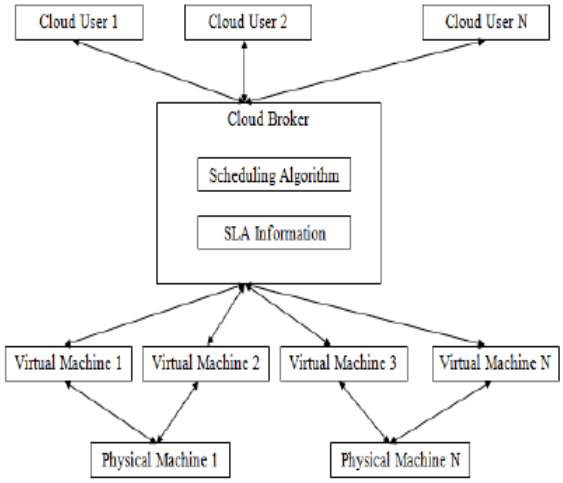

Fig. 2 Role of Broker in Cloud Scheduling [2]

1.3 Scheduling in cloud computing:

Scheduling is the one of the most prominent activities that performs 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 utmost profit. The main objective of the scheduling algorithms in cloud environment is to consume the resources properly while managing the load between the resources so that to get the minimum execution time [3].

1.3.1 Types of Scheduling:

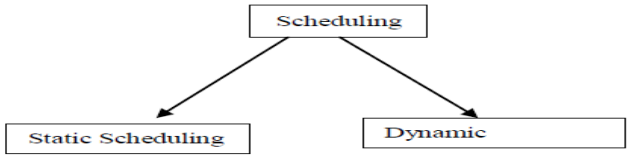

Fig 3: Types of Scheduling [3]

1.3.2 Procedure of Scheduling:

Scheduling process in cloud can be generalized into three phases namely–

Resource discovering and filtering –

Datacenter Broker discovers the resources declare in the network system and collects status information related to them.

Resource selection – Target resource is selected relayed on certain parameters of task and resource. This is deciding stage.

Task submission -Task is submitted to resource selected [4].

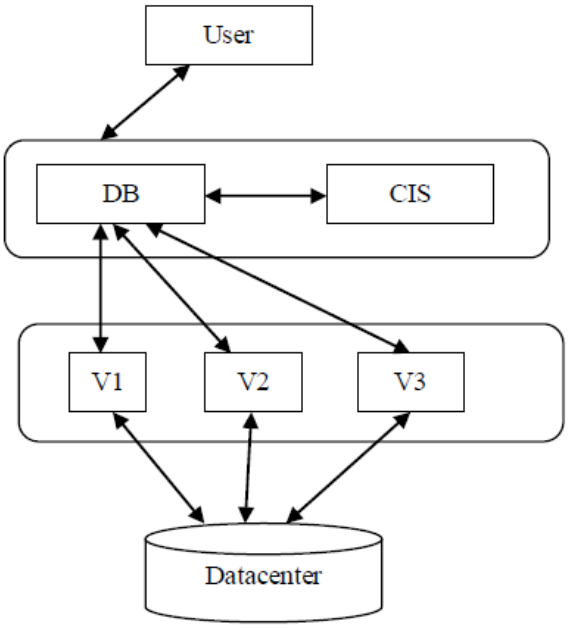

Fig 4: stages of scheduling

DB: - Datacenter Broker
CIS: - Cloud Information services
V1, V2, V3 are the virtual machines [3]

1.3.3 Various scheduling algorithms:

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

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

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

Min-Min Algorithm: Min-Min algorithm choose the smaller tasks to be executed first.

Max-Min algorithm: Max-Min algorithm chooses the bigger tasks to be executed first. i.e, Heuristic algorithms [3].

1.4 The Following scheduling algorithms are currently prevalent in clouds.

1.4.1 A Compromised-Time-Cost Scheduling Algorithm: A novel compromised-time-cost scheduling algorithm which considers the characteristics of cloud computing to accommodate request-intensive cost-constrained workflows by compromising execution time and cost with user input enabled on the fly. The simulation has established that CTC (compromised-time-cost) algorithm can achieve lower cost than others while meeting the user-designated time limit or reduce the mean execution time than others within the user-designated effecting cost. The tool used for simulation is SwinDeW-C (Swinburne Decentralised Workflow for Cloud) [1].

1.4.2 Task Scheduling Algorithm [2]:

1.4.2.1 Task Scheduling in cloud computing: role of task scheduler and terminology associated with task scheduling:

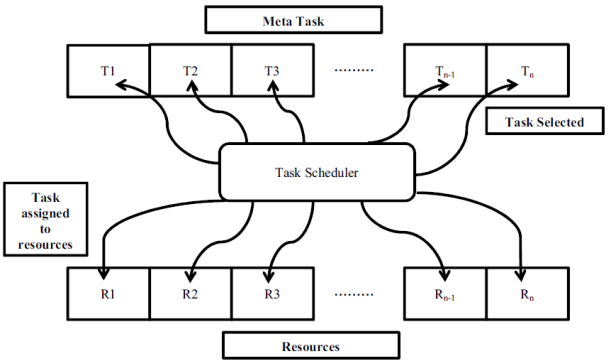

Fig 5.Role of task schedular

There are different task scheduling algorithms:

1. Opportunistic load balancing (OLB)
2. Minimum execution time (MET)
3. Minimum completion time (MCT)
4. Min-Min Task scheduling
5. Max-Min Task scheduling algorithm
6. Load balanced Min-Min algorithm [5]

1.4.3 Workflow Scheduling:

In workflow scheduling, diverse sub tasks of a bigger task are allocated resources in such a way that some pre-defined objective criteria is met. A workflow application is generally represented as a Directed Acyclic Graph (DAG) such as G (V, E) where V is the number of jobs and E is the information regarding data dependencies among tasks. A task which does not have any parent task is named entry task and a task which does not have any child task is called an exit task [5].

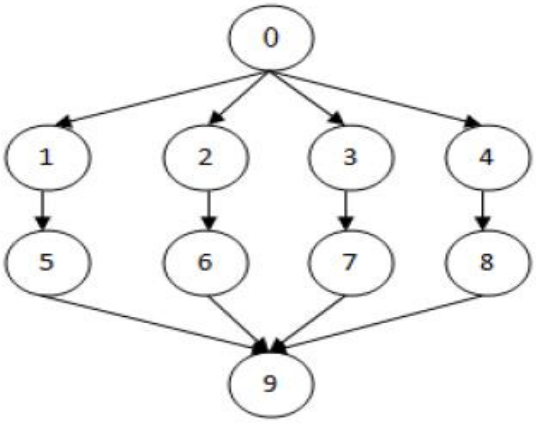

Fig. 6 A Workflow represented in the form of a graph [2]

2. LITERATURE SURVEY

Sujit Tilak et al. (2012) Cloud computing environments provide scalability for applications by supplying virtualized resources dynamically. Cloud computing is built on the support of distributed computing, and virtualization, grid computing. User applications may need huge data retrieval very often and the system competence may degrade when these applications are scheduled taking into account only the 'finishing time'. In addition to optimizing system efficiency, the cost derive from data transfers among resources as well as execution costs must also be taken into account whilst scheduling. Moving applications to a cloud computing atmosphere triggers the need of scheduling as it enables the consumption of various cloud services to facilitate execution**.**

Lovejit Singh et al (2013). Cloud Computing refers to a paradigm whereby services are offered by internet using pay as you go model. Services are deployed in data centers and the pool of data centers is together referred to as "Cloud". Data centers make use of scheduling techniques to optimally allocate resources to diverse jobs. Different scenarios require different scheduling algorithms. The selection of a specific scheduling algorithm depends upon various factors like the parameter to be optimized (cost or time), eminence of service to be provided and information available concerning various aspects of job. Workflow applications are the applications which involve various sub-tasks to be executed in a particular fashion in order to complete the entire task. These tasks have parent child association. The parent task needs to be executed before its child task. Workflow scheduling algorithms are imaginary to preserve dependency constraints implied by their nature and structure. Resources are allocated to diverse sub-tasks of the original task by keeping into account these constraints. In this paper, diverse workflow scheduling algorithms have been studied. Some algorithms have been found to optimize cost, some have been found to optimize time, some stress on reliability, some stress on availability, some focuses on energy efficiency, some focuses on load balancing or some stresses on a combination of these parameters. A lot of work has already been done in the part of workflow scheduling but still, think that there is a require and lot of scope in applying other optimization techniques, akin to intelligent water drops, to schedule workflow applications [7].

Er. Shimpy et al. (2014) Cloud computing is a capable technology for future computing platforms and is estimated to provide easier access to remote computational resources that are typically locally limited. Scheduling with meta-heuristic algorithms is one of the active research area in cloud computing. The aim of cloud task scheduling is to achieve high system throughput and to allocate various computing resources to applications. The hardness of scheduling problem increases with the size of the task and becomes greatly difficult to solve efficiently. Many different techniques have been proposed to solve this problem. Some of the methods are relayed on heuristic techniques that provide an optimal or near optimal solution for tasks i.e, large in size. In this paper author study diverse scheduling algorithms in different environments with their respective parameters.

Dr. Amit Agarwal et al. (2014) Cloud computing is an emerging technology in distributed computing which facilitates pay per model as per user demand and requirement. Cloud consists of a collection of virtual machine which includes both computational and storage facility. The primary plan of cloud computing is to provide efficient access to remote and geographically distributed resources. Cloud is mounting day by day and faces many challenges, one of them is scheduling. Scheduling refers to a set of policies to direct the order of work to be performed by a computer system. A good scheduler adapts its scheduling approach according to the changing environment and the type of task. In this research paper we presented a Generalized Priority algorithm for competent execution of task and comparison with FCFS and Round Robin Scheduling. Algorithm should be veteran in cloud Sim toolkit and result shows that it gives better performance compared to other traditional scheduling algorithm [8].

Navdeep Kaur et al (2015) in this research paper, additional constrains have been considered to evolution a holistic analysis based algorithm based on Max-Min algorithm, which work on principle of sorting jobs(cloudlets) based on conclusion time of cloudlets. The improved algorithms here also reviews the job

characteristics in method of size, outline, payload ratio and available storage blocks in particular cluster of contribution of file systems. The observations show no important overload due to addition of these constrains, as sorting operation remains same and capable. Storage allocation helps in getting better performance [6].

Liyun Zuo et al (2015). For task-scheduling troubles in cloud computing, a multi-objective optimization method is proposed here. First, with an aim toward the biodiversity of resources and tasks in cloud computing, author propose a resource cost model that defines the demand of tasks on resources with more details. This model reflects the association between the user's resource costs and the budget costs. A multi-objective optimization scheduling process has been proposed based on this resource cost model. This method considers the make span and the user's budget costs as constraints of the optimization problem, achieving multi-objective optimization of both performance and cost. An enhanced ant colony algorithm has been proposed to solve this problem. Two constraint functions were used to assess and provide feedback regarding the performance and budget cost. These two restraint functions made the algorithm adjust the quality of the solution in a timely manner based on feedback in order to achieve the finest solution. Some simulation experiments were designed to evaluate this method's performance using four metrics: 1) the make span; 2) cost; 3) deadline violation rate; and 4) resource utilization. Experimental results show that relayed on these four metrics, a multi-objective optimization method is better than other similar methods, particularly as it increased 56.6% in the best case scenario [9].

RASA Workflow scheduling [2]	Batch mode	make span	Grouped tasks	1.It is used to reduce make span	Grid Environment	GridSim
Innovative transaction intensive cost-constraint scheduling algorithm [5]	Batch Mode	Execution cost and time	Workflow with large number of instances	1.To minimize the cost under certain user-designated Deadlines. 2. Enables the compromises of execution cost and time	Cloud Environment	SwinDeW-
SHEFT workflow scheduling algorithm [6]	Dependency Mode	Execution time, scalability	Group of tasks	1. It is used for optimizing workflow execution time. 2. It also enables resources to scale elastically during workflow execution.	Cloud Environment	CloudSim
Multiple QoS Constrained Scheduling Strategy of Multi-Workflows [7]	Batch/dependency mode	Scheduling success rate,cost,time, make span	Multiple Workflows	1. It is used to schedule the workflow dynamically. 2. It is used to minimize the execution time and cost	Cloud Environment	CloudSim

Scheduling Algorithm	Scheduling Method	Scheduling Parameters	Scheduling factors	Findings	Environment	Tools
A compromised-Time-Cost Scheduling Algorithm [3]	Batch mode	Cost and time	An array of workflow instances	1. It is used to reduce cost and cost	Cloud Environment	SwinDeW-
A Particle Swarm Optimization-based Heuristic for Scheduling [4]	Dependency mode	Resource utilization, time	Group of tasks	1. it is used for three times cost savings as compared to BRS 2.It is used for good distribution of workloa onto resources	Cloud Environment	Amazon EC2
Improved cost-based algorithm for task scheduling [1]	Batch Mode	Cost, performance	Unscheduled task groups	1.Measures both resource cost and computation performance 2. Improves the computation/communication ratio	Cloud Environment	Cloud Sim

Table 1: Comparison between Existing Scheduling Algorithms

3. Conclusion

Scheduling is one of the major issues in the management of application execution in cloud environment. Surveyed the various accessible scheduling algorithms in cloud computing. Also noticed that disk space management is critical in virtual environments. The heuristic based strategy is used to schedule EMAN, a bio-imaging workflow application. It results into 1.5 to 2.2 time better optimization of make span and load balance. Genetic algorithm was used to find the schedule for workflow application that meet up the user defined budget and deadline. Multi-objective MGrid resource service composition and optimal-selection (MO-MRSCOS) problem is resolved by PSO. It minimizes execution time, cost, and maximize the reliability.PSO algorithm that assigned Cloud resources to workflow application. It consider both computation cost and data transmission cost when finding schedule. PSO attained 3 times cost saving as compared with BRS (Best Resource Selection. Existing scheduling algorithms does not believe reliability and availability. Therefore there is a need to implement a scheduling algorithm that can advance the availability and reliability in cloud environment.

REFERENCES

[1]. Sujit Tilak , Prof. Dipti Patil," A Survey of Various Scheduling Algorithms in Cloud Environment", Volume 1, Issue 2 (September 2012) PP: 36-39.

[2]. Lovejit Singh, Sarbjeet Singh," A Survey of Workflow Scheduling Algorithms and Research Issues", Volume 74– No.15, July 2013.

[3]. Er.Shimpy, Mr. Jagandeep Sidhu,"Different Scheduling Algorithms In Different Cloud Environment", Vol. 3, Issue 9, September 2014

[4]. Dr. Amit Agarwal, Saloni Jain," Efficient Optimal Algorithm of Task Scheduling in Cloud Computing Environment", volume 9 number 7– Mar 2014
[5]. Gaurang patel,Rutvik Mehta,Upendra Bhoi,"Enhanced load balance Min-Min Algorithm for static meta task scheduling in cloud computing",2015
[6]. Navdeep Kaur, Khushdeep Kaur,"Improved Max-Min Scheduling Algorithm", Volume 17, Issue 3, Ver. 1 (May – Jun. 2015), PP 42-49
[7]. Antony Thomasa , Krishnalal G, Jagathy Raj V P,"Credit Based Scheduling Algorithm in Cloud Computing Environment",2014
[8]. Liyun Zuo, Lei Shu, (Member, IEEE), Shoubin Dong, Chunsheng Zhu, (Student Member, IEEE), and Takahiro Hara, (Senior Member, IEEE),"A Multi-Objective Optimization Scheduling Method Based on the Ant Colony Algorithm in Cloud Computing", December 2015
[9]. Mohd Rahul, Impact of Cloud Computing on IT Industry: A Review & Analysis, International Journal of Computer and Information Technology (ISSN: 2279 – 0764) Volume 01– Issue 02, November 2012