



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Survey of Task Scheduling Algorithms in Cloud Computing

Manpreet Kaur

M.Tech Student, Dept. of Computer Science, SGGSWU Fatehgarh Sahib, Punjab, India

ABSTRACT: Cloud computing means storing and accessing data and programs over the internet Instead of our computer's hard drive. There are number of issues define in task scheduling, such as performance, cost, execution time, security and privacy. In these types the main drawback is allocation time is large. This can be overcome with the help of environment types i.e. Static and Dynamic allocation of VMs. Working on the existing QoS based algorithm which is integrated with the improved shortest job Task grouping algorithm i.e. tasks are assigned to the resources in the groups.

KEYWORDS: Cloud Computing, Task Scheduling, Task Scheduling Algorithms.

I. INTRODUCTION

1.1 Cloud Computing

Cloud Computing: In the modern era of IT if we talk about the "Cloud computing" it is one of the hottest topics of the industry. A lot can be done on this topic. If we talk about the term "cloud" it arises from the network that represents the internet. We can't define Cloud computing easily because there are many definitions for it, which explains the same meaning: the Internet. Cloud computing is a way to use the Internet in the daily life from your PC and Laptop.[1] Cloud Computing came into action to know what happens when our data is moved to internet that is in "cloud".

1.2 Characteristics of Cloud Computing:-

- **Shared Infrastructure:** - In the cloud computing we are having shared resources like storage and networking.
- **Dynamic Provisioning:** - It means its having the capability to scale up his devices or we can say infrastructure. Means user need not to worry about the data.
- **Network Access:** - It means user can access it from anywhere. User just needs an internet connection and it through mobile, Laptop, PC, Tab etc.[3]

1.3 Service Models:- Service models of the cloud computing provides the various models in terms of services to their consumers. We have three types of the service models in the cloud computing that are explained below:

- **Software as a Service (SaaS)** - In software as a Service, service providers host applications and clients access it through the World Wide Web. What actually SaaS do to see that we can also see the services provided by companies like Microsoft (Microsoft Office 365) and (Google doc). The service providers provide all the required programs by users at central location, and then customers can access it through the World Wide Web.
- **Platform as a Service (PaaS)** – In platform as a service as the names suggesting its meaning itself that by this service cloud providers provide the platform to the customers where they can develop, deploy and manage their own services. Examples of the PaaS can be the Google App Engine and Microsoft Windows Azure.
- **Infrastructure-as-a-service (IaaS)** - In infrastructure as a service as the names suggesting its meaning itself that by this service cloud providers provide the Infrastructure which is the whole implementation of the cloud environment on the consumer's system where they can do any work according to their requirement. Salesforce CRM (Customer Relationship Manager) is the providers of the IaaS [12].

1.4 Deployment Models

There are different types of clouds that you can subscribe to depending on your needs. As a home user or small business owner, you will most likely use public cloud services.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

- **Public Cloud** - A public cloud can be accessed by any subscriber with an internet connection and access to the cloud space.
- **Private Cloud** - A private cloud is established for a specific group or organization and limits access to just that group.
- **Hybrid Cloud** - A hybrid cloud is essentially a combination of at least two clouds, where the clouds included are a mixture of public, private, or community.
- **Community Cloud** - A community cloud is shared among two or more organizations that have similar cloud requirements.[2]

II. SCHEDULING

In the simple language if we define scheduling it is assigning number of resources to jobs or tasks to be executed. Now if we talk about the Cloud Computing field Job Scheduling and Load Balancing is one of the main security issues in the today's world. For this there are number of algorithms are exist in the cloud computing environment. There are so many algorithms for scheduling in cloud computing. The main advantage of scheduling algorithm is to obtain a high performance.[4]

Throughput – It is the total number of processes that complete their execution per time unit.

Waiting Time - The time the process remains in the ready queue.

Let's discuss brief introduction of some scheduling algorithms.

First Come First Serve Algorithm

In this algorithm jobs are executed on the basis of first come and first out. It is easy to understand and implement It's not having high performance and having high waiting time.

Shortest Job First (SJF)

Shortest job first algorithm is further having two categories i.e. Preemptive and Non-Preemptive. Let's discuss both these with help of example.

Preemptive SJF In this algorithm when the processes comes to queue it will sort it and while executing any process if any new process comes with the less burst time from the remaining time of the current executing process then it will preempt.

SCHEDULING PROCESS:-

Scheduling process basically contains three steps that are the following:-

- Resource discovering and filtering: - In first step Datacentre broker looks for the available resources in the network system and collects the status related to them.
- Resource Selection: - In the second step Target resource is selected based upon some parameters of the task and resource.
- Task Submission: - In the last step task is submitted to the resource selected. [1]

TASK SCHEDULING

Task Scheduling plays a key role to improve flexibility and reliability of systems in cloud. The main reason behind scheduling tasks to the resources in accordance with the given time bound. Cloud consists of a number of resources that are different with one other via some means and cost of performing tasks in cloud. Using resources of cloud is different so scheduling of tasks in cloud is different from the traditional methods of scheduling and so scheduling of tasks in cloud need better attention to be paid because services of cloud depends on them[4].

SCHEDULING TYPES: - There are two types static and dynamic.

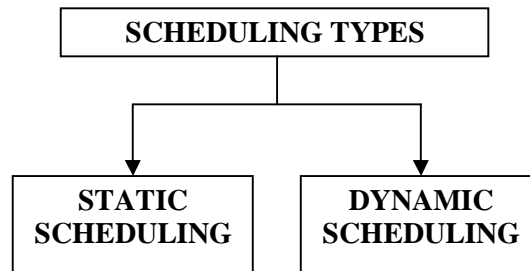


International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017



Static scheduling schedule tasks in known environment i.e. it already has the information about complete structure of tasks and mapping of resources before execution, estimates of task execution/running time.

Dynamic scheduling must depend on not only the submitted tasks to cloud environment but also the current states of system and computer machines to make scheduling decision [4].

ISSUES IN TASK SCHEDULING:-

a) Performance

Performance is the biggest issue in cloud adoption. The cloud must provide improved performance when a user moves to cloud computing infrastructure. Performance is generally measured by capabilities of applications running on the cloud system. Poor performance can be caused by lack of proper resources viz. disk space, limited bandwidth, lower CPU speed, memory and network connections etc.

b) Cost

With cloud computing, business gets the ability to save money on hardware and/or software but still requires spending more on the bandwidth. It is almost impossible to fully exploit the services of cloud computing without high speed communication channels.

c) Security and Privacy

Security, Performance and Availability are the three biggest issues in cloud adoption. The critical challenge is how it addresses security and privacy issues which occur due to movement of data and application on networks, loss of control on data, heterogeneous nature of resources and various security policies.

d) Execution Time

The execution time or CPU time of a given task is defined as the time spent by the system executing that task, including the time spent executing run-time or system services on its behalf. The mechanism used to measure execution time is implementation defined. It is implementation defined which task, if any, is charged the execution time that is consumed by interrupt handlers and run-time services on behalf of the system.

III. REVIEW OF LITERATURE

- **Upendra and Purvi (2013)** In this paper authors proposed a new algorithm with the modification in the Max-Min algorithm. This algorithm selects the task with average execution time. In Enhanced Max-min Task Scheduling Algorithm in Cloud Computing the task with the average or nearest greater than average task is selected and then assigned to the resource with the minimum completion time. Authors compared the results of the proposed algorithm with the Max-Min and showed that the overall makespan is reduced with the efficient performance. [6].
- **Rajiv Ranjan and Rajkumar Buyya** Hosting Internet-based application services. These applications have different composition, configuration, and deployment requirements. The simulation framework has the following novel features: (i) support for modelling and instantiation of large scale Cloud computing infrastructure, including data centers on a single physical computing node and java virtual machine; (ii) a self-contained platform for modelling data centers, service brokers, scheduling, and allocations policies; (iii) availability of virtualization engine, which aids in creation and management of multiple, independent, and co-hosted virtualized services on a data center node.[7]
- **O.M. Elzeki et al. (2012)** Proposed the new algorithm with the modification of the Max-Min algorithm. This algorithm based upon the expected execution time instead of expected completion time. In this algorithm expected



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

completion time for all the submitted tasks are calculated first then the task with the maximum completion time is selected and then submitted to the resource with the minimum execution time. They compared the results of the proposed algorithm with the Max-Min and RASA algorithm their results showed that their algorithm is working efficiently. They reduced the overall makespan [8].

- **Tarun Kumar Ghosh, Rajmohan Goswami(2012)** The traditional Min-Min and Max-Min algorithms are simple algorithms that produces a schedule that minimizes the makespan than the other traditional algorithms in the literature. In real scenario of Min-Min and Max-Min failed to produce a load balanced schedule. The proposed method has two-phases. In the first phase the traditional Max-Min algorithm is executed and in the second phase the tasks are rescheduled to use the un utilized resources effectively.[9]

- **Parsa and Maleki (2009)** In cloud computing authors proposed a new algorithm i.e. RASA. They took the two basic algorithms i.e. Max-Min and Min-Min. These two algorithms work on the estimated time for execution and completion of the task. In these execution time of each task is calculated on the resources. In MIN-MIN task with the minimum completion time is selected first and assigned to the resource with the minimum execution time. This procedure is followed with all the tasks. [10].

- **Jia and Jacky (2013)** In this paper purposed a new scheduling algorithm in which they integrated task grouping, priority aware and Shortest Job First algorithm. In this algorithm they used Gaussian distribution to generate the tasks [11].

- **Jia Ru** This project introduces an effective scheduling algorithm, which attempts to maximize cloud resources utilization, improve the computation ratio, and reduce makespan, overhead and delay in a cloud-based software system.[12]

- **Pinal Sanlot (2013)** In this paper author discussed about various existing scheduling algorithms. He categorized the job scheduling algorithms in two types' i.e. batch mode heuristic scheduling algorithms and Online Mode heuristic algorithms [13].

- **Sunil kumar Nakum, C. Ramakrishna, Amit Lathigara (2014)** Grid computing collects physically distributed resources from tablets, smart phones etc. to super computers to perform job requests by customers. Job scheduling refers to allocating grid user jobs to resources. Job scheduling in such an environment is major problem. Efficient job scheduling algorithms are required to make efficient use of the marvelous capability of the grid. In addition to this grid user may decide QoS he wants. This paper introduces reliability requirement percentage of grid user to Max-Min, Min-Min and RASA along with their comparative study[14].

- **S.Devipriya and C.Ramesh (2013)** In cloudcomputing, many tasks need to execute at a time by the available resources in order to achieve better performance, minimum completion time, shortest response time, resource utilization etc. So the scheduling tasks within cloud environment using Improved Max-min can achieve low rmakespan rather than original Max-min.[15]

- **Elhossiny Ibrahim and Fatma A. Omara(2016)** The Cloud Computing is a most widely spreading platform for executing tasks using virtual machines (VMs) as processing elements. In this paper, an enhancement task scheduling algorithm on the Cloud Computing environment has been introduced to reduce the make-span, as well as, decrease the price of executing the independent tasks on the cloud resources.[16]

- **S.K. Jeya Brindha, J. Angela Jennifa Sujana, T. Revath (2016)** Cloud provider provides services to run the tasks in cloud and maintain the quality of service for its customers. To ensure the cloud provider's QoS, we propose a novel agent-based scheduling mechanism to run the customers' tasks in cloud and provide dynamic resources. In agent-based scheduling, the authors have introduced a bidirectional announcement based bidding mechanism to allocate tasks and resources dynamically. In addition, it consists of three phases, i.e., basic matching phase, forward announcement-bidding phase and backward announcement-bidding phase[17].

- **Xiao-long Zheng, Ling Wang (2016)** A Pareto based fruit fly optimization algorithm (PFOA) is proposed to solve the task scheduling and resource allocating (TSRA) problem in cloud computing environment. First, a heuristic based on the property of minimum cost is proposed for initializing the population. Second, a resource reassign operator is designed to generate non-dominated solutions. Third, a critical path based search operator is designed to improve the exploitation capability [18].

- **Chaokun Zhang, Yong Cui, Rong Zheng, E Jinlong, Jianping Wu (2016)** The authors have investigated the scheduling problem with multi-resource allocation in cloud computing environments. They formulated the problem



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

of Multi-Resource Partial-Ordered Task Scheduling (MR-POTS) to minimize the makespan. In the first stage, the proposed Dominant Resource Priority (DRP) algorithm decides the collection of subtasks for resource allocation by taking into account the partial order relationship and characteristics of subtasks. [19]

- **Ekta S. Mathukiya, Piyush V. Gohel. (2015)** This paper introduces the multi-objective task scheduling algorithm for optimization of throughput of the system. This algorithm not only performs multi objective task scheduling but also performs non-dominated sorting for ordering of tasks. Thus, aim of this research work is to prove the effectiveness of the optimization method.[20]

IV. TASK SCHEDULING ALGORITHM

4.1 First Come First Serve Algorithm

In this algorithm jobs are executed on the basis of first come and first out. It is easy to understand and implement.. Job in the queue which come first is served. This algorithm is simple and fast. [13]

4.2 Round Robin algorithm

In the round robin scheduling, processes are dispatched in a FIFO manner but are given a limited amount of CPU time called a time-slice or a quantum. If a process does not complete before its CPU-time expires, the CPU is preempted and given to the next process waiting in a queue. The preempted process is then placed at the back of the ready list. [13]

4.3 Most Fit Task Scheduling Algorithm

In this algorithm task which fit best in queue are executed first. This algorithm has high failure ratio. [13]

4.4 RASA Algorithm

Saeed Parsa and Reza Entezari-Maleki [2] proposed a new task scheduling algorithm RASA. It is composed of two traditional scheduling algorithms; Max-min and Min-min. RASA uses the advantages of Max-min and Min-min algorithms and covers their disadvantages. [20]

4.5 Max-Min Algorithm

This algorithm chooses large tasks to be executed firstly, which in turn small task delays for long time. [13] Max-Min [9] algorithm follows the following procedure:

Phase1: First computes the completion time of every task on each machine and then for every task chooses the machine which processes the tasks in minimum possible time.

Phase 2: Among all the tasks in Meta Task the task with maximum completion time is selected and is assigned to machine. The task is removed from the list of Meta Task and the procedure continues until Meta Task list is empty. [21]

4.6 Min-Min Algorithm

This algorithm chooses small tasks to be executed firstly, which in turn large task delays for long time.[13] Min-Min[9] algorithm has following procedure:

Phase 1: First computes the completion time of every task on each machine and then for every task select the machine which processes the tasks in minimum possible time.

Phase2: Among all the tasks in Meta task the task with minimum completion time is selected and is assigned to machine on which minimum execution time is expected. The task is removed from the list of Meta Task and the procedure continues until Meta Task List is empty. [21]

V. CONCLUSION

This paper discusses various techniques and approaches related to the scheduling of the tasks to the virtual machines in the cloud computing. Various techniques like first in first out, priority scheduling etc. have been used by the authors in the past schedule the tasks to the virtual machines. In the "Multi objective task scheduling" the authors have used non dominated sorting of the tasks to schedule them to the virtual machines. In the future we would like to continue the study over this multi objective study to further improve the scheduling approach.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

VI. FUTURE SCOPE

We have done the non-dominated sorting of virtual machines (VMs) according to MIPS. In future we will take other parameters of VMs also to sort them for better performance.

REFERNCES

- [1] Sumit Arora, Sami Anand "Improved Task Scheduling Algorithm in Cloud Environment" International Journal of Computer Applications (0975 – 8887) Volume 96– No.3, June 2014
- [2] Alexa Huth and James Cebula "The Basics of Cloud Computing" 2011 Carnegie Mellon University. Produced for US-CERT, a government organization.
- [3]] Introduction to Cloud Computing, White Paper by Dialogic, 2010
- [4] Er. Shimpy1, Mr. Jagandeep Sidhu2 "Different scheduling algorithm in diffrent cloud environment" International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 9, September 2014
- [5] Raja Manish Singh#1, Sanchita Paul*2, Abhishek Kumar #3 "Task Scheduling in Cloud Computing: Review"(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 7940-7944
- [6] Upendra Bhoi1, Purvi N. Ramanuj "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing"(IJCSIT) International Journal of Computer Science and Information Technologies2 Volume 2, Issue 4, April 2013
- [7] Calheiros Rodrigo N., Ranjan Rajiv, De Rose César A. F., and Buyya Raj Kumar "CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services"
- [8] Elzeki O. M., Reshad M. Z., Elsoud M.A "Improved Max-Min Algorithm in Cloud Computing" International Journal of Computer Applications (0975 – 8887) Volume 50 – No.12, July 2012, pp.22-27
- [9] Ghosh Tarun Kumar, Goswami Rajmohan, Bera Sumit, Barman Subhabrata "Load Balanced Static Grid Scheduling Using Max-Min Heuristic" 2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing, pp.419-423
- [10] Parsa Saeed and Maleki Reza Entezari "RASA: A New Task Scheduling Algorithm in Grid Environment" 2009 World Applied Sciences Journal 7 (Special Issue of Computer & IT), pp. 152-160
- [11] Ru Jia and Keung Jacky "An Empirical Investigation on the Simulation of Priority and Shortest-Job-First Scheduling for Cloud-based Software Systems" 2013 22nd Australian Conference on Software Engineering, pp.78-87
- [12] Ru Jia "An Investigation on Scheduling Policies for Cloud-based Software Systems"
- [13] Salot Pinal "A Survey of Various Scheduling Algorithm in Cloud Computing Environment" 2013 IJRET Volume: 2 Issue: 2, pp.131-135
- [14] Sunil kumar Nakum, C. Ramakrishna, Amit Lathigara, "Reliable RASA scheduling algorithm for grid environment", International Conference on Computer Communication and Systems, 2014 IEEE.
- [15] S.Devipriya, C.Ramesh "Improved Max-Min heuristic model for task scheduling in Cloud" 2013 IEEE
- [16] Elhossiny Ibrahim, Fatma A. Omara "Task Scheduling Algorithm in Cloud Computing Environment Based on Cloud Pricing Models" 2016 IEEE
- [17] S.K. Jeya Brindha, J. Angela Jennifa Sujana, T. Revathi, "Agent based bidirectional bidding mechanism for efficient scheduling of real time tasks in cloud computing", International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), (2016)
- [18] Xiao-long Zheng, Ling Wang, "A Pareto based fruit fly optimization algorithm for task scheduling and resource allocation in cloud computing environment", IEEE Congress on Evolutionary Computation (CEC), 2016.
- [19] Chaokun Zhang, Yong Cui, Rong Zheng, E Jinlong, Jianping Wu, "Multi-Resource Partial-Ordered Task Scheduling in cloud computing", 24th International Symposium
- [20] Ekta S. Mathukiya, Piyush V. Gohel. "Efficient Qos Based Tasks Scheduling using Multi-Objective Optimization for Cloud Computing" International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 8, August 2015

BIOGRAPHY



I, Manpreet Kaur currently doing M.Tech (Integrated) from Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India. My areas of interests are Cloud Computing and Data Mining.