

HEURISTICS BASED MULTI QUEUE JOB SCHEDULING FOR CLOUD COMPUTING ENVIRONMENT

V.Rajeshram¹, C.P. Shabariram²

¹ME, Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Tamilnadu, India

²Assistant Professor, Department of CSE, Sri Shakthi Institute of Engineering and Technology, Tamilnadu, India

Abstract

With the influx of large number of users, delivering the services to them and to meet their quality requirements is still challenging. Cloud computing is built upon the advancements of virtualization technology and distributed computing. Cloud computing itself is a grid computing with additional property of elasticity and scalability. Since cloud usually contains a heterogeneous pool of resources, scheduling plays a vital role in cloud computing. Scheduling in cloud computing denotes the execution of current jobs in a given constraint. QoS is inevitably needed to deal with scheduling in cloud computing environment. Efficient scheduling mechanism should meet the QoS parameters and should enhance resource utilization. Traditional scheduling algorithms such as FCFS, SJF, EASY, round robin and backfilling possesses fragmentation problem. The proposed heuristics based multi queue job scheduling method overcomes the problem of fragmentation by providing an on-demand strategy for scheduling. On-demand service provided by introducing priority. Priority is a comprehensive concept, computed based on QoS parameters. The proposed system, by considering the priority concept reduces the problem of fragmentation and also overcomes "starving to death" problem in scheduling. The proposed method thus can efficiently executes the user requirements and can manifest better performance.

Keywords: Multi-Queue Scheduling; Heuristics

1. INTRODUCTION

Cloud computing is an emerging technology and also a trend which provides services at any time anywhere through internet connection with on-demand and pay-as-you-go model. Cloud is a pool of resources where users can rent the resources as per needs and pay for what they use. It has less capital expenditure for setting up an organization. There is no need for huge volume of resources. It is enough to have small amount of physical resources from which users can set to use virtualized resources. Resource sharing provides an important advantage in cloud computing that leads to job scheduling process. Job scheduling is one of the important challenge in cloud computing environment. There are multiple job scheduling algorithms are available in the cloud computing environment like FCFS, SJF, Min-Max and Round Robin.[7] Cloud computing is an enhanced concept of basic web services and grid computing. It contains three major services. They are IaaS (Infrastructure-as-a-Service), PaaS (Platform-as-a-Service) and SaaS (Software-as-a-Service). All the services combined together and will be known as Everything-as-a-Service (EaaS). This is the service needed by every organization. It helps to reduce the maintenance cost of all types of organization. The major service providers of cloud are Amazon (AWS), Google (Google App Engine) and Microsoft (Azure).

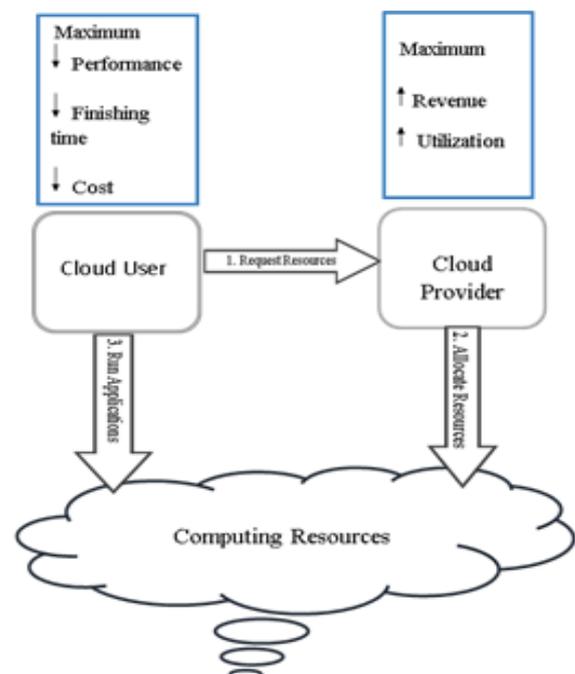


Fig -1: Cloud usage scenario

They provide stream less service to their clients. The cloud deployment model can be categorized as public cloud, private cloud, hybrid cloud and community cloud. Public cloud denotes every users can access the service provided by the third party, that correctly defines all the users may share the space in same location. Private cloud is created and maintained by an organization for their own purpose and public users cannot access that. Hybrid cloud is the

combination of public and private cloud. If set of organizations have similar requirements then they share the context known as community cloud. The main features of cloud computing are elasticity, economy, reliability and on-demand. [6] This paper presents a bird's eye view on the various algorithms for scheduling cloud resources.

Private cloud environment has been developed by organization using many tools or environments. OpenStack is the most preferable tool to develop public and private cloud environment. Because of it is open-source, scalable, easy to integrate with public cloud, integrated management tools and it has both option for public and private cloud. [8] Generally OpenStack is known as cloud operating system.

The paper is organized as follows: Section "Scheduling landscape" gives a brief introduction to scheduling. Section "Heuristic scheduling" discusses the ways to achieve scheduling efficiency based on past history. Section "Scheduling in Cloud" discusses the various scheduling algorithms in cloud. Section "Multi Queue job scheduling" narrates about scheduling using multiple queues. Finally the concluding remarks are presented in the "Conclusion and Future work" section.

2. SCHEDULING LANDSCAPE

Scheduling denotes the arrangement of process, threads, tasks or jobs based on some criteria and user requirements. In job scheduling the jobs allocated to resource. It is important because it manages the resource utilization and reduces the cost.

There are several classifications in the job scheduling strategies. The two major types are preemptive and non-preemptive. Preemptive scheduling depends on the priority. The current executing process state is changed when the new job to be executed interrupts the current process. Examples for preemptive scheduling are round-robin and priority based scheduling. Non-preemptive scheduling denotes the state of the process cannot be changed. i.e., the new process could not change the current executing process state. [4] Examples for non-preemptive scheduling are FCFS and SJF. Priority scheduling is another aspect where the priority is assigned internally or externally.

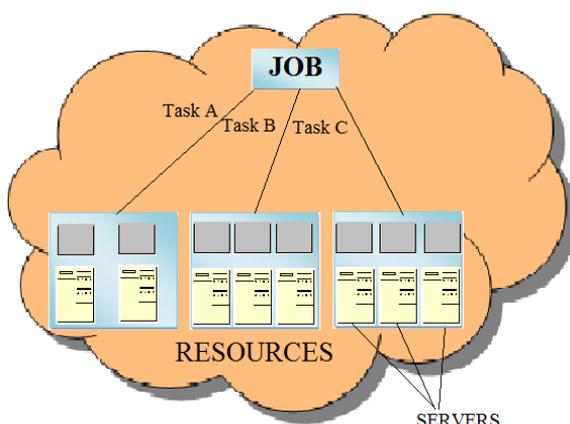


Fig -2: Scheduling tasks to available resources

Internal priority is assigned by measurable qualities or quantities to compute process.

3. HEURISTICS SCHEDULING

Heuristics based job scheduling can be under both of Preemptive and non-preemptive. It denotes that the jobs are scheduled based on their previous history. It is again classified as batch mode heuristics and online mode heuristics. In the batch mode, the jobs are collected and arranged; finally it is executed in the system. [3] The main examples for batch mode heuristics algorithm are First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Min-Min algorithm and Max-Min algorithm. In the online mode, the order is not changed i.e., the order in which the jobs are executed is the same. [4] The example for online mode scheduling is Most Fit Task scheduling algorithm (MFTF).

4. SCHEDULING PROCESS IN CLOUD

Scheduling in cloud is done in three stages. The stages are

- ❖ **Discover the resources and filter it:** Datacenter Broker search the resources present in the cloud and gather the status information related to the resource.
- ❖ **Select the suitable resource:** Destination resource is selected based on desired parameters of task and resource. This stage is the deciding stage.
- ❖ **Submit the task to the resource:** Task is submitted to resource selected depending on the parameter.

5. SCHEDULING ALGORITHMS IN CLOUD COMPUTING

There are several scheduling algorithms are available in cloud computing environment. One important strategy is backfill.

5.1 BACKFILL

Backfill is an important category in scheduling where the jobs can be moved if they will not delay the ahead job in the queue. It is known as Backfill strategy. Two types of backfilling are EASY (Extensible Argonne Scheduling System) backfilling and conservative backfilling. [3]

5.1.1 Easy Backfilling and Conservative Backfilling

In case of EASY backfilling the jobs can be moved if they do not delay first waiting job. In the conservative backfilling the jobs can be moved if they do not delay any jobs in the queue.

5.1.2 Aggressive Backfilling

Aggressive backfilling also same as backfilling strategy but it ensures the first job of the queue must not be delayed. [3]

5.1.3 CBA Backfilling

CBA denotes the Combinational Backfill Algorithm where a group of small jobs selected to backfill the resource gap and

increases the resource utilization. Combination of more than one jobs to be backfilled is known as CBA. If there is no available jobs then that again lead to EASY backfilling. The above all creates fragment in the scheduling which leads to the heuristic based multi-queue job scheduling. [2]

5.1.4 SJF-BF

SJF-BF denotes Shortest job first Backfill. The job with minimum execution time move ahead in the queue and the backfill fills the space created by high priority jobs. [5]

6. MULTI-QUEUE JOB SCHEDULING

Multi Queue job scheduling is a new strategy that uses multiple queues to scheduling the jobs. In the Multi-queue job scheduling the scheduler groups the jobs based on the burst time. The jobs submitted by clients are sorted in the ascending order and then proceed to the different queues. The jobs sorted in the ascending order gives equal importance to all the jobs.

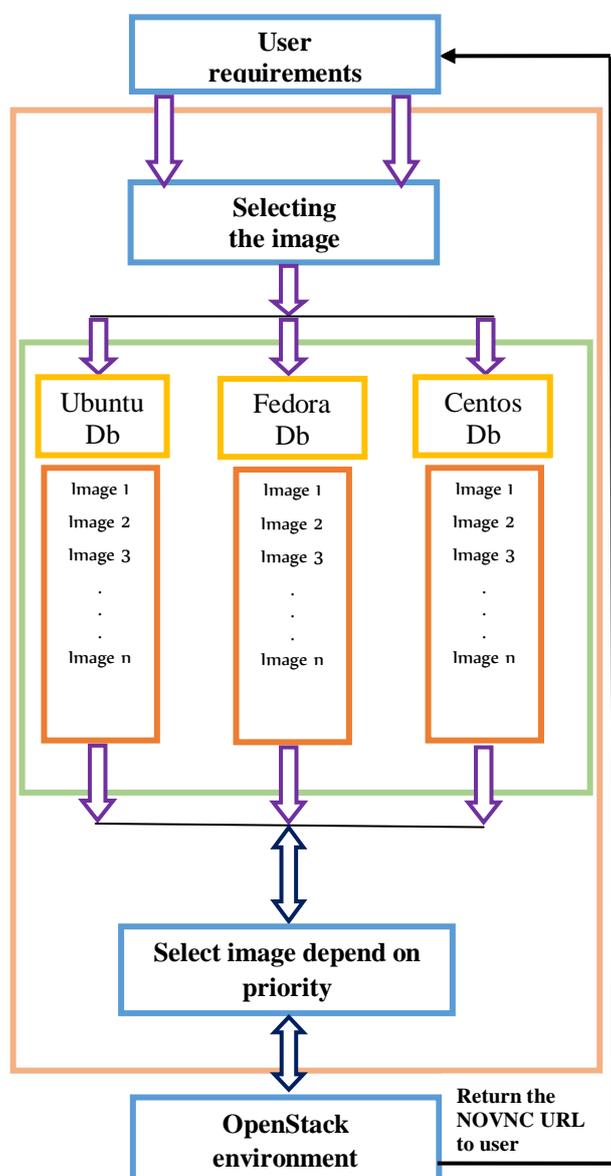


Fig -3: Multi-queue job scheduling algorithm

That increases the client satisfaction because the client requirements are varied based on their needs. The dynamic allocation gives the benefit of reducing the starvation. Queue manager is a part of the cloud computing that manages the utilization of all the resources. It also balances the load by distributing jobs among the meta-scheduler and its disposal. There are three queues in the system. The queues are small, medium and large.

- Small queue contains the first 40% of the jobs.
- Medium queue contains the next 40% of jobs.
- Long queue contains the remaining 20% of the jobs. [1]

From these three different queues jobs are selected dynamically which eliminate the starvation that is "starve to death". These dynamically selected jobs are entered into the cloud environment. The best allocation reduces the time and allocation of the jobs. It provide better and satisfactory environment when compared to the traditional algorithms. It provides optimum use of resources and attains high resource utilization and provides Quality of System in the cloud environment. The multi queue job scheduling algorithm used to reduce the cost of reservation and on-demand plan. Here jobs are put it into three different queues and selected randomly from the queue. That provides optimum solution the scheduling problem and efficient utilization of unused free spaces. The Multi-queue scheduling algorithm can be extended to the on demand and reservation category. When the client requests a new job, priority needs to be assigned so that the queue that contains the sorted jobs needs to be rearranged again.

7. IMPLEMENTATION

In the implementation part OpenStack used here for developing cloud computing environment. It is an open source software tool. OpenStack is one of the best development environment to develop private cloud. Here Platform-as-a-Service (PaaS) service will be provided and the instances will be scheduled instead of job. In the private cloud environment only limited resources available. Due to the limited resource the number of instances also be limited. Within the number of resources administrator need to do the scheduling and allocation to the user. Here different operating systems provided as an instance. They called as instance in OpenStack Environment. For each operating system the number of instances will be limited and predefined. The admin create instances respect to the predefined numbers for each operating system. Instance scheduling done here which is one of the job in cloud environment. The instance scheduled here and will be provided to the user. The new instance creation performed depend on the resource available. If it is not then the free instance will be allocated to the user. The multiple queues used to maintain different operating system category. Each operating system contains multiple images on the same property. The property defined here as a RAM size, Memory size and OS version. Among the multiple images one image will be return to user. The criteria used to select image is Most Frequently Used (MFU). That is referred from Most Frequently Used (MFU) algorithm which is used in page

replacement problem. The number of user request exceeds the available instances means the user must wait until instance will be free. Likewise if all the resources free means among them which one need to be schedule to the user also done here. The selection process done by which is in shutoff status and also has maximum usage count. Maximum usage count ensures the Most Frequently Used criteria which is based on the heuristics of the instance.

8. CONCLUSION AND FUTURE WORK

Cloud computing is an emerging topic in distributed computing and it is the door to the future of information technology. Scheduling is one of the major issue among the several issues that cloud computing is facing today. Heuristics based scheduling is a system for solving the problem of scheduling tasks in cloud environment based on multi queue concept. Scheduling is one main problem that researchers are discussing these days. The system been developed aims at maximizing resource utilization in scheduling based on MQS. The parameter chosen for scheduling is burst time. Here the incoming jobs are sorted on ascending order based on burst time. Dynamically select jobs from queue that proceed for processing. This system obtains the optimum use of resources for cloud computing and provide Quality of Service in cloud environment.

REFERENCES

- [1]. AV.Karthick, R.Ganapathy Subramanian, Dr.E.Ramaraj, "An Efficient Multi Queue Job Scheduling for Cloud Computing," World Congress on Computing and Communication Technologies, ISSN 978-1-4799-2876-7/13, March 2013.
- [2]. Lawson, Barry G., Smirni, Evgenia,"Multiple-queue Backfilling Scheduling with Priorities and Reservations for Parallel Systems" Department of Computer Science, College of William and Mary Williamsburg, VA 23187-8795, USA
- [3]. A.D.Techiouba, G.Capannini, Ranieri Baraglia, D.Puppin, M.Pasquali, Laura Ricci (2002), "Backfilling strategies for scheduling streams of jobs on computational farms", European CoreGRID NoE, IST-2002-004265.
- [4]. Anand K Chaturvedi and Rajendra Sahu," New Heuristic for Scheduling of Independent Tasks in Computational Grid," International Journal of Grid and Distributed Computing Vol. 4, No. 3, September, 2011.
- [5]. Anurag Mishra, Dharmender Singh Kushwaha and Shakti Mishra, "An Improved Backfilling Algorithm: SJF-BF," Int. J. on Recent Trends in Engineering & Technology, Vol. 05, No. 01, Mar 2011.
- [6]. Hassan Rajaei And Mohammad B. Dadfar, "Comparison Of Backfilling Algorithms For Job Scheduling In Distributed-Memory Parallel System," American Society for Engineering Education Annual Conference & Exposition, 2006.
- [7]. Cited <http://www.explainthatstuff.com/cloud-computing-introduction.html>
- [8]. Cited <http://getcloudify.org/2014/07/10/what-is-openstack-tutorial.html>

BIOGRAPHIES



V. Rajeshram was born in Aruppukottai on 04th May 1992. He received his B.E. (CSE) from MIET Engineering College of, Trichy, Tamil Nadu in 2013. He is currently pursuing his M.E. (CSE) in Sri Shakthi Institute of Engineering and Technology, Tamil Nadu, India. He has published a paper in international journal and has attended one international and one national level conferences.



Mr. C.P.Shabariram received the BE degree in Computer Science and Engineering from Anna University Tiruchirappalli and ME degree in Computer Science and Engineering from Kathir College of Engineering, Coimbatore. He is assistant professor at Sri Shakthi Institute of Engineering and Technology. His area of interests include Software Engineering, Cloud Computing and Big Data. He has published numerous conference papers and journal articles on various aspects of improving software quality.