A REVIEW OF VARIOUS SCHEDULING ALGORITHMS ON CLOUD COMPUTING

Preeti Rawal

Dept of CSE & IT, BPSMV, Khanpur kalan,

(preetirawal3@gmail.com)

Abstract

Cloud computing is set of IT service that are provided to customer over a network and these services are delivered by third party provider who own the infrastructure. Cloud computing is general term used to describe a new class of network based computing over internet. Scheduling the task is quite a challenging in cloud environment. Scheduling algorithms need to be proposed to overcome the problem due to network scheduling strategy is the key technology in cloud computing. It describes various scheduling algorithms.

Keyword

Introduction**,** Characteristics**,** Various Scheduling algorithms

1. INTRODUCTION

Cloud computing today widely used concept in field of technology. Mainly cloud computing used for servicing memory as well as time. For storage of photos online instead of home computer or using social networking sites or web mails cloud computing is used. Cloud computing provide online invoicing service. Mainly cloud computing refers to delivery of computing service over the internet instead of keeping data on own hard drive and updating needs of the service. Examples of cloud computing online file storage, social networking sites, webmail's and online business application. Cloud computing allow access of information and computer resources from anywhere connection is available.

Cloud computing is becoming one of the most explosively expanding technologies in the computing industry today [1]. It enables users to Cloud computing model promotes availability and is composed of five essential characteristics, three service models and four deployment models [2].

1.2. Characteristics

Cloud computing has many characteristics like on demand self-service, Broad network access, resource pooling and measured services, etc.

On Demand self-service means that customer can request or manage their own computing resources. Broad Network Access allows service

to be offered over internet of private networks Such as phones, laptops, computers and personal digital assistants.

Pooled Resources means that customer draw from pool of computing resources usually remote data centers [1]. The provider's computing resources are used to serve multiple consumers using a multi-tenant model having different physical and virtual resources dynamically assigned and reassigned according to consumer demand [2].

Rapidly and elastically can be scaled larger or smaller, and use of service is measured and customers are billed accordingly. It allows the subscriber to increase or either decrease services. Mainly capabilities available often appear too unlimited to the subscriber and can purchased in any quantity at any time [2].

2. Various scheduling algorithms

2.1. FCFS

First come First serve basis means that task that come first will be execute first [3]. Shortest job next is advantageous because of its simplicity and it minimizes the average amount of time. Each process has to wait until its execution is complete. It manages ready queue which serve as first in first out means that the first job will be processed first without other preferences.

Algorithm FCFS:

1. Initialize Tasks.

2. First task assigned to the queue and add tasks up to n numbers.

3. Add next task 'I' at last position in the main queue [4].

2.2. Min-Min Algorithm

Min-Min algorithm selects the smaller tasks to be executed first [3]. An algorithm depends on the original Min-min algorithm. It is called QoS guided Min-min and it assigns tasks with high bandwidth before others. QoS acts similar to Min-min when available tasks have the same bandwidth so it preferred to use QoS guided. Min-min whenever submitted tasks have large bandwidth. At that moment QoS guided Min-min produces better results. Similar to QoS guided Min-min new algorithm called QoS priority grouping scheduling that is proposed by F. Dong et al. QoS Suffrage is new task scheduling algorithm presented by E. Ullah Munir. This algorithm considers network bandwidth and assigns tasks based on their bandwidth requirement. It achieves smaller make span compared to Max-min, Min-min; QoS guided Min-min and QoS priority grouping algorithms. K. Etminani et al. provided a new algorithm which uses Max-min and Min-min algorithms to select one of these two algorithms that depend on standard deviation of the expected completion times of the tasks on each of the resources. RASA uses the strategy to execute small tasks before large ones that is Min-Min and applies the Max-min strategy to avoid delays in the execution of the large tasks which support concurrency in the execution of large and small tasks [5].

2.3. Max-Min algorithm

Max-Min algorithm selects the bigger tasks to be executed first. The Max-min algorithm is used in

distributed environment which begins with a set of unscheduled tasks. Max-Min algorithm calculates the expected execution matrix and expected completion time of each task on available resources. Next choose the task with overall maximum expected completion time and assign it to the resource with minimum overall execution time. Finally recently scheduled task is removed from the meta-tasks set and update all calculated times, then repeat until meta-tasks set become empty. In the Max-min algorithm, rj represents the ready time of resource Rj to execute a task, while Cij and Eij represent the expected completion time and Execution time respectively. Task Tk with maximum expected completion time is chosen to be assigned for corresponding resource Rj that gives minimum execution time. Each of Max-min, Min-min and RASA algorithms have running time complexity of O (mn2), where m is the number of resources currently in the system and n is the number of submitted tasks which is to be scheduled [5].

2.4. Round-Robin algorithm (RRA)

It is one of the oldest, simplest, and fairest and most widely used scheduling algorithms, especially designed for timesharing system. A small unit of time called time slices or quantum is defined. The CPU scheduler goes around the queue, allocating the CPU to each process for a time interval of one quantum. The CPU scheduler picks the first process from the queue, sets a timer to interrupt after one quantum then dispatches the process. If the process is still running at the end of the quantum, the CPU is preempted and the process is added to the tail of the queue. The process finishes before the end of the quantum, the process itself releases the CPU voluntarily [4]. Mainly advantage of RR algorithm is that it utilizes all the resources in a balanced order. Scheduler allocates equal number of VMs to all the nodes which ensure fairness. And the major drawback of using RR algorithm is that the power consumption will be more as many nodes will be kept turned-on for a long time. If single node runs the three resources then all the three nodes will be turned on when Round Robin is used which will consume a significant amount of power [6].

2.5. Priority Scheduling Algorithm

This Scheduling algorithm is preemptive in which all things are based on the priority in this scheduling algorithm each process in the system is based on the priority whereas highest priority job can run first whereas lower priority job can be made to wait, the biggest problem of this algorithm is starvation of a process [4].

2.6. Shortest Response Time First

In SRTF each process is assigned a priority and Equal-Priority processes are scheduled in FCFS manner. The shortest-Job-First (SJF) algorithm is a special case of priority scheduling algorithm. An SJF algorithm is a priority algorithm where the priority is the inverse of the next CPU burst. As longer the CPU burst it lower the priority and vice versa. Priority can be defined internally or either externally. Internally defined priorities use some measurable quantities [6]. RR Scheduling algorithm is deal

with different approach in this algorithm CPU is allocated to the process with least burst time [4].

3. Conclusion

Cloud Computing cover wide area for providing services to the subscriber. In this paper we discuss about cloud computing types, characteristics, various scheduling algorithms, how we manage our resources and discussing about providing services to the subscriber. We concluded that a defined scheduling algorithm gives high throughput and cost effective but they do not consider reliability and availability. So we need more algorithms to improve efficiency of cloud computing. For better resource management in cloud computing we need better scheduling algorithms and management technique.

References

[1].Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, Sonia Lopez-Alarcon, Warren Carithers,"Efficient Resource Management for Cloud Computing environments",pp.1-8.

[2].Shirley Radack, "a review of features, benefits, and risks, and recommendations for secure, efficient implementations", itl Bulletin, June 2012, pp.1-7.

[3].Er.Shimpy, Mr. Jagandeep Sidhu "different scheduling algorithms in different cloud environment",International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 9, p.p.1-4, September2014.

[4].Rajveer Kaur, Supriya Kinger "Analysis of Job Scheduling Algorithms in Cloud Computing", International Journal of Computer Trends and Technology(IJCTT), volume 9, pp.1-8, March 2014.

[5].O.M.Elzeki, M.Z.Reshad, M.A.Elsoud "Improved Max-Min Algorithm in Cloud Computing", International Journal of Computer Applications (0975 – 8887) Volume 50- No.12, July 2012,pp. 1-6.

[6]. Vignesh V, Sendhil Kumar KS, Jaisankar N, "Resource Management And Scheduling In Cloud Environment" ,International Journal of Scientific and Research Publications, Volume 3, Issue 6,pp.1-6, June 2013.