Cloud Computing – An Important Aspect

Prashant Saxena1
1 HOD Computer Science & Mathematics
1Kimberley The Boarding School, Panchkula, Haryana

Abstract—Cloud computing is known as a source of energetic services using very large scalable and virtualized resources over the Internet. Due to innovation of cloud computing field, there is no many regular task scheduling algorithm used in cloud environment. Especially that in cloud, there is a high communication expenditure that prevents well known task schedulers to be applied in large scale distributed environment. Today, researchers attempt to build job scheduling algorithms that are well-matched and relevant in Cloud Computing environment Job scheduling is most important task in cloud computing environment because user have to pay for resources used based upon time. Hence competent consumption of resources must be important and for that scheduling plays a imperative role to get highest advantage from the resources. In this paper we are studying various scheduling algorithm and issues related to them in cloud computing.

Keywords: cloud computing, scheduling algorithm

I. INTRODUCTION

Cloud computing is known as a source of dynamic services using very large scalable and virtualized resources over the Internet. Various definitions and interpretations of “clouds” and / or “cloud computing” exist. With particular respect to the a range of usage scopes the term is engaged to, we will try to give a representative (as opposed to complete) set of definitions as recommendation towards potential usage in the cloud computing related research space. We attempt to imprison an abstract term in a way that best represents the technological aspects and issues related to it. In its broadest form, we can define a 'cloud' is an elastic effecting environment of resources involving numerous stakeholders and providing a metered service at several granularity for a particular echelon of quality of service. To be more precise, a cloud is a podium or infrastructure that enables execution of code (services, applications etc.), in a managed and elastic fashion, whereas “managed” means that consistency according to pre-defined quality parameters is automatically ensured and “elastic” implies that the resources are put to use according to actual current requirements observing overarching requirement definitions – implicitly, elasticity includes both up- and downward scalability of resources and data, but also load-balancing of data throughput. Job scheduling is one of the major activities performed in all the computing environments. Cloud computing is one the upcoming latest technology which is developing severely. To proficiently increase the working of cloud computing environments, job scheduling is one the tasks performed in order to gain maximum profit. The goal of scheduling algorithms in distributed systems is spreading the load on processors and maximizing their utilization while minimizing the total task execution time. Job scheduling, one of the most well-known optimization problems, plays a key role to improve flexible and reliable systems. The main purpose is to schedule jobs to the flexible resources in accordance with flexible time, which involves finding out a proper sequence in which jobs can be executed under transaction logic constraints. There are main two categories of scheduling algorithm (BMHA) and online mode heuristic algorithms. In BMHA, Jobs are queued and collected into a set when they arrive in the system. The scheduling algorithm will start after a fixed period of time. The main examples of BMHA based algorithms are; First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Min–Min algorithm and Max–Min algorithm. By On-line mode heuristic scheduling algorithm, Jobs are scheduled when they arrive in the system. Since the cloud environment is a heterogeneous system and the speed of each processor varies quickly, the on-line mode heuristic scheduling algorithms are more appropriate for a cloud environment. Most fit task scheduling algorithm (MFTF) is suitable example of On-line mode heuristic scheduling algorithm.

1) FCFS Algorithm: Job in the queue which come first is served. This algorithm is simple and fast.

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.

3) Min–Min algorithm: This algorithm chooses small tasks to be executed firstly, which in turn large task delays for long time.

II. SCHEDULING

There has been various types of scheduling algorithm exist in distributed computing system. The majority of them can be applied in the cloud environment with appropriate verifications. The main advantage of job scheduling algorithm is to achieve a high performance computing and the best system throughput. Traditional job scheduling algorithms are not able to provide scheduling in the cloud environments. According to a simple classification, job scheduling algorithms in cloud computing can be categorized into two main groups; Batch Mode Heuristic Scheduling algorithms (BMHA) and online mode heuristic algorithms. In BMHA, Jobs are queued and collected into a set when they arrive in the system. The scheduling algorithm will start after a fixed period of time. The main examples of BMHA based algorithms are; First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Min–Min algorithm and Max–Min algorithm.

By On-line mode heuristic scheduling algorithm, Jobs are scheduled when they arrive in the system. Since the cloud environment is a heterogeneous system and the speed of each processor varies quickly, the on-line mode heuristic scheduling algorithms are more appropriate for a cloud environment. Most fit task scheduling algorithm (MFTF) is suitable example of On-line mode heuristic scheduling algorithm.
4) *Max – Min algorithm*: This algorithm chooses large tasks to be executed firstly, which in turn small task delays for long time.

5) *Most fit task scheduling algorithm*: In this algorithm task which fit best in queue are executed first. This algorithm has high failure ratio.

6) The basic idea is straightforward: each process is assigned a priority, and priority is allowed to run. Equal-Priority processes are scheduled in FCFS order. The shortest-Job-First (SJF) algorithm is a special case of general priority scheduling algorithm. An SJF algorithm is simply a priority algorithm where the priority is the inverse of the (predicted) next CPU burst. That is, the longer the CPU burst, the lower the priority and vice versa. Priority can be defined either internally or externally. Internally defined priorities use some measurable quantities or qualities to compute priority of a process.

A. Scheduling Process

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

- Resource discovering and filtering – Datacenter Broker discovers the resources present in the network system and collects status information related to them.
- Resource selection – Target resource is selected based on certain parameters of task and resource. This is deciding stage.
- Task submission -Task is submitted to resource selected.

![Cloud Communication process](image)

Fig. 1: Cloud Communication process.

III. EXISTING SCHEDULING ALGORITHM

The Following scheduling algorithms are currently prevalent in clouds.

A. Resource-Aware-Scheduling algorithm (RASA):

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. Though the deadline of each task, arriving rate of the tasks, cost of the task execution on each of the resource, cost of the communication are not considered. The experimental results show that RASA is outperforms the existing scheduling algorithms in large scale distributed systems.

B. RSDC (RELIABLE SCHEDULING DISTRIBUTED IN CLOUD COMPUTING):

Arash Ghorbannia Delavar,Mahdi Javanmard, Mehrdad Barzegar Shabestari and Marjan Khosravi Talebi[1] proposed a reliable scheduling algorithm in cloud computing environment. In this algorithm major job is divided to sub jobs. In order to balance the jobs the request and acknowledge time are calculated separately. The scheduling of each job is done by calculating the request and acknowledges time in the form of a shared job. So that efficiency of the system is increased.

C. An Optimal Model for Priority based Service Scheduling Policy for Cloud Computing Environment:

Dr. M. Dakshayini, Dr. H. S. Guruprasad [3] proposed a new scheduling algorithm based on priority and admission control scheme. In this algorithm priority is assigned to each admitted queue. Admission of each queue is decided by calculating tolerable delay and service cost. Advantage of this algorithm is that this and the maximum profits of the cloud computing service provider are achieved with this algorithm.

D. Improved Cost-Based Algorithm for Task Scheduling:

Mrs.S.Selvarani, Dr.G.Sudha Sadhasivam [7] proposed an improved cost-based scheduling algorithm for making efficient mapping of tasks to available resources in cloud. The improvisation of traditional activity based costing is proposed by new task scheduling strategy for cloud environment where there may be no relation between the overhead application base and the way that different tasks cause overhead cost of resources in cloud. This scheduling algorithm divides all user tasks depending on priority of each task into three different lists. This scheduling algorithm measures both resource cost and computation performance, it also Improves the computation/communication ratio.

E. Performance and Cost evaluation of Gang Scheduling in a Cloud Computing System with Job Migrations and Starvation Handling:

Ioannis A. Moschakis and Helen D. Karatza has proposed a gang scheduling algorithm with job migration and starvation handling in which scheduling parallel jobs, already applied in the areas of Grid and Cluster computing. The number of Virtual Machines(VMs) available at any moment is dynamic and scales according to the demands of the jobs being serviced. The aforementioned model is studied through simulation in order to analyze the performance and overall cost of Gang Scheduling with migrations and starvation handling. Results highlight that this scheduling strategy can be effectively deployed on Clouds, and that cloud platforms can be viable for HPC or high performance enterprise applications.

IV. COMPARISION

<table>
<thead>
<tr>
<th>Scheduling Algorithm</th>
<th>Scheduling Method</th>
<th>Scheduling Parameter</th>
<th>Scheduling Factor</th>
<th>Findings</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource – Aware- Scheduling</td>
<td>Batch Mode</td>
<td>Make Span</td>
<td>Grouped Task</td>
<td>It is used to reduce make span</td>
<td>Grid Environment</td>
</tr>
<tr>
<td>Algorithm (RASA)</td>
<td>Scheduling Mode</td>
<td>Performance Metric</td>
<td>Handling</td>
<td>Cloud Environment</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>----------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>RSDC</td>
<td>Batch Mode</td>
<td>Process time</td>
<td>Grouped Tasks</td>
<td>Cloud Environment</td>
<td></td>
</tr>
<tr>
<td>An optimal Model for Priority based service scheduling</td>
<td>Batch Mode</td>
<td>Quality of Service, Service request time</td>
<td>An array of workflow instances</td>
<td>High QoS, High throughput</td>
<td>Cloud Environment</td>
</tr>
<tr>
<td>A priority based job scheduling algorithm in cloud computing</td>
<td>Dependancy mode</td>
<td>Priority to each mode</td>
<td>An array of job time</td>
<td>Less finish time</td>
<td>Cloud environment</td>
</tr>
<tr>
<td>Extended Max-Min scheduling using PetriNet and load balancing</td>
<td>Batch mode</td>
<td>Load balancing finish time</td>
<td>Grouped Task</td>
<td>It is used for efficient load balancing and Petri net is used to remove limitation of Max-Min Algorithm</td>
<td>Cloud environment</td>
</tr>
<tr>
<td>Improved Cost Based Algorithm for task scheduling</td>
<td>Batch mode</td>
<td>Cost performance</td>
<td>Unscheduled task group</td>
<td>Measures both resource cost and computational performance</td>
<td>Cloud Environment</td>
</tr>
<tr>
<td>Performance and Cost evaluation of Gang Scheduling in a cloud computing System with job</td>
<td>Batch mode</td>
<td>Performance, Cost</td>
<td>Workflow with large number of job</td>
<td>The application of migrations and starvation handling had a significant effect on the model, Improves</td>
<td>Cloud Environment</td>
</tr>
</tbody>
</table>

Table 1: Comparison of various algorithms

V. CONCLUSION

Scheduling is one of the most important task in cloud computing environment. In this paper we have analyze various scheduling algorithm and tabulated various parameter. We have noticed that disk space management is critical issue in virtual environment. Existing scheduling algorithm gives high throughput and cost effective but they do not consider reliability and availability. So we need algorithm that improves availability and reliability in cloud computing environment.

REFERENCES