IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 4, June 2014.
www.ijiset.com
ISSN 2348 - 7968

TBP:A Threshold Based Priority Scheduling in Cloud Computing Environment

Dharmalingam.K[1] , Dhavamani.A[2] , Ravi.S[3]

[1]PG Scholar, Department of Computer Science & Engineering ,Hindustan University, Padur,Chennai, India.

[2]PG Scholar, Department of Computer Science & Engineering ,Hindustan University, Padur,Chennai, India.

[3]Assistant Professor, Department of Computer Science & Engineering ,Hindustan University, Padur,Chennai ,India.

***Abstract* - Cloud Computing is the delivery of computing as a service rather than a product, whereby software, shared resources and information are provided to computers and other devices as a utility over the internet. It exhibits characteristics like reliability, scalability multi-tenancy, empowerment of end users and device or location independence. Scheduling is the process of deciding how to commit resources between a variety of possible tasks. Existing algorithms like round robin, pre-emptive priority and Shortest job first algorithms are taken to compare with the proposed system. In round robin scheduling an additional load on the scheduler to decide the size of the quantum and it has longer average waiting time ,higher context switches, turnaround time and low throughput. Number of context switches are very high. It selects the load on random basis and leads to the situation where some nubs are heavily loaded and some are lightly loaded. Service providers negotiate with the users on the requirements of tasks including network bandwidth, complete time, task costs and reliability of task. In pre-emptive priority scheduling a pre-emption policy is normally for scheduling high priority activities when a capacity shortage appears. Pre-emption has been investigated fairly extensively relative to scheduling single-capacity resources. Threshold based priority scheduling is proposed algorithm which reduces average waiting and turnaround time ,high throughput and improves overall performance of the system.**

***Index Terms*-Cloud Computing Environment, Resource Scheduling, Round Robin, Pre-emptive Priority, Shortest Remaining Time First and Threshold based Priority.**

1.INTRODUCION

Cloud computing is internet based computing, whereby shared servers provides software, resources and data to the users and other services on demand. In case of providing such services, scheduling the request is most important. The services rendered by cloud to users grows day-by-day by an enormous amount. So, cloud service provider need to provide the services to the users request effectively and efficiently. Job scheduling is one of the major activities performed in all the computing environment. Cloud computing is one the upcoming and current trend latest technology which is developed drastically. To efficiently and effectively increase the working of cloud computing environment, job scheduling is one the tasks performed in order to gain maximum profit. The main goal of scheduling algorithms in distributed systems is spreading the load on processors and while minimizing their utilization. Job scheduling ,one of the most famous optimization problem, plays a key role to improve reliable and flexible systems. The main purpose is to schedule jobs according with adaptable time, which involves finding out a proper sequence in which jobs can be executed under transaction logic constraints. There are major two categories of scheduling algorithm.1).Dynamic Scheduling and 2).Static Scheduling algorithm. Both

have their own limitation and advantage. Static scheduling algorithm have lower performance than Dynamic scheduling algorithm but has a lot of overhead compare to it.

2.SCHEDULING

There has been various types of scheduling algorithms present in shared computing environment. Most of them can be applied in the cloud computing environment with suitable information.The major advantage of job scheduling algorithm is to achieve a best system throughput and high performance computing .Traditional job scheduling algorithms are not able to provide scheduling in the cloud environment. According to a simple allocaton, job scheduling algorithms in cloud computing can be categorized into two major groups: Online Mode Heuristic Algorithm(OMHA) and Batch Mode Heuristic Algorithm(BMHA) .In BMHA, when jobs are arrive in the system they are queued and collected into a list. The scheduling algorithm will start after fixed period of time. The major example of BMHA based scheduling algorithms are: First Come First Served Algorithm(FCFS),Round Robin(RR),Max-Min and Min-Min algorithm. By OMHA ,when jobs are arrive in the system and they are scheduled. Since the cloud computing is a heterogeneous system and the speed of each processor differs quickly, the OMHA are more suitable for cloud computing environment. Most Fit Task Scheduling algorithm (MFTS) is suitable example of Online Mode Heuristic Algorithm.

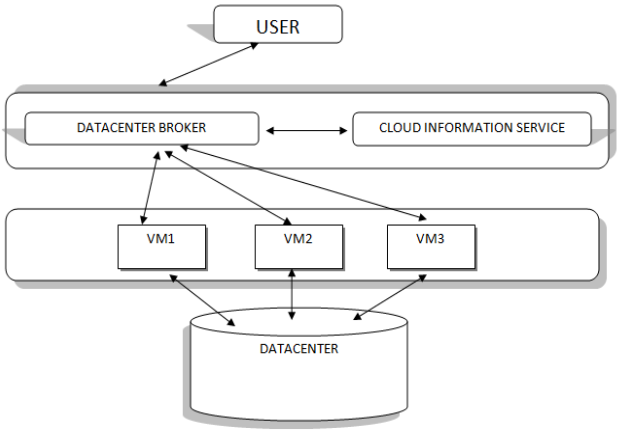
2.1 Scheduling Process

Scheduling process in cloud computing can be classified into three stages namely:

1.Resource Analysis and Cleaning: Datacenter broker discovers the resources present in the network system and collects status information related to them.

2.Resource Selection: Target Resource is selected based on certain parameters of resource and task.

3.Submit Task:Task is submitted to resource selected.



3.RELATED WORK

Venkatesa kumar.V and S.palaniswami et.al[6], in their paper have proposed overall resource utilization is high and processing cost has been reduced.Their experimental results clearly shows that proposed pre-emptive scheduling algorithm is effective in this regard.In this study, we present a novel turnaround time utility scheduling approach which focuses on both the high priority and the low priority tasks that arrive for scheduling.

Vinothina,Dr.R.Sridaran et.al[7],Survey on Resource Allocation Strategies in Cloud Computing clouds can make it possible to access applications and associated data from anywhere. Organizations are able to rent resources from cloud for storage and other computational purposes so that their infrastructure cost can be reduced significantly. However one of the major difficulty in cloud computing is related to optimizing the resources being allocated. Because of uniqueness of the system, resource allocation is performed with the objective of minimizing the costs and maximizing throughput associated with it.

Pinal Salot et.al[8],A survey of various scheduling algorithm job scheduling is most important task in cloud computing environment because user have to

pay-per-use for resources based upon time. Hence efficient and effective utilization of resources must be important and for that scheduling plays a important role to get maximum benefit from the resources.

Sujit Tilak et.al[9], A Survey of various scheduling algorithms is to make effective use of the tremendous capabilities of the cloud, efficient and effective scheduling algorithms are required. These scheduling algorithms are most commonly applied by cloud resource manager to optimally dispatch tasks to the cloud resources. There are relatively a large number of scheduling algorithms to maximize the total response time of the tasks in distributed systems. These algorithms try to minimize the overall completion time of the tasks by finding the most suitable resources to be allocated to the tasks. It should be noticed that minimizing the overall completion time of the tasks does not necessarily result in the minimization of the processing time of each and every individual task.

Shalmali Ambike et.al[10] An Optimistic Differentiated Job scheduling system for cloud computing job scheduling is one the tasks performed in order to gain maximum profit. A web application is developed which provides authenticated users two types of service uploading and downloading respectively. Multiple user requests are processed by the use of non-pre-emptive priority algorithm.The service provider's main aim is to provide fast services to the multiple requests.

Ronak Patel et.al[5],Survey on Resource Allocation Strategies in cloud computing multiple cloud users can request number of cloud service simultaneously. Provision that all resources are made available to requesting user in efficient manner to satisfy their need. Dynamic resource allocation in cloud computing is shown based on topology(TARA),Linear Scheduling Strategy for Resource allocation and Dynamic Resource Allocation for Parallel Data Processing.

Vignesh V et.al[2],Resource management and scheduling algorithm job-oriented resource scheduling in a cloud computing environment. Resource management is scheduled for the process which gives the available resources based on user preferences. The computing resources can be assigned according to the rank of the job. Resource scheduling constructs the analysis of resource scheduling algorithm. These three algorithms are compared to analyse the time through Round robin, Pre-emptive Priority, Shortest Remaining Time First have been taken into consideration. From this analysis SRTF has been computed that lowest time parameters in all aspects and is the most efficient algorithm for resource scheduling.

4.PROPOSED WORK

In order to efficiently and effectively assigning computing resources scheduling becomes a very difficult task in a cloud computing environment where many alternative computers with different capacities are available. Efficient and Effective task scheduling mechanism can meets users requirements and improve the resource utilization

4.1 Round Robin

It is the simplest algorithm that uses the concept of time segments. Here the time is divided into multiple segments and each nub is given a particular time segment or time interval and in this segment, the nub will perform its operations. The resources of the service provider are provided to the client on the basis of this time segment. In round robin scheduling the time segment play vital role for scheduling ,because if the segment is very large then Round Robin scheduling algorithm is same as the FCFS scheduling .If the time segment is extremely too small then Round Robin scheduling is called as Processor scheduling algorithm and number of context switches are very high.It selects the load on random basis and leads to the situation where some nubs are heavily loaded and some are lightly

loaded.Though the algorithm is very simple,there is an additional load on the scheduler to decide the size of segment and it has longer average waiting time,higher context switches higher turnaround time and low throughput.The Round Robin algorithm mainly focuses on distributing the load equally to all the nubs.Using this algorithm,the scheduler assigns one VM to a nub in a cyclic manner.The Round Robin scheduling in the cloud is very sililar to the round robin scheduling is used in the process scheduling.The scheduler starts with a nub and moves on to the next nub,after a VM is assigned to that nub.This is repeated until all the nubs have been allocated atleast one VM and then the scheduler returns to the first nub again.Hence,in this case ,scheduler does not wait for the exhaustion of the resources of a nub before moving on to the next. As an example ,if there are three nubs and three VMs are to be scheduled ,each nub would be allocated one VM,provided all the nubs have enough available resources to run the VMs. The main advantage of this algorithm is that it utilizes all the resources in a balanced order. An equal number of VMs are allocated to all the nubs which ensure fairness. However, the major drawback of using this algorithm is that the power consumption will be high as many nubs will be kept turned-on for a long time. If three resources can be run on a single node, all the three nubs will be turned on when Round Robin is used which will consume a significant amount of power.

4. 2. Pre-emptive Priority

Priority of jobs is an important issue in scheduling because some jobs should be serviced earlier than other those jobs can't stay for a long time in a system. A suitable job scheduling algorithm must be considered priority of jobs. To address this problem some researchers have considered priority of jobs scheduling algorithm. Those researches have focused on a few criteria of jobs in scheduling. In cloud environments we always face a wide variety of attributes that should be considered. Priority is an important issue of job scheduling in cloud environments. In this paper we have proposed a priority based job scheduling algorithm which can be applied in cloud environments. Also we have provided a discussion about some issues related to the proposed algorithm such as complexity, consistency and finish time. Result of this paper indicates that the proposed algorithm has reasonable complexity. In addition, improving the proposed algorithm in order to gain less finish time is considered as future work. As a scheduling policy, preemption has wide applications in many areas (e.g. Process scheduling, bandwidth allocation, manufacturing scheduling). Most basically, preemption can be seen as a process that removes (un-schedules, suspends or aborts) one or more previously scheduled activities according to certain criteria and re-allocates freed resource capacity to a new activity. A preemption policy is normally used for scheduling high priority activities when a capacity shortage appears. Preemption has been investigated fairly extensively relative to scheduling single-capacity resources. CPU scheduling, which is central to operating system design, is preventative example, The CPU is single-capacity resource, which can be time-shared to accommodate multiple tasks by algorithms (such as round robin) that repeatedly allocate time slices to competing tasks. Here, a pre-emptive scheduling policy provides a means for reallocating time slices as new, more important jobs arrive for processing. Pre-emptive scheduling is much more complex in the context of cumulative or multi-capacity resources, and this problem has received much less attention in the literature. The principal complication concerns the selection of which activity (or activities) to pre-empt. In the case of multi-capacity resources, the number of candidate sets of activities increases exponentially with resource capacity size, while only a single activity must be identified in the single-capacity case.

4. 3 Shortest Response Time First

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 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. Instead, we have to check the event clock at each cycle, to see whether pre-emption is taking place, and if so, update the process data accordingly. The main thrust of the changes is to refine the process data to include a variable to count how much time to the next wait, and to the event module to explicitly cycle the clock, and see if an event is due at the current time. SJF policy selects the job with the shortest (expected) processing time first. Shorter jobs areal ways executed before long jobs. One major difficulty with SJF is the need to know or estimate the processing time of each job (can only predict the future!)Also, long running jobs may starve; because the CPU has a steady supply of short jobs.

5. PROPOSED ALGORITHM

5.1 Threshold based Priority Scheduling Algorithm

Threshold based priority scheduling algorithm:

J – Job process; Process Time- P_R ; J_n- Job event;

T_i - Time Process for Job pool based on

Job event ;

Event pool: P_n ;

Event pool P_1, P_2 -....Allocate the job for partial process, convert as a Clustering

To estimate the Process time or performance time (PSJF):

T_n - duration of n^{th} *CPU* Process time;

Q_{n+1} - predicted duration of the $(n+1)^{st}$

CPU Process time;

$Q_{n+1} = a * T_n + (1- a)Q_n$;

Turnaround time

Threshold Priority scheduling:

$P_n = 1; P_n + 1;$

for (P_n =1; P_n < n+1; P_n);

{ for (Jn = 1; Jn <10; Jn ++);

{ PSJF $_{TT}$ = waiting time + service time ; PSJF performance time = PSJF $_{TT}$ / service time;

Q_{n+1}= PSJF performance time ;

while (J_1; P_R, process time)

{ Proc = sqrt ($PSJF_1$ performance time - J_1 Process time);

Threshold value = Proc / J_1 Process Time;

Threshold Priority based on Job:

Threshold Priority $_{Job}$ (0.1 < Jn < Ti); // Second or Minutes of Duration for Job event

} $P_n = 1;$

Threshold Limit = Σ (Threshold value) ;

P_n *end ; }*

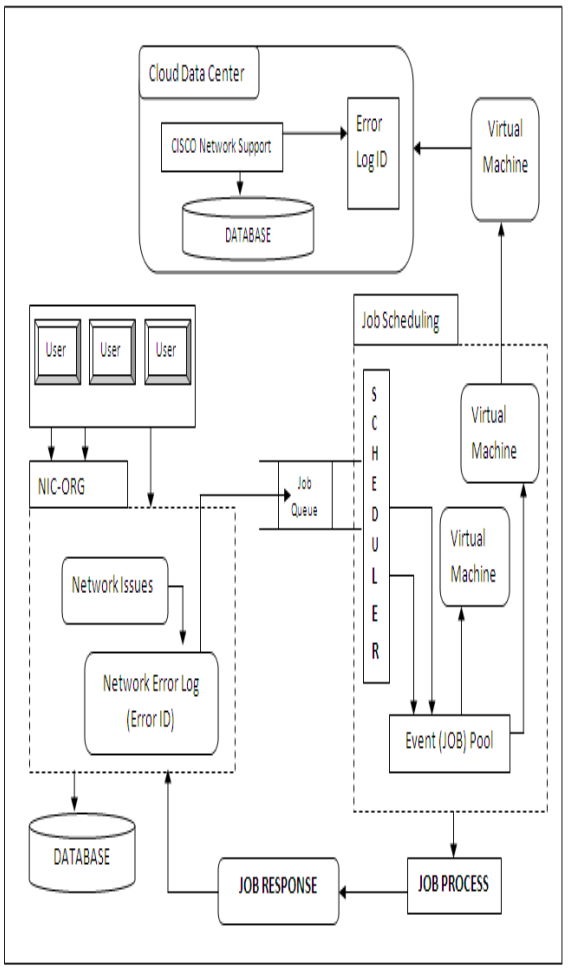


Fig :1.1 TBP Architecture

Threshold priority for Pool:

if (0.1 < Threshold Priority Pool < P_n +1)

{

P_n = Virtual Machine$_1$;

// Virtual machine Capable is less amount of

Job Event Process time.

}

else

{

P_n = Virtual Machine $_2$

// Virtual machine Capable is greater amount of

Job Event Process time.

}

Swap VM_1 =VM_2 ; // While the Job Process is completion.

VM_2 = VM_1 ;

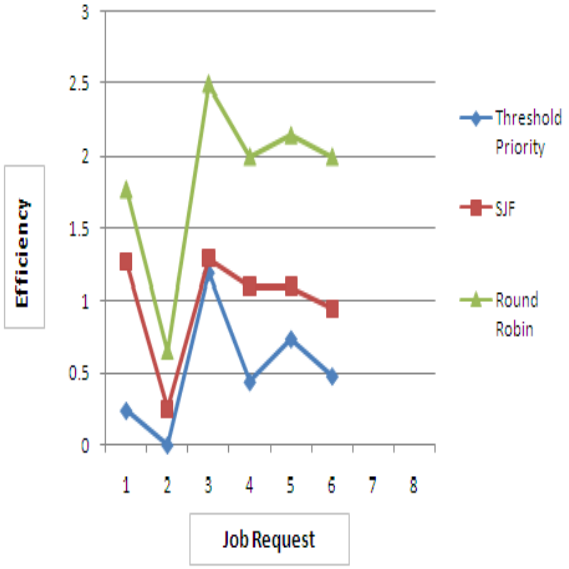
6. RESULT and ANALYSIS

Based on the observation from the following five request and based on their submission time the priority has been assigned .The priority has been assigned based on the submission time or type of job request.The execution time has been determined by the cloud provider.Based on the five job request undergone by the above algorithms we attain the above values that show table 1. For round robin the time quantum is assigned as 5.

Job Request ID	Job Queue	Job Process Time (Sec)	Job Turn Aroud Time (Sec)	Efficency
#J1	1	5	4.80	0.24
#J2	3	12	12.0	0.0
#J3	4	11	10.85	1.19
#J4	4	7	6.35	0.44
#J5	7	9	8.25	0.74
#J6	9	7	6.90	0.48

Table: 1.1 Table of Threshold Priority Efficency

Based on the above observation that have been put under the above scheduling algorithm we attain the average waiting and average turnaround time as show in graph 1.Eventually graph1 which show the total turnaround time and waiting time . After all the performance analysis we have been through shortest remaining time first has efficient performance compare with the round robin and pre-emptive priority Shortest job first and TBP . The Average and Total times have been calculated and presented for the three resource scheduling algorithms in Table 1.



. From the analysis, we can conclude that, Threshold based priority algorithm shows excellent performance and computes the resource scheduling with relatively less waiting time and turnaround time.

7. CONCLUSION

TBP scheduling ha significant improvement over algorithms in cloud environment. It minimizes Makespan and balance the load properly. Finally, it provides semi optimal or optimal solution. In future work, we can extend the TBP algorithm by security, adding deadline to the tasks and priority among the tasks. Communication cost between task and resource are ignored. It can be another future work. Tasks are non-pre-emptive in nature. It means once a resource is assigned, it cannot be removed before execution of the task is over. Pre-emption is effective, if machines are very close to each other. This study illustrates with some scheduling heuristics, performance metrics, comparisons and future work.

8.REFERENCES

[1].Hitoshi Matsumoto,Yutaka Ezaki,"Dynamic Resource management in cloud environment",July 2011,FUJITSU science & Tech journal,Volume 47,No:3,page no:270-276.

[2].Shalmali Ambike,Dipti Bhansali,jaee Kshirsagar,Juhi Bansiwal",An Optimistic Differentiated Job Scheduling System for Cloud Computing "International Journal of Engineering Research and Applications(IJERA) ISSN:2248-9622 Vol.2,Issue 2,Mar-Apr 2012,pp.1212-1212 1212.

[3].Vignesh V,Sendhil Kumar KS,Jaisankar N,"Resource Management and Scheduling in Cloud Environment",International Journal of Scientific and Research Publications, 1 ISSN 2250-3153 Vol. 3 Issue 6,June 2013.

[4].M.Sasitharagai,A.Padmashree,T.Dhanalakshmi &S.Gowri,"Dynamic Resource Management in Large Cloud Environments using Distributed Gossip Protocol",Intenational Journal of Internet Computing ISSN No:2231-6965,Vol.1,Issue 4,2012.

[5].Pinal Salot,Purnima Gandhi,"Job Resource Ratio Based Priority Driven Scheduling in Cloud Computing" IJSRD – International Journal for Scientific Research & Development Vol.1,Issue 2,2013 ISSN(online):2321-0613.

[6].Ronak Patel,Sanjay Patel,"Survey of Various Resource Allocation Strategies in Cloud Computing",International Journal of Engineering Research&Technology(IJERT) Vol.2,February 2013 ISSN:2278-0181.

[7].Venkatesh Kumar.V and S.Palaniswami",Dynamic Resource Allocation Method for Parallel Data Processing in Cloud Computing",2012,Journal of Computer Science 8(5),ISSN 1549-3636,Science Publications,page no:780-788.

[8].V.Vinothina,Dr.R.Sridaran,Dr.Padmavathi Ganapathi",A Survey on Resource Allocation Strategies in Cloud Computing",(IJACSA) International Journal of Advanced Computer Science and Applications,Vol.3,No.6,2012.

[9].Pinal Salot "A Survey of Various Scheduling Algorithm in Cloud Computing Environment",ISSN:2319-1163 Vol.2,Issue 2 131-135.

[10].Sujit Tilak,Prof.Dipti Patil",A Survey of Various Scheduling Algorithms in Cloud Computing Environment",International Journal of Engineering Inventions ISSN:2278-7461,Vol.1,Issue 2, Sep 2012,pp:36-39.