

A Survey on Several Job Scheduling Techniques in Cloud Computing

Dr. Karambir

Faculty of Computer Science and Engineering, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Kurukshetra University, 136119, Kurukshetra, Haryana, India.

Charul Chugh

M. Tech. (Computer Engineering), University Institute of Engineering and Technology, Kurukshetra University, 136119, Kurukshetra, Haryana, India.

Abstract – Cloud computing is one of the upcoming latest new computing paradigm where applications and data services are provided over the Internet. Cloud computing is attractive to business owners as it eliminates the requirement for users to plan ahead for provisioning, and allows enterprises to start from the small and increase resources only when there is a rise in service demand. While dealing with cloud computing, a number of issues are confronted like heavy load or traffic while computation. Job scheduling is one of the answers to these issues. The main advantage of job scheduling algorithm is to achieve a high performance computing and the best system throughput. Job scheduling is most important task in cloud computing environment because user have to pay for resources used based upon time. This paper aims to do the insight study of various existing job scheduling algorithms and proposes a hybrid job scheduling algorithm for enhancing the performance in cloud computing environment.

Index Terms – Cloud Computing, Job Scheduling in Cloud, Hybrid Job Scheduling.

1. INTRODUCTION

Cloud computing is a way of leveraging the Internet to consume software or other IT services on demand. Users share processing power, storage space, bandwidth, memory, and software. With cloud computing, the resources are shared and so are the costs. Users can pay as they go and only use what they need at any given time, keeping cost to the user down. Cloud computing is very much a business model as well.

Providers of cloud computing solutions, whether they are software, hardware, platform, or storage providers, deliver their offerings over the Internet. There are no shrink wrapped boxes containing discs or hardware for you to buy and set up yourself. Cloud providers typically charge monthly recurring fees based on your usage [1].

Cloud computing is the delivery of computing services over the Internet. Cloud services allow characters and businesses to use software and hardware that are managed by third parties at remote locations. Specimens of cloud services include online file storage, social networking sites, webmail, and online

business presentations [2].

A. Advantages of Cloud Computing

- Cloud computing makes it easier for enterprises to scale their services – which are increasingly reliant on accurate information – according to client demand. Since the computing resources are managed through software, they can be deployed very fast as new requirements arise [3] [23].
- It dramatically lowers the cost of entry for smaller firms trying to benefit from compute-intensive business analytics that were hither to available only to the largest of corporations. These computational exercises typically involve large amounts of computing power for relatively short amounts of time, and cloud computing makes such dynamic provisioning of resources possible [3] [23].

B. Essential characteristics of Cloud Computing

The essential characteristics of cloud computing:

On-demand self-service: Registering could resources be gained and utilized whenever without the requirement for human association with cloud administration suppliers. Computing resources include processing power, storage, virtual machines, etc [4] [5].

Resource pooling: Cloud administration suppliers pool their resources that are then imparted by numerous clients. This is alluded to as multi-tenure where, for instance, a physical server may have a few virtual machines having a place with distinctive clients [4] [5].

Rapid elasticity: A client can rapidly gain more resources from cloud by scaling out and can scale back in by discharging those resources once they are no more needed [4] [5].

Broad network access: The beforehand said resources could be gotten to over a system utilizing heterogeneous gadgets, for example, laptops or mobiles telephones [4] [5].

C. Job Scheduling in Cloud Computing

The job scheduling algorithm controls the order in which each task will run and also controls the allocation of resources. The Scheduler in which order the jobs are scheduled have significant effect on system performance. The Job management is the fundamental concept of cloud computing systems job scheduling issues are identifies with the effectiveness of the entire cloud computing framework. Job scheduling will be a mapping component from client assignments to the proper determination of assets and its execution [6]. Job scheduling is adaptable and helpful. Jobs and job streams can be planned to run at whatever point required, taking into account business capacities, needs, and requirements. Job torrents and procedures can set up every day, week after week, month to month, and yearly ahead of time [7]. The Need of Job Scheduling in Clod Computing is explained as:

Best Running time:

Jobs can be divided into different categories according to the needs of users, and then set the best running time on the basis of different goals for each task. It will improve the QoS of task scheduling indirectly in a cloud environment [8].

Enhance Throughput:

Mainly for cloud computing systems, throughput is a measure of system task scheduling optimizing performance, and it is also a target which has to be considered in business model development. Increase throughput for users and cloud providers would be benefit for them both next is advantageous because of its simplicity and because it minimizes the average amount of time each process has to wait until its execution is complete. It is one of the simplest scheduling algorithms we have it allocate the CPU in the order in which the process arrive [8].

Quality of Service –

The cloud is mainly to provide users with computing and cloud storage services, resource demand for users and resources supplied by provider to the users in such a way so that quality of service can be achieved. When job scheduling management comes to job allocation, it is necessary to guarantees about QoS of resources [8].

2. LITERATURE SURVEY

Dugki Min et.al, 1995 [24] In this paper, author described the arrangements have tried to attain good job response times with little system fragmentation of processing incomes. Since most schemes hazy focussed on approaches for mesh computer allocation, the schemes have used First-Come-First-Serve as the job scheduling discipline. Though, it has been beforehand reputable that job scheduling procedures for parallel and distributed computing organizations can have a large bearing

on the system consumption and job response time. Systems that use multiple queues, which reorder the classification of jobs allotted to the parallel system, can be very effective in educating the system performance.

Manoj Kumar Mishra et.al, 2011 [12] analyzed and exploration on job scheduling can be carried out to increase the performance of assemblage based scheduling algorithm in grid computing. This study means to achieve better performance by ranging the thought of group based job scheduling. Therefore, this paper proposed “A modified combination based job scheduling in computational grid” with the autonomous of reducing overhead time and computation time, thus dropping overall processing time of jobs.

Saad Bani-Mohammad et.al, 2012 [2] In this paper, author described the concert of non-contiguous provision can be knowingly affected by the job scheduling approach used for determining the order in which jobs are particular for execution. In this paper, the routine of the well-known Greedy Offered Busy List non-contiguous apportionment strategy for 2D mesh-connected multi computers is re-visited considering several significant job scheduling strategies. These were the First-Come-First-Served, Out-of-Order, and Window-Based job scheduling strategies. They were likened using detailed flit-level imitations. General simulation consequences based on synthetic and real assignment models indicate that the Window Based job scheduling approach exhibits good presentation when the scheduling window size was large and weighty system masses.

PinkyRosemarry et.al, 2012 [1] Grid computing resolved the multifaceted compute troubles between manifold. Grid computing resolved the great scale computational load in a elevated arrangement computing atmosphere. The main meaning in the grid computing is given to the reserve association & the job scheduler. The aim of the job scheduler was to exploit the supply operation & decrease the dispensation time of the jobs. Breathing come up to of grid scheduling had not gave a great deal significance on the presentation of a grid scheduler in dispensation time stricture. Schedulers assigned property to the occupation to be performing using the first come first serve algorithm. In this paper, author provided an optimized algorithm to row of the scheduler by a assortment of scheduling methods similar to direct job first, first in first out, round robin.

Wei Tang et.al, 2012 [13] proposed an adaptive metric-aware job scheduling approach. First, they proposed metric aware scheduling which enables the scheduler to stability challenging scheduling goals characterized by different metrics such as job waiting time, fairness, and system utilization. Second, they improved the scheduler to adaptively adjust scheduling strategies based on comment information of monitored metrics at runtime.

VaishaliChahar et.al, 2013 [22] Multilevel Queuing & Multilevel Feedback Queuing was common in the CPU scheduling methods used in operating systems. These methods were common but still have some subjects & have a wide scope of development. Very less literature review was near on these scheduling techniques. So in this paper dissimilar methods for scheduling these techniques future by different authors had been composed & discussed

Chengkuan ZENG et.al, 2013 [14] purposed of this investigation is to clarify and enhance the job shop scheduling problematic below the environment of mixture operations of machining. The machining operations of the jobs could not be ended in own plant and thus inter-cell productions were considered. They clarify the features of the above job shop scheduling difficult in detail and propose a new optimization criterion which accords with fact problem.

M. Sharkh et.al, 2013 [10] explained that cloud computing is a utility computing paradigm which provides various service portfolios that differ in resource configurations and provided services. Authors described Resource allocation algorithms to accomplish the task of scheduling virtual machines on the servers residing in data centers and as a result scheduling network resources while complying with the problem limitations.

Rajveer Kaur et.al, 2014 [8] Cloud computing is rich day by day & it will carry on in just beginning phase until computer & internet era is in endurance. While commerce with cloud computing, a numeral of matter are face up to like weighty load or transfer while computation. Job scheduling was single of the answer to these matter. It was the process of map job to available resource. In this research is one converses about cloud computing & scheduling. It gave details about job scheduling in cloud computing. Its accessible algorithms for job scheduling were discussed, existing algorithms were compared & lastly conclusion & future work were discussed

S.K. Aparnaa et.al, 2014 [9] considered the memory requirement of each collection, which is one of the main reserves for scheduling data exhaustive jobs. Due to this the job disappointment rate is also very high. To deliver a solution to that problem Greater Adaptive Scoring Job Scheduling algorithm is familiarized. The jobs were recognized whether it is data concentrated or computational intensive and based on that the jobs were scheduled. The jobs were allocated by computing Job Score along with the memory condition of each cluster. Due to the vibrant nature of grid environment, each time the position of the resource changes and each time the Job Score was calculated and the jobs were allocated to the most suitable resources.

Chunsheng Zhu et.al, 2014 [15] described as, first analyzed the characteristics of job scheduling with respect to CC-WSN integration and then studies two job scheduling procedures .

Further, two novel job scheduling algorithms, namely priority-based two phase Min-Min and priority-based two phase Max-Min, are proposed for CC integrated with WSN

Chen, Huangning and Wenzhong Guo et al, 2015 [11] In this paper, a soft real-time task scheduling algorithm proposed based on particle swarm optimization approach for cloud computing. Simulation results show the proposed algorithm can effectively minimize deadline missing ratio, maximize the profit of cloud service provider and achieve better load balancing compared with baseline algorithms..

Dazhao Cheng et.al, 2015 [16] In this paper, they proposed, a Resource and Deadline-aware Hadoop job Scheduler that takes coming resource availability into consideration when minimizing job deadline misses. They formulated the job scheduling problem as an online optimization problem and solve it using an efficient receding horizon control algorithm. To aid the control, they design a self-learning model to estimate job completion times and use a simple but effective model to predict future resource availability.

Abhishek Gupta et.al, 2015 [17] intended to present the performance comparison analysis of various pre-existing job scheduling algorithms considering various parameters. This paper discussed about cloud computing and job scheduling concept in cloud computing had been elaborated. Further existing algorithms for job scheduling are discussed, and were compared in a tabulated form with respect to various parameters.

3. EXISTING JOB SCHEDULING TECHNIQUES

FCFS (First Come First Serve) Scheduling Algorithm:

It is one of the simplest Scheduling algorithms we have it allocate the CPU in the order in which the process arrive. It assumed that ready queue is managed as first in first out which means that the first job will be processed first without other preferences [18].

Algorithm FCFS:

- Initialize Tasks.
- First task assigned to the queue and add tasks up to n numbers.
- Add next task 'I' at last position in the main Queue.

Shortest Job First Scheduling Algorithm:

Shortest job First (SJF) also known as Shortest Job Next (SJN) or Shortest Process Next (SPN) is a scheduling technique that selects the job with the smallest execution time [19]. The jobs are queued with the smallest execution time placed first and the job with the longest execution time placed last and given the lowest priority. This Scheduling algorithm is deal with different approach in this algorithm CPU is allocated to the process with least burst time.

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 [18].

Round Robin Scheduling 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. So the basic idea behind Round Robin is that in this algorithm, time slices are assigned to each process in equal portions and in circular order.

In [20] the round robin algorithm is used with the multiple feedback scheduling algorithm to reduce the response time of jobs and improve the throughput. This Hybrid Round Robin Approach is the average response time of jobs can be reduced, but not the throughput for the cluster. For the round robin algorithm with single queue, if the size of jobs in the cluster is equal, for each job, each job will get the slots every fixed cycle. Assume that each job will get one slot for every round, and the number of jobs in the cluster is N , the running time of each task if t , so the average turnover time is $N*t$. We can see that when the number of jobs in the cluster get larger, the turnover time will also become larger, and the same as the waiting time for each job. The number of the jobs in the cluster can be very large, at this situation; the completion time of each job can be very long, especially for the jobs in the tail. And the throughput of the cluster will be brought down as the number of jobs grows. So we use multiple ready queues to hold all the jobs in the cluster. When the multiple queues with different priorities are applied, the jobs will be transferred from one queue to the other.

Priority Scheduling Algorithm:

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 [21].

Multilevel Feedback Queue Technique:

Multilevel feedback queue scheduling is similar to the ordinary multilevel queue scheduling described above, except

jobs may be moved from one queue to another for a variety of reasons like if the characteristics of a job change between CPU-intensive and I/O intensive, then it may be appropriate to switch a job from one queue to another. Aging can also be incorporated, so that a job that has waited for a long time can get bumped up into a higher priority queue for a while. It uses several complete queues & connections a dissimilar main concern with every queue. The Algorithm decides to procedure with uppermost priority from the unavailable queue & run that procedure either preemptive or non-preemptively. After execution into the first queue the processes either gets completed or they are scheduled to run into the second queue. But before shifting to the next lower queue their remaining CPU burst and waiting time is calculated. In second queue they are again scheduled with Round Robin CPU scheduling technique with a suitable time quantum value and again these parameters are updated and the process go as on [22].

4. PROPOSED SCHEDULING TECHNIQUE

The proposed algorithm is a hybrid of two techniques known as Multi Level Scheduling Algorithm with P-Thread algorithm. The objective of using the P threads scheduling feature is we can designate how threads share the available processing power. We may decide that all threads should have equal access to all available CPUs, or you can give some threads preferential treatment. In some applications, it's beneficial to give those threads that perform important tasks an advantage over those that perform background work.

For instance, in a process control application, a thread that responds to input for special devices could be given priority over a thread that simply maintains the log. Both of them are get hybrid in such a way as to increase the performance in cloud network. The overall working is divided into two different phases.

The proposed hybrid approach working can be summarized in two phases as in the first phase with the help of multi queue approach all the possible queues are managed. The execution of all the process in the different queues is being handled. In the second phase handling of jobs are carried out in the different job queues. The P Thread algorithm will be responsible for this handling because this algorithm aims in reducing the job execution time by executing all the jobs in the queue with multi-threading technique. The proposed approach will aim to decline the waiting time for cloud user jobs so that throughput of the overall system can be magnified.

5. CONCLUSION AND FUTURE WORK

Cloud computing has fuelled concerns on a fundamental point for the success of information systems, communication, virtualization, data availability and integrity, public auditing, scientific application, and information security. A better scheduling method definitely has a higher potential to enhance the performance of most parts of a cloud computing system.

Job scheduling algorithms is one of the most demanding area in the cloud computing domain because it solve major issues in this environment. In this paper efforts have been put to study various scheduling algorithms in cloud environment. The study which has been done may be useful for development of efficient proposed hybrid scheduling algorithm so as to enhance the overall efficiency of cloud computing environments in future. In future, the work can be extended to implement the proposed approach in a real time cloud environment to enhance cloud environment performance.

REFERENCES

- [1] Pinky Rosemary, Ravinder Singh (2012), "Grouping Based Job Scheduling Algorithm using Priority Queue and Hybrid Algorithm in Grid Company", International Journal of Grid computing and Applications (IJGA).
- [2] Bani-Mohammad Saad (2012), "On the performance of job scheduling for noncontiguous allocation in 2D mesh- connected multicomputers", IEEE Electro technical Conference (MELECON).
- [3] Avram, M. G(2014), " Advantages and Challenges of Adopting Cloud Computing from Enterprise Perspective" Procedia Technology, vol.12,Issue(1), pp. 529-534.
- [4] Puthal D, B. Sahoo, S. Mishra, and S. Swain(2015), "Cloud Computing Features, Issues, and Challenges: a Big Picture", IEEE International Conference on Computational Intelligence & Networks.
- [5] Sasikala(2013), "Research Challenges and Potential Green Technological Applications in Cloud Computing" International Journal of Cloud Computing, pp. 1-19.
- [6] Tang Wei, Dongxu Ren, Zhiling Lan, and Narayan Desai(2012), "Adaptive Metric-Aware Job Scheduling For Production Supercomputers" 41st IEEE International Conference on Parallel Processing Workshops (ICPPW), pp. 107-115.
- [7] Zeng, Chengkuan, Jiafu Tang, and Huabo Zhu(2013), "Two heuristic algorithms of job scheduling problem with inter- cell production mode in hybrid operations of machining." 25th Chinese Control and Decision Conference (CCDC).
- [8] Rajveer Kaur and Supriya Kinger(2014), "Analysis of Job Scheduling Algorithms in Cloud Computing", International Journal of Computer Trends and Technology (IJCTT), Vol. 9 , Issue 7 .
- [9] Aparnaa, S. K., and K. Kousalya(2014), "An Enhanced Adaptive Scoring Job Scheduling Algorithm For Minimizing Job Failure In Heterogeneous Grid Network". IEEE International Conference on Recent Trends in Information Technology (ICRTIT).
- [10] M. Sharkh, M. Jammal, A. Shami and A. Ouda(2013), "Resource Allocation In A Network-Based Cloud Computing Environment: Design Challenges", IEEE Communications Magazine, Vol. 51, pp. 46-52. .
- [11] Chen, Huangning, and Wenzhong Guo(2015), "Real-Time Task Scheduling Algorithm for Cloud Computing Based on Particle Swarm Optimization." in Cloud Computing and Big Data, , Springer International Publishing,pp. 141-152.
- [12] Mishra, Manoj Kumar, Prithviraj Mohanty, and G. B. Mund(2011), "A Modified Grouping-Based Job Scheduling in Computational Grid" IEEE International Conference (NUICONE).
- [13] Tang Wei(2012), "Adaptive metric-aware job scheduling for production supercomputers." , Parallel Processing Workshops (ICPPW), 41st IEEE International Conference .
- [14] Zeng Chengkuan, Jiafu Tang, and Huabo Zhu(2013), "Two Heuristic Algorithms of Job Scheduling Problem with Inter Cell Production Mode in Hybrid Operations of Machining" IEEE Control and Decision Conference (CCDC).
- [15] Zhu Chunsheng(2014), "Job Scheduling for Cloud Computing Integrated with Wireless Sensor Network", IEEE 6th International Conference on Cloud Computing Technology and Science (Cloud Com).
- [16] Cheng Dazhao(2015), "Resource And Deadline-Aware Job Scheduling In Dynamic Hadoop Clusters" IEEE International Parallel and Distributed Processing Symposium (IPDPS).
- [17] Abhishek Gupta(2015), "A Theoretical Comparison Of Job Scheduling Algorithms In Cloud Computing Environment, "IEEE International Conference on Next Generation Computing Technologies (NGCT).
- [18] Lu Huang, Hai-shan Chen and Ting-ting Hu(2013), "Survey on Resource Allocation Policy and Job Scheduling Algorithms of Cloud Computing" ,Journal of Software, Vol. 8, Issue 2, pp. 480-487.
- [19] Van den Bossche R, Vanmechelen K, Broeckhove J(2010), "Cost Optimal Scheduling in Hybrid IaaS Clouds for Deadline Constrained Workloads", 3rd IEEE International Conference on Cloud Computing, Miami.
- [20] R. Rao, and Y. Wang(2014), "A Round Robin With Multiple Feedback Job Scheduler In Hadoop" IEEE International Conference on Progress in Informatics and Computing, pp. 471-475.
- [21] Vaishali Chahar(2013), "A Review of Multilevel Queue and Multilevel Feedback Queue Scheduling Techniques" IEEE International Journal of Advanced Research in Computer Science and Software Engineering.
- [22] Anca Apostu, Florina Puican, Geanina Ularu, George Suciu, and Gyorgy Todoran(2014), Study on advantages and disadvantages of Cloud Computing – the Advantages of Telemetry Applications in the Cloud" Recent Advances in Applied Computer Science and Digital Services.
- [23] Min Dugki, and Matt W. Mutka(1995), "Efficient job scheduling in a mesh multicomputer without discrimination against large jobs." Parallel and Distributed Processing, Seventh IEEE Symposium.