

# Survey on Job Scheduling Algorithms in Grid Computing

Seema Kumari  
Student  
Lovely Professional University  
India

Gulshan Kumar  
Assistant Professor  
Lovely Professional University  
India

## ABSTRACT

Grid computing is a type of distributed computing that distributes the tasks to a group of network computers where computer are linked together to provide a better computational resource. Scheduling plays a vital role in gaining high consumption of resources in the field of grid computing and achieving utilization of resources in grid environments. The use of the scheduling is to increase the throughput, efficiency and decrease the average waiting time. In this paper we have been surveyed various job scheduling algorithm in grid computing that will benefit for the researchers to carry out the future work in that area and develop a better algorithm. It is found that the resource allocation and job scheduling plays a vital role to improve performance and can proficiently and efficiently allocate workloads to proper resources.

## Keywords

Grid Computing, Job Scheduling, Virtualization, Resource Scheduling.

## 1. INTRODUCTION

Grid computing is a pool of computational no. of resources, resources are different type of computers that is interconnected with a network to solve computational problem. There are being very used in data mining for a large no. of research-based experiments. The scientists uses the grid computing for their research work. The grid computing is to useful to makes a different type of machines that are located in different locations that may be in different physical locations. Clusters is also helpful to increase the available resources at physical level but at the different location to work together. They needs peer to peer communication and internet to allow the clusters and to control the multiple clusters at different locations that you requires a single Grid scheduling process where that process can use peer to peer communication tool .In cluster calculation have been done by CPU. This is mostly used to improve the performance of the resources.

The main purpose of Grid computing is to distribute the computational power, storage data and network resource to solve the particular problem across dynamic and geographically organizations. Grid is a dispersed system that enables the distribution, selection, and geologically dispersed resources enthusiastically at runtime i.e. depending on their usability, ability, execution and quality of service according to the user's specification. The capability of current computer and processor networks has amplified exponentially as compared with customary mainframes. This enhancement in their performance, in general terms wastage of computational grid resources because mostly the central processing unit sits free. Grid utilizes this free central processing unit cycles to processing the computation when it is requested by the grid

applicant. Grid computing makes an optimized utilization of the resources the CPU cycles otherwise would have wasted.

According to this the users can get greater computation resource and can process their large-scale computational issues thus solving a complex problem to the computational level of a supercomputer. A scheduler plays a very important role in grid computing. Scheduler is helpful for managing the jobs and available resources in grid computing.

The main motive of the scheduler should be maximize utilization of the resources and minimize the time of the job execution. The scheduler select suitable resources for job to complete its execution.



Figure1 Grid Computing

Grid scheduling are mainly classified on the basis of two types:

1. Static scheduling: In this scheduling the jobs are assigned to the available resources before execution and continue the execution without interruption for these resources.
2. Dynamic Scheduling: In dynamic scheduling rescheduling is allowed and executing job can be transferred according to their dynamic status about the workload of the resources.

## 1.1 Scheduling Fundamentals

Scheduling is a process that manages the execution of inter-dependent tasks on distributed resources. It launch the assigning suitable resources to workflow tasks so that the execution can be finished to fulfil target objective functions that is pointed out by users. When an activity should start and end, depending on its (1) duration, (2) Processing-time, (3) Start-time (4) Finish-time (5) Expected execution time

## 2. CURRENTLY DISTRIBUTED ENVIRONMENTS

1. *Cluster Computing:* A cluster computer is a type of computing that is a group of tightly coupled computers works teamwise so that it can be viewed as a single computer.
2. *Grid Computing:* Grid computing (or the use of a *computational grid*) is applying the resources of multiple computers on a network to a single problem at the identical time - usually to a technical problem that needs a great number of computer processing cycles or entrance to large amounts of data.
3. *Cloud Computing:* Cloud computing is a form of computing which means storing and access your data on the Internet instead of client's hard disk.

## 3. GRID ARCHITECTURE

To understand the grid computing we have to understand the grid layers and architecture. In architecture user wants to do some kind of application for processing. The grid user send the application for processing to the grid but user can't interact with grid system. So this application is send to the grid broker and broker goes to GIS (Grid Information System) i.e. a kind of directory information of all the resources .Here grid broker knows about the resources and select the suitable resources for the processing of application. Grid broker send all the data to the resources, resources process the data and send back result to the broker. The broker send back result to the grid user

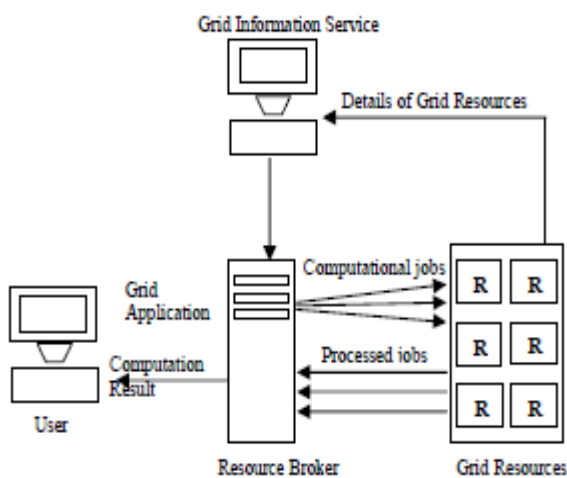


Figure 2: Grid Computing Architecture

## 4. TECHNIQUES AND SCHEDULING ALGORITHM USED IN GRID COMPUTING

In this section we have been surveyed various scheduling algorithm in grid computing.

### 4.1 Job Grouping Strategy in Grid Computing

Job grouping strategy [7] is used to improve that increase the utilization of the resource and decrease the processing time of jobs in grid computing environment. The resource and job scheduling model is divided into three levels user level, global level and local level (cluster level).Grid computing is

a combination of different clusters and a combination of nodes is called as cluster. In this model user enters the jobs for performing the grid system at user level and scheduler collects the information about the all available resources from GIS , then the scheduler distribute the jobs for processing to the selected suitable resources. Now when jobs are entered at global level in grid then first of all it search that cluster which have higher computational power. If the selected cluster have highest computational power than that of the submitted jobs then next job enters and now compare the total computational power with computational power of the selected cluster. If requirement of the group of jobs is less than or equal to of selected cluster then jobs are submitted to higher computational power of the cluster. So global level makes the load balancing across the cluster and local level. The local scheduler is used to maximize the utilization of the grid resources adequately and decrease the processing time of the jobs.

### 4.2 Adaptive Scheduling

Adaptive algorithm proves a genetic algorithm that is based on evolutionary algorithm (G-EA).G-EA provides an algorithm for solving the scheduling problem for reducing the job execution time in grid computing. G-EA as compared with Round-Robin, Min-Min scheduling algorithms and perform better throughput. In adaptive scheduling the parameters that is used to make a scheduling decisions change dynamically according to the previous, resource status in grid computing.

### 4.3 Firefly Algorithm

The Firefly algorithm [12] is a meta heuristic algorithm which is focused for find the global optimal solution to complete the allocated job within minimum makespan and flow time. It propagates based on intensity level which have higher knowledge that more glow. All firefly to move towards each other on basis on intensity level and the amount of movement depends upon attractiveness parameter which is further depend on distance between them. It focus to the efficiently utilization of the processor in the grid. In firefly algorithm first of all the No. of available resources ,no. of jobs allocated and parameters are identified .after then it found the brightness of each firefly using fitness function then calculate the distance and light intensity is updated. Now the current global best is recognized and fireflies are ranked.

### 4.4 Ant Colony Optimization Algorithm (ACO)

The ant colony optimization algorithm [4] is dependent on the behavior on ants moving in search of food. Each ants share their experience by releasing a chemical called as pheromone. This pheromone level evaporates moves with time and ants follow the patch in which there is maximum pheromone level. The pheromone is helpful to find the at least way and the value of pheromone of each resource is depends on their job processing time.. The higher processing time of the resource will receive a large no. of pheromone. When a resource is successfully done its job then its higher value of the pheromone will be added each time for processing of the next job .If a resource fails to complete its job execution then it will be penalty by adding lower pheromone value. The problem is that where the possibility of jobs are assigned to same resources having high pheromone value .The algorithm is follows as:-

(1) The applicant will forward a job for execution (2) The resource broker will accept the request sent by the applicant and find an appropriate resource based on higher value in pheromone for the job (3) When job is assigned to the resource then the local pheromone update is done. (4) When a resource is successful to complete its job then the global pheromone is updated. (5) After then the result will be sent to the client.

#### 4.5 Heuristic Algorithm for Job Scheduling

Heuristic Algorithm for Job Scheduling [1] is a type of computational method that modifies a problem with the help of iterations trying to improve a performance of the individual. This heuristic formulates the fitness of each character and based on their fitness character are added in the list of optimization. The new population is examined for next hop or iteration of algorithm. When acceptable fitness level is reached then algorithm stops.

#### 4.6 Prioritized Deadline Scheduling Algorithm (PDSA)

In Prioritized Deadline based scheduling [5] algorithm most of the researchers have not completely considered the deadline perspective for job processing. But PDSA generate a new concept that meet the deadline as indicated by users and users are extremely to processed the no. of jobs in appropriate way. In this algorithm input is determined by its deadline, arrival time and burst time. After then calculate the time delay in increasing order for per job and then select the job for the processing on the basis of least time delay. The performance of the PDSA as comparison to the previous algorithm like EDF scheduling algorithm and RR scheduling algorithm is more efficient. and greatest tardiness.

#### 4.7 Load Balancing Technique

In this Load Balancing Technique [11] the scheduler gets a job which is allocated by a client, it will send a request to the Information Service to attain the compulsory information like the idleness of the Central processing unit percentage of each resource, clusters average load and system's average load. After the section of the cluster, it will evaluate the average load of the selected cluster with the average load of the organization. The load average of the cluster is defined by the average load of each resource in particular cluster. Weighted sum of the square method for the measure of the load is used.

#### 4.8 Task Scheduling in Job Scheduling

In this researcher proposed technique for task scheduling [2] in grid computing, which is an enhanced task scheduling algorithm using triangle module operator to integrate Quality of Service of a user with system's based performance of the computations? The result are simulated on Grid Simulator, and the results powerfully show the proposed algorithm performs better, resource load balancing and Quality of Service effectiveness than minimum heuristic which does not regarded as good Quality of service.

#### 4.9 (MFQ) algorithm

In MFQ algorithm provides [10] the pre-emptive and non-pre-emptive algorithms can be used together in Multi-Level queue scheduling and it divides the ready queue into a number of queues with different scheduling policies. The proposed MFQ algorithm in multilevel scheduling, with the help of simulator for various group of jobs and scheduling policies, In this algorithm preferred RR for the fair use of

CPU in 1<sup>st</sup> queue and SJFRR in next queues (which we called as SJFRR) to reduce the average waiting time and turned around time.

#### 4.10 Non-preemptive Priority (NPRP)

The NPRP algorithm [3] is combination of the Backfill algorithm and Earliest Deadline algorithm. Job Scheduling and Resource Management are the fundamental issues in Grid Computing. This algorithm is integrated with Care Resource Broker (CRB) that supports virtualized. CRB helps to entire decisions to prioritizing the jobs and reducing the turnaround time, waiting time and response time of the jobs i.e available in the queue and the user's quality of service has enhance to make as large the virtualized grid environment.

#### 4.11 Multi-Agent Brokering Approach and Firefly Algorithm

Multi-agent brokering approach [14] proposed algorithm is based on firefly algorithm. The main work of the firefly algorithm is helpful to reduce the makespan time. Now in this algorithm addition jump option was introduced to jump to new situation to arrive high probability i.e called as a jumper firefly and it is more efficient than that of other heuristic algorithm and firefly algorithm.

### 5. ACKNOWLEDGMENT

I might want to express my deepest much obliged and gratefulness to my guide Mr. Gulshan Kumar for this excellent guidance, and who provided for me the brilliant chance to compose this paper on the topic Survey on Job scheduling algorithms in grid computing. I am sincerely thankful to him for sharing his truthful perspectives views and I came across large number of new ideas related to the research.

### 6. CONCLUSION

Grid computing having a wide area of application like as medical, science, research and others. In there so many challenges arise in grid computing i.e job scheduling and resource scheduling are the main in grid computing. In this paper presents such type of algorithm which is useful for the scheduling computational resource to the jobs in appropriate way. There are a lot of algorithm have been developed to improve the throughput, maximum utilization of the resource, reduce processing time and execute more no. of the jobs on the sever. It will be of consideration how a most of the researchers would discover new techniques to develop grid. In future the researchers can take a idea and use these facts to develop a best algorithm as comparison of them. Our research work aims to produce a more efficient scheduling algorithm to reduce effort by the grid in the procedure to execution of jobs.

### 7. REFERENCES

- [1] Wang, Jing, Gongqing Wu, Bin Zhang, and Xuegang Hu. "A heuristic algorithm for scheduling on grid computing environment." In China Grid Annual Conference (ChinaGrid), 2012 Seventh, pp. 36-42. IEEE, 2012.
- [2] Wu, Run-ze, Ming-shan Wu, Xiang Mi, and Qi An "Task Scheduling Algorithm Based on Triangle Module in Grid Computing." In Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on, pp. 1-4. IEEE, 2012.
- [3] G.Kannan and Dr.S. Thamarai Selvil "Non-preemptive Priority (NPRP) based Job Scheduling Model for

- Virtualized Grid Environment” 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE) V4-377-381 IEEE, 2010
- [4] Ku-Mahamud, Ku Ruhana, and Husna Jamal Abdul Nasir. "Ant Colony Algorithm for job scheduling in grid computing." *Mathematical/Analytical Modelling and Computer Simulation (AMS)*, 2010 Fourth Asia International Conference on. IEEE, 2010.
- [5] Abba, Haruna Ahmed, et al. "Deadline Based Performance Evaluation of Job Scheduling Algorithms." *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, 2012 International Conference on. IEEE, 2012.
- [6] Raksha Sharma, Vishnu Kant Soni, Manoj Kumar Mishra and Prachet Bhuyan, "A Survey of Job Scheduling and Resource Management in Grid Computing", in *World Academy of Science, Engineering and Technology* 40, 2010.
- [7] Raksha Sharma, Vishnu Kant Soni, Manoj Kumar Mishra "An Improved Resource Scheduling Approach Using Job Grouping strategy in Grid Computing" 2010 International Conference on Educational and Network Technology (ICENT 2010) 2010 IEEE pp 94-96
- [8] S. Gomathi and Dr.D.Manimegalai, "An adaptive grouping based Job Scheduling in Grid Computing", 2011 International Conference on Signal-Processing, Communication, Computing and Networking Technologies (ICSCCN 2011), IEEE 2011.
- [9] Zhang, Yu, Shuwei Chen, and Ziqian Hu. "A Scheduling Algorithm for Many-Task Computing Optimized for IO Contention in Heterogeneous Grid Environment." In *Computational and Information Sciences (ICCIS)*, 2013 Fifth International Conference on, pp. 1541-1544. IEEE, 2013.
- [10] Dr. Sanjay. K. Dwivedi and Ritesh Gupta " Simulator based performance analysis of Multilevel Feedback Queue Scheduling" 2014 5th International Conference on Computer and Communication Technology 2014 IEEE
- [11] Manimala, R., and P.Suresh. "Load balanced job scheduling approach for grid environment." In *Information Communication and Embedded Systems(ICICES)*, 2013 International Conference on, pp. 336-339. IEEE, 2013.
- [12] Adil Yousif, Abdul Hanan Abdullah, Sulaiman MohdNor and Adil Ali Abdelziz, "Scheduling algorithm on Grid Computing Using Firefly Algorithm", in *Journal of Theoretical and Applied Information Technology* 2011. Vol. 33 No.2.
- [13] Abba, Haruna Ahmed, Nordin B. Zakaria, and Anindya J. Pal. "Comparative Analysis of Some Grid Scheduling Algorithms on Real Computational Grid Environment." (2012)
- [14] Prajapati, Harshadkumar B., and Vipul A. Shah. "Scheduling in Grid Computing Environment." *Advanced Computing & Communication Technologies (ACCT)*, 2014 Fourth International Conference on. IEEE, 2014.
- [15] Nithya, G., and R. M. S. Engels. "Multi-agent brokering approach and firefly algorithm for job scheduling in grid environment." *Electronics and Communication Systems(ICECS)*, 2014 International Conference on. IEEE, 2014.