International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 5 | ISSN : 2456-3307

An Analysis of different scheduling approaches used in Grid Computing

Aman Baria[1] , Dharam Singh Meena[2] , Raksha Pandey[3]
[1,2] B.Tech Student, Department of Computer Science and Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya ,Bilaspur(C.G), India
[3] Assistant Professor , Department of Computer Science and Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya ,Bilaspur(C.G), India

ABSTRACT

Grid computing is considered as a very high performance computing environment to solve complex computational problems. It comprises of job scheduling, security, resource management, information management. The scheduling concept is needed for better management of resources. Scheduling refers to order of jobs which satisfy the metrics like satisfaction of user, completion time etc. Better scheduling policy in grid computing helps in improving performance, computational cost, load balancing and increases reliability and availability of resources. This paper end up with brief description of existing scheduling algorithms by considering different metrics used to prove importance of existing techniques.

Keywords : Grid Computing; Load Balancing; Makespan; Resource Utilization; Scheduling; User Deadline; Latency; Cost.

I. INTRODUCTION

Grid is a type of wide scale distributed computing of different resources which are related to different organizations. These resources are selected based on their cost, usage, availability etc. [8].Grid systems can be divided into following two categories: data grid and compute grid. The role of data grid is to manage the data distributed over various locations while in compute grid, the primary resource that is handled is compute cycle[11].

To execute resource management and job scheduling in grid computing, a resource/meta scheduler has to be used. The main task of resource scheduler is to characterize and select the best resource for the incoming job[12].

There exists two types of schedulers in grid computing: local scheduler and global scheduler.Scheduling of jobs and resource management at a single node is referred to as local scheduler while the role of global scheduler is to select proper jobs mapping or a local site over the domain which is selected[13].

Grid is considered as a system which is distributed and contains huge number of files. It is very different from other computing environments where each node is asked to perform a different job.Computers in grid are very heterogeneous and widely dispersed. Grid are used to perform various applications and middleware software's are used to construct large sized grids[14].

This paper is structured as follows : Section 1 describes the introduction to grid computing. Section 2 describes basic concepts of grid computing . Section 3 describes different scheduling approaches . Section 4 ends up with conclusion.

II. BASIC CONCEPTS OF GRID COMPUTING

A. Architecture of Grid Computing

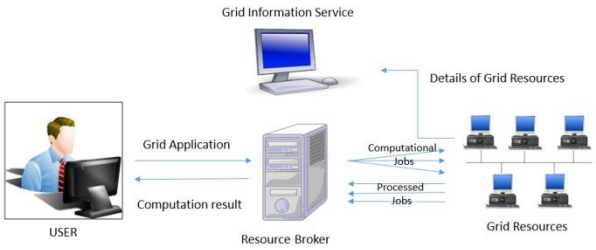

Figure 1: A Grid Architecture [15]

Components of grid architecture:

User: They are the ones who submit their applications to grid along with their metrics to boost execution.

Grid Information Service (GIS): It maintains the information of all resources including their id, capacity, capability, load balance etc. that is needed by the broker for scheduling the jobs efficiently.

Scheduler/Resource Broker:
It gathers information about resources from GIS and helps in scheduling the jobs in accordance with their QoS and resource characteristics. Scheduler provides best allocation of resources by giving benefits like efficient resource utilization, high throughput, less cost .

Resources: Hard disk, Personal Computers, CD are called resources. Wired/wireless form of connection is used to connect various resources.

b) Grid Computing - advantages[16]

1. Exploitation of resources (underutilized)
   Busy/idle machines are used to run an existing task.
2. CPU capacity (Parallel)
   Various fields like modelling related to finance, chemical etc. use this.
3. Collaborating organizations and virtual resources
   Distributed computing has been able to reach wider audience to some extent.
4. Accessing other required resources
   Access to CPU , storage spaces and other needed resources.
5. Resource balancing
   Grid enabled applications can provide resource balancing by scheduling jobs on machines with low utilization.
6. Reliability
   Expensive hardwares which are built using chips helps in increasing reliability.
7. Management
   Grid helps in handling heterogeneous systems for managing a larger IT infrastructure.

***b*) Modes of Scheduling**.

1. Batch Mode Scheduling Algorithms:

They schedule the tasks that are parallel. Job Execution time is less and there is specific order for scheduling of resources.[17]

2. Online Mode Scheduling Algorithms:

The tasks are scheduled and executed as soon as they are entered in the grid. High-level of resource management is required for these types of algorithms.[15]

III. VARIOUS SCHEDULING APPROACHES IN GRID COMPUTING ENVIRONMENT

A. Grid Computing for Effective Performance of Job Scheduling(GCEPJS) [1]

Jobs are classified based on its "Job Type". The job types are "Data Intensive" i.e. the jobs which require more data access power and "Computational Intensive" i.e. the jobs which require more CPU power. The resources are classified based on the CPU speed, baud rate and failure rate. The failure rate is calculated based on the number of successfully executed jobs and number of failures.

Computational Intensive jobs are scheduled to resources that have more CPU speed and less failure rate while Data Intensive jobs are scheduled to resources that have high baud rate and less failure rate.

Advantages:
a) Jobs are allocated more efficiently.
b) Reduces makespan and increases throughput.
Disadvantages:
a) There is a chance of resource failure.
b) Resources are not efficiently utilized.

B. An Efficient Resource Scheduling Algorithm Using Knapsack(GCEPJS) [2].

In this paper, Efficiency of resource is high. In this Resource works as Knapsack. Max Heap tree is used and highest execution time required job will be on top of the tree. First move is done from top node to the left side node and job is assigned to a resource .In some case when job is not assigned then backtracking is used and right node is selected.
Advantages:
a) Efficient Resource Utilization.
b) It reduces total processing cost.

Disadvantages:

a) An efficient resource scheduling algorithm using knapsack performs less as the no. of jobs is decreased.

C. Fault Tolerance in Grid Computing with Improved Resource Utilization and Maximum Efficiency (FTGCIRUME) [3].

This paper gives a method to improve the resource utilization with maximum efficiency and throughput even in occurrence of fault in system.

It contains the phase of Job Analysis:

- Resource allocation phase
- Job execution phase
- Job analysis

The basic objective of all the entities is to offer fault tolerance in grid computing environment by using best method with low cost. The rate of increase in network bandwidth is increasing at a rate faster than that of processer speed that make best use of computing power.

Advantages:

a) Improved resource utilization
b) Maximize efficiency
c) Fault tolerance reduced

Disadvantages:

a) Cost is more
b) Failure still exist

D. An Adaptive Grouping based on Job Scheduling in Grid Computing(AGJSGC)[4].

This paper introduces a model in job scheduling in grid computing environments. A dynamic scheduling algorithm is proposed to maximize the resource utilization and minimize processing time of the jobs .The proposed algorithm is based on job grouping .The results shows that the proposed scheduling algorithm efficiently reduces the processing time of jobs.Grouping based on resources status according to processing capabilities ,bandwidth and memory size of the available resources.

After gathering jobs and resources ,the system selects jobs in FCFS order to form different job groups .Jobs are put into a job group one after another until sum of the requirements of the jobs in that group is less than or equal to amount of resources available at the resource site.

Advantages :

a) Improve processing of fine grained jobs .
b) Reduce execution time.
c) High time complexity
d) Reduces waiting time

Disadvantages:

a) There are some specific set of jobs that require only to a specific set of resources for assignment.
b) Some jobs may require the processing capabilities of more than a resource.

E. Scheduling in Grid Computing Environment (SGCE)[5].

This paper discusses distributed computing environments in current use ,presents fundamental concepts of scheduling ,and relate scheduling in grid with scheduling in systems.It discusses various sub - systems related to grid computing and also presents discussion about grid simulation tools .It covers scheduling algorithms related to grid computing falling in to two main types :

Task and Resource. It Focuses on the technology used for scheduling algorithm in both environment real as well as simulation.

ADVANTAGE:

a) Provides Good knowledge of both real as well as simulation environment of grid Computing.

DISADVANTAGE:

a) Theoretical approach.

F. Workflow Heuristics for Resource Scheduling in Grid Computing Supporting QoS(WHRSGCS)[6].

This paper is based on FCFS and EBF algorithm. In This Resource Utilization is average but it is expected that the utilization of resource will be more when the size of jobs will increase. It shows it is having better performance when compared to FCFS with respect to Workloads. There are many several limits to this work such as Scheduling heuristic which this paper have proposed has been implemented on a simulator i.e. on

GridSim 5.0. It is known that the simulation does not provide the real experience. In advance they try to add more objectives such as load balancing etc.

ADVANTAGES.
a)Better performance than FCFS.
b)Low cost.

DISADVANTAGES.
a)Average utilization of Resource.
b)Failure still exist.

G. Bio-Inspired Optimization Techniques for Job Scheduling In Grid Computing(BIOTJSGC)[7].

Various bio-inspired optimization algorithms have been discussed. Due to arising lot of issues in scientific and engineering applications grid scheduling has emerged as a challenge. Bio-inspired techniques are much more efficient in solving even multi-objective problems in relation to decentralized computation environments. They can be smoothly modified with other stand alone approaches. These Algorithms are very much popular among researchers because of the simplicity & efficiently finding optimal solutions.

ADVANTAGES.
a) Full- fills large scale computer demands.
b) It is effective for computing optimization problem.

DISADVANTAGES.
a) Scheduling in such environment is difficult.
b)High cost.

H. FastPGA based Scheduling of Dependent Tasks in Grid Computing to provide QoS to Grid users(FSDTGCQG)[8].

The mechanism is multiple objective approach that optimizes total schedule length to minimize the overall execution time, penalty cost and job tardiness time to provide good quality of service to users of Grid Computing. The proposed scheduling mechanism enables the grid users to execute their applications within provided budget and deadline. The users have to pay minimum penalty cost if the resource usage cost goes beyond the predetermined budget of the user. The input file which is generated in standard task graph. The implementation done in ALEA 3.0 grid scheduling simulator. This paper shows that FastPGA performs better than other heuristics like NSGA-II and GA.

ADVANTAGES:
a)Good performance.
b)Multiple jobs at same time.

DISAVANTAGES:
a)High cost.
b)Scheduling in such environment is crucial task.

I. Task Scheduling in Grid Computing using Genetic Algorithm(TSGCGA)[9].

Genetic Algorithm provides the solutions of difficult problems by simultaneously searching various regions of solution space. Each individual is represented as possible solution. The solutions are the schedulers for efficiently allocating jobs to resources in a Grid system

ADVANTAGES:
a) Task is taken into consideration.
b) Helps to find out most Optimal Solution.

DISADVANTAGES:
a) Not perform multiple task at same time.
b) Cost is High.

J. Incremental Checkpoint Based Failure-Aware Scheduling Algorithm in Grid Computing(ICFASAGC)[10].

This research paper, represent a solution to improve performance under failure of Grid Computing System. The solution is to use a fault tolerant environment and design failure aware task scheduling. The fault tolerant environment used is incremental checkpoint or restart mechanism. First calculate the expected wasted time due to failure for each task on every node by using incremental checkpoint or restart mechanism. Then re-calculate the computational capacity of Grid system nodes. Finally, schedules the tasks according to new recalculated node capacity. This paper proposes a failure–aware scheduling algorithm which takes into consideration both performance factor and failure rate of the nodes while scheduling tasks. This paper also compare proposed failure-aware scheduling algorithm with speed-only scheduling algorithm and experimental results show that the improvement of the proposed

failure-aware scheduling algorithm is significant for Grid computing, where failure rate is considerable.

ADVANTAGES.

a)Improved performance.
b)Improvement of the proposed failure aware scheduling algorithm.
c)Improved in Performance Ratio, Failure ratio, Success ratio, Average Response time.

DISADVANTAGE.
a)There are still chances of failure.

Table I. Comparision of Different Jobs and Resource Scheduling Algorithms.

PAPER	Res.Ti.	Reso.Uti.	Lo.Bl.	COST	Sim.Kit
GCEPJS	L	H	H	H	GridSim
ERSAK	L	H	H	L	---
FTGCIRUME	L	H	AVG	H	---
AGJSGC	L	H	H	H	GridSim
SGCE	L	H	H	H	Grid Simulator
WHRSGCS	L	AVG	AVG	L	GridSim
BIOTJSGC	H	H	H	H	MATLAB
FSDTGCQG	L	H	H	H	ALEA
TSGCGA	L	H	H	AVG	---
ICFASAGC	L	H	AVG	AVG	---

Abbreviations: Res.Ti-Response Time, Reso.Uti.-Resource Utilization, Lo.Bl. – Load Balance, Sim.Kit-Simulation Kit, H-High, AVG-Average, L-Low

IV. CONCLUSION

In this paper, a review of resource management system in grid computing is presented. Several types of models for resource management in grid computing are discussed. Various scheduling algorithm in grid computing have been analysed. A comparison on various parameters like response time, load balancing, resource utilization, cost and the simulation kit used is done on different types of job scheduling.

V. REFERENCES

[1]. B. Kalaiselvi, Dr.M.Thangamani, "Grid Computing for effective performance of Job Scheduling",2016,ISJCRESM Vol 1 ,Issue 1,March-May 2016.ISSN :2456-1134.

[2]. Rashmi Kumari,Raksha Pandey, "An efficient Resource Scheduling algorithm using Knapsack",IJCSEITR ISSN(P):2249-6831;ISSN€:2249_7943 Vol. 7,Issue 2 Apr 2017,25-28.

[3]. Minakshi Memoria,Dr.Mukesh Yadav,"Fault Tolerance in Grid Computing with Improved Resource Utilization and Maximum Efficiency",ISSN 2321 3361.

[4]. S. Gomathi, Dr.D.Manimegalai, "An adaptive grouping based Job Scheduling in Grid Computing",Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN 2011).

[5]. Harshadkumar B. Prajapati, Vipul A. Shah, "Scheduling in Grid Computing Environment",2014 Fourth International Conference on Advanced Computing & Communication Technologies.

[6]. Neha Bhardwaj. Karambir Singh, Ajay Jangra,"Workflow Heuristics for ResourceScheduling in Grid Computing Supporting QoS",IEEE2015.

[7]. Reetika Grover, Amit Chabbra,"Bio-Inspired Optimization Techniques for Job Scheduling In Grid Computing", IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016.

[8]. Mandeep Kaur, "FastPGA based Scheduling of Dependent Tasks in Grid Computing to provide QoS to Grid users",IOTA, 22 Jan - 24 Jan, 2016.

[9]. Ujjwal Prajapati,Prof. Dr. Subarna Shakya, "Task Scheduling in Grid Computing using Genetic Algorithm",2015 IEEE.

[10]. Manjeet Singh, "Incremental Checkpoint Based Failure-Aware Scheduling Algorithm in Grid Computing"International Conference on Computing, Communication and Automation ICCCA2016.

[11]. K. Jairam Naik, Dr. A Jagran , Dr. N Satya Narayana, "A Novel Algorithm for fault tolerant job Scheduling and load balancing in grid computing environment", ICGCIoT,2015.
[12]. Anubha Chauhan , Smita Singh, Sarita Negi, Shashi Kant Verma, "Algorithmfor Deadline based Task Scheduling in Heterogeneous Grid Environment", NGCT-2016, October 2016.
[13]. Khusboo Yadav, Deepika Jindal, Ramandeep Singh, "Job Scheduling in Grid Computing", IJCA (0975-8887), Vol. 69- No. 22,May 2013.
[14]. Grid Computing.[Online]. Available :https://en.wikipedia.org/wiki/Grid_computing.
[15]. Heena Setia , Abhilasha Jain, "Literature Survey on Various Scheduling Approaches in Grid Computing Environment", ICPEICES-2016.
[16]. Shruti N. Pardeshi, Chitra Patil, Snehal Dhumale, "Grid Computing Architecture and Benefits", IJSRP,Vol.3, Issue 8,August 2013.
[17]. Jia Yu, Rajkumar Buyya, Kotagiri Ramamohanarao, "Workflow Scheduling Algorithms for Grid Computing", Springer – Verlag Berlin Heidelberg ,SCI 146, pp. 173-214, 2008.