Job Scheduling in Grid Computing

Khushboo Yadav   Deepika Jindal   Ramandeep Singh

ABSTRACT

Job scheduling is used to schedule the user jobs to appropriate resources in grid environment. The goal of scheduling is that it achieves highest possible system throughput and match the application need with the available computing resources. In this paper, we will review the definition of grid computing, types of Grids, framework of Grid scheduling and characteristics of Computational Grid. Grid is a system in which machines are distributed across various organizations. Grid computing is a form of distributed computing. It involves sharing of resources that are heterogeneous and geographically distributed to solve various complex problems and develop large scale applications. Grid computing is broad in its domain of application and raises research questions that span many areas of distributed computing and of computer science in general. In this paper, we will explain job scheduling and resource scheduling algorithms that will benefit the researchers to carry out their further work in this area of research.

Keywords

Grid computing, Job Scheduling, Resource Scheduling.

1. INTRODUCTION

Definition: GRID computing has become apparent as the next generation parallel and distributed computing methodology. Its instance is to provide a service-oriented infrastructure to enable easy access to and coordinated sharing of geographically distributed resources for solving various kinds of large-scale parallel applications. Now a days, grid computing has been widely accepted, study, and given attention to by researchers [1]. Unlike the traditional file exchange, as supported by the Web or peer-to-peer systems, users in the grid can access the required resource or service in a transparent way as if they were to use local resources or services. However, it gives rise to any of two or more ideas conflict between grid users and resource providers in usage policy of the local resources. For users, in addition to simplicity and easiness, to get desirable service functionalities, some quality of service (QoS) targets associated with the service, such as grid service reliability [2], the financial cost of the resource, and the efficiency of grid service, may be specified when a service is submitted. On the other hand, resource providers receive the compensation from grid users for the consumed resources at the price of sacrificing local task executions [3].

1.1 Types of Grids

A grid can be classified into three categories: Computational grid, Data grid and Service grid. Fig. 1 shows the categories of grid systems. A computational grid provides access to huge pool of shared processing power suitable for high throughput applications and computation intensive computing. Depending on how the capacity is utilized, computational grid is classified into Distributed Supercomputing grid and High

Throughput grid. A supercomputing grid executes applications on multiple machines in parallel to reduce the computation time of a task. A data grid provides an infrastructure for synthesizing information from data archives such as digital libraries or data warehouses that are distributed in a wide area of network. European data grid project and Globus are data grid initiatives, working on developing large scale data organizations. A service grid provides services that cannot be accomplished by any single machine. The service grid further categorized into on-demand grid, collaborative grid and multimedia grid systems [4].
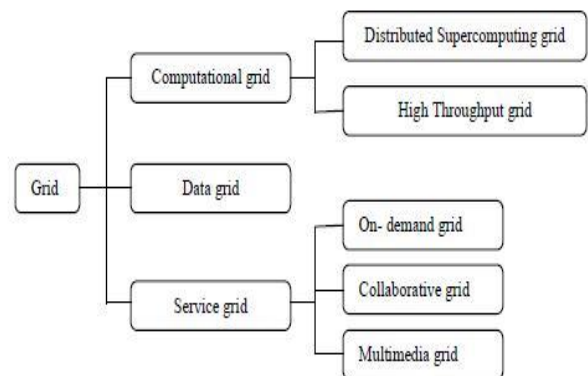

Fig. 1 Different categories of Grid systems

1.2 Framework of Grid scheduling

In the various types of resources the grid may be shared and used, they are usually accessed via an executing application or job. A job or meta-task or application is a set of atomic tasks that will be computed on a set of resources. Jobs can have sub-jobs or tasks, and sub-jobs can themselves be decomposed further into atomic tasks. The grid system is responsible for sending a job to a computing element in the grid. More advanced grid systems include a job scheduler of some kind that automatically finds the most suitable machine on which to run any given job that is waiting to be executed. There are two kinds of schedulers in grid environment: local scheduler and grid scheduler. A scheduler or local scheduler is responsible for scheduling of jobs and managing resources at a single site or node or cluster or resource. A global scheduler or grid scheduler is responsible for selecting appropriate local site and mapping of jobs onto the selected site or domain. A grid scheduling framework has been shown in Fig. 2. Due to the heterogeneous nature of grid, information about the status of available resources in the grid is very important for global scheduler to make scheduling decisions. The Grid Information Service (GIS)[5] provides information about resources such as CPU capacity, availability, [7]. memory size, network bandwidth and load of site in a particular period to grid scheduler. A local scheduler resides in a resource domain, it schedules not only jobs from exterior grid users but also jobs from the users of that local site and it reports the information to GIS. There can be one or more local schedulers running

within a site with specified resource management policies. Some of the local schedulers are Condor [6] and Open PBS [7].

1.3. Characteristics of Computational Grid

There are many desirable properties and features that are required by a grid to provide users with a computing environment. They are as follows:

- Heterogeneity:-The grid involves a number of resources that are varied in nature and can encompass a large geographical distance through various domains.

- Scalability:-The grid should be tolerant to handle a large number of nodes without any performance degradation.

- Adaptability or Fault Tolerant:-In a grid unexpected computational aborts, hardware or software faults etc are high. These faults are generally handled by Resource Managers.

- Security:-All the user participating computers should be protected from any malicious manipulations or interventions.

2. JOB SCHEDULING

Generally the objective of job scheduling is to have good load balancing among the processors, whereas for the later minimization of overall execution time is the main concern [8]. Job scheduling and resource scheduling are the two main necessities in grid computing. In job scheduling, the job scheduler has to find the appropriate resource for the job that the user submits [9]. It has to find the best machine in grid to process the user job.

Grid has two main schedulers such as local schedulers and grid schedulers. The local schedulers work in local computational environment and hence it is reliably, fast connection, works in uniform environment and also takes full control of the homogeneous resources [10]. Grid Schedulers also called as meta-schedulers are the top level schedulers. They are responsible for orchestrating resources that are managed by different local schedulers [11]. Scheduling can also be classified into static and dynamic scheduling. In static scheduling, before execution the jobs are assigned to the suitable machines and those machines will continue executing those jobs without interruption. In dynamic scheduling, the rescheduling of jobs is allowed. The jobs executing can be migrated based on the dynamic information about the workload of the resources [12]. In grid, there may be lots of resources to run a job. The main focus is to find the appropriate resource for the job that is to schedule the job. The methods for job scheduling are centralized, hierarchical and decentralized.

- In centralized scheduling, there will be a centralized scheduler and it is responsible for scheduling the jobs. It is very useful when all the resources have same objective.

- In hierarchical, there will be central scheduler. All jobs will be submitted to the central scheduler. The central scheduler redirects the jobs to the global scheduler.

- In decentralized, there is no central scheduler. Distributed schedulers coordinate with each other to schedule jobs [13].
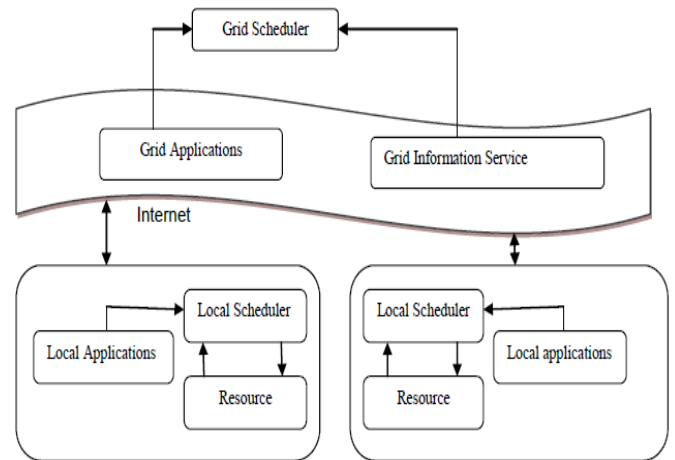

Fig. 2. Grid Scheduling Framework

3. DIFFERENT METHODS OF JOB SCHEDULING ALGORITHMS

Job Scheduling The aim of job scheduling system is to select the best machines in a grid for various users' jobs. For each machine, job schedulers are generated by the management and scheduling system. Effective computation and Job Scheduling is becoming one of the main challenges in grid computing.

3.1. Efficient Utilization of Computing Resources using Highest Response Next Scheduling in Grid (HRN)

In Highest Response Next Scheduling schema [14] jobs are allotted to number of processors based on job's priority and processor's capability. HRN scheduling algorithm was proposed to correct the weaknesses of both Shortest Job First and FCFS. This algorithm provided for response with time, memory and CPU requirement. The advantages are it utilizes the resources efficiently and HRN model is much adaptive for

Grid environment. The disadvantages are that it contains high turnaround time, memory and CPU wastage.

3.2. Node Allocation in Grid Computing using Optimal Resource Constraint (ORC) Scheduling

ORC scheduling algorithm includes the combination of both the Best fit allocation and Round Robin scheduling to allocate the jobs in queue pool [15]. This algorithm improved the efficiency of load balancing and dynamicity capability of the grid resources. The advantages are reduces waiting time of job and turnaround time and

increases processing time of jobs. The disadvantages are high communication overhead.

3.3. Hierarchical Job Scheduling For Clusters of Work Station (HJS)

Hierarchical job scheduling approach used two levels Scheduling global scheduling & local scheduling [16]. The global scheduler uses separate queues for different type of the

jobs for scheduling with the FCFS, SJF and first fit (FF) and the local scheduler uses same queue for different type of the jobs. The advantages are it uses multi queue to allow more control over the allocation of resources to jobs and reduces turn6around time. The disadvantages are this algorithm does not consider the dynamic behavior of the grid resources and chance of under utilization of grid resources.

3.4. Scheduling Framework for Bandwidth Aware Job Grouping Based Scheduling in Grid Computing (SFBAJG)

The use of bandwidth in scheduling framework was explored to enhance the performance of job scheduling [17]. The scheduling frame work consisted of job scheduler that obtained information form grid information service (GIS). About the resources on the basis of this information job scheduling was used for grouping of jobs & selecting resources. The framework has a grouping and selection service where matching of jobs is performance based on the information gathered from information collector the information collector will information about the network bandwidth to reach each resource, then this information is used for grouping and selection service to gather the necessary information which is required for job selection the GIS contains all the resources and provide information about them jobs to the respective resources. The advantages are improves the dissemination of jobs of jobs in grid computing. This framework is a good alternative for delivering job scheduling process with better job scheduling performance and reduces total processing time. The disadvantages are does not consider memory size constraint, does not take dynamic resource characteristics.

3.5. Grouping Based Fire Grained Job Scheduling in Grid Computing (GB FJS)

The grouping based fire grained job scheduling algorithm [18] starts with obtaining info about the resources. In this algorithm, light weight jobs are grouped as coarse grained jobs. The groping based algorithm utilized resourced efficiently of integrates greedy algorithm of FCFS algorithm. To improve processing undertake of fire grained jobs. The advantages are it reduces total processing time, reduces execution time of the jobs and maximize the utilization of the resources. The disadvantages are it does not consider memory size constraint and preprocessing time of job grouping is high.

3.6. A Bandwidth- Aware Job Grouping – Based Scheduling On Grid Environment (BAJGS)

In bandwidth aware job grouping based scheduling algorithm the job grouping concept is explored coupled with bandwidth aware scheduling [19]. This algorithm focuses on grouping independent jobs having small processing requirement into jobs with larger processing requirements and then schedules them according to network conditions. The concept of bandwidth was used for performing load balancing at stream control transmission protocol (SCTP) layer. Its main objective was to provide the in-order delivery over multiple paths. This approach reduces total job processing time as compared to job scheduling without grouping. The advantages are this algorithm focuses on grouping jobs with small processing requirements into jobs with larger processing requirements and minimizes total processing time. The disadvantages are

does not implement load balancing scheme and this algorithm does not include QOS requirements.

3.7. An Agent Based Dynamic Resource Scheduling Model with FCFS Job Grouping Strategy in Grid Computing (ABDRS)

The agent based dynamic resource scheduling strategy focuses on maximize the processing time of jobs [20]. The process of selecting a job is based on maximum heap tree. The processed outsource model is a hierarchical two layer approach in which top layer is called grid level an d other is called cluster level. The Advantages are this algorithm enhances scalability, robustness and load balancing availability of the Grid, Maximize resource utilization and Minimize processing time of jobs. The disadvantages are in this algorithm does not implement job and resource scheduling with genetic algorithm to increase the performance.

3.8. A Dynamic Job Grouping Based Scheduling for Deploying Application with Fine Grained Task on Global Grid (DJGBS)

Dynamic job grouping based scheduling algorithm group the jobs according to MIPS of the resource [21]. It selects resource in first come first serve order. It selects jobs and group the jobs and assigns the group jobs in FCFS order and compare to resource if group job MI is less that to resource MIPS of this process continues until the resource MIPS is less to group job. The advantages are it maximizes the utilization of the resource and it reduces the total processing time of jobs. The disadvantages are it does not consider the bandwidth and memory size constraints.

3.9. Virtual Computing Grid Using Resource Pooling (VCGRP)

Virtual computing grid means that the system can choose a resource of allocate tasks to it [22]. It is a single point web based access known as virtual computing grid portal and the virtual computing grid monitor is a central resource manager for the system. The advantages are it is cost effective model. The disadvantages are it is not much reliable and since it is cost effective solution quality of service has been play done in the prototype model.

4. CONCLUSION AND FUTURE SCOPE

Grid computing can be solve complex tasks in shorter time and utilizes the hardware efficiently. The best job scheduling strategies have to be employed to make the grid work efficiently. Job scheduling is the foremost step in grid computing where the users' jobs are scheduled to different machines. The various strategies have been studied and classified. The advantages and disadvantages of the algorithms have been studied. In this paper we have discussed various job scheduling algorithms that can be used to schedule different jobs. This paper also describes the various components of the grid scheduling system. The future work will be concerned with the development of the better scheduling algorithm which is heterogeneous and works in dynamic environment.

5. REFERENCES

[1] I. Foster, 2002, "The Grid: A new infrastructure for 21st century science,"Physics Today.

[2] Y. S. Dai, M. Xie, and K. L. Poh, 2006, "Reliability of grid service systems," Computers and Industrial Engineering.

[3] S. C. Guo, H. Wan, G. B. Wang, and M. Xie, 2010, "Analysis of grid resource compensation in market-oriented environment," Eksploatacja INiezawodnooeæ— Maintenance and Reliability.

[4] Klaus Krauter, Rajkumar Buyya and Muthucumaru Maheswaran, 2002, "A taxonomy and survey of grid resource management systems for distributed computing", Software-Practice and Experience.

[5] Karl Czajkowski, Steven Fitzgeraldt, Ian FosterB and Carl Kesselman, 2001, "Grid Information Services for Distributed Resource Sharing".

[6] Condor http://research.cs.wisc.edu/condor/manual/v7.8/index.html

[7] Sterling, T,"PBS: Portable Batch System " , Beowulf Cluster Computing with Windows, MIT Press.

[8] Gyung, L.P., 2004, "Performance Evaluation of a List Scheduling Algorithm in Distributed Memory Multiprocessor Systems".

[9] R.Buyya and M.Murshed, 2002, "Gridsim: a toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing," Concurrency and Computation: Practice and Experience.

[10] Raksha Sharma, Vishnu Kant Soni, Manoj Kumar Mishra and Prachet Bhuyan, 2010, "A Survey of Job Scheduling and Resource Management in Grid Computing", in World Academy of Science, Engineering and Technology.

[11] E. Huedo, R.S. Montero, and I.M. Llorente. 2005, The Grid Way framework for adaptive scheduling and execution on grids. Scalable Computing: Practice and Experience.

[12] M. Chtepen, 2005, "Dynamic scheduling in grids system," Sixth Firw PhD Symposium, Faculty of Engineering, Ghent University.

[13] Mr. Rakesh Kumar and Navjot Kaur, "Job Scheduling in Grid Computers".

[14] K.Somasundaram, S.Radhakrishnan M.Gomathynayagam, 2007, "Efficient Utilization of Computing Resources using Highest Response Next Scheduling in Grid" 6 (5): 544-547, Asian Journal of Information Technology.

[15] K.Somasundaram, S.Radhakrishnan, 2008, "Node Allocation in Grid Computing Using Optimal Resource Constraint (ORC) Scheduling",VOL.8 No.6, IJCSNS International Journal of Computer Science and Network Security.

[16] J.Santoso; G.D. van Albada; B.A.A. Nazief and P.M.A.Sloot, 2000, "Hierarchical Job Scheduling for Clusters of Workstations", ASCI 2000, pp. 99-105. ASCI, Delft.

[17] Ng Wai Keat,Ang Tan Fong,Ling Teck Chaw,Liew Chee Sun, 2006, "Scheduling Framework For Bandhwidth-Aware Job Grouping Based Scheduling In Grid Computing", Malaysian Journal Of Computer Science.

[18] Quan Liu, Yeqing Liao, 2009, "Grouping-Based Fine-grained Job Scheduling Grid Computing", IEEE .

[19] T.F Ang, W.K.Ng, T.C Ling, 2009, "A Bandwidth-Aware Job Grouping Bas Scheduling on Grid Environment", Information Technology Journal.

[20] Kumar Mishra, "An Agent Based Dynamic Resource Scheduling Model with FCFS-Job Grouping Strategy In Grid Computing", Waset, ICCGCS 2010. In Press.

[21] Nithiapidary Muthuvelu, Junyang Liu, "A Dynamic Job Grouping- Based Scheduling for Deploying Application with Fine-Grained tasks on Global Grids, Australasian Workshop on Grid Computing and e- Research, vol. 44, AusGrid -2005.

[22] Alpana Rajan, Anil Rawat, Rajesh Kumar Verma, 2008, " Virtual Computing Grid using Resource Pooling", IEEE, International Conference on Information Technology.

[23] K. Krauter, R. Buyya, and M. Maheswaran, 2002, "A taxonomy and survey of grid resource management systems for distributed computing," Software—Practice and Experience.

[24] C. L. Li and L. Y. Li, 2007, "Multiple QoS modeling and algorithm in computational grid," Journal of Systems Engineering and Electronics.