International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-5, Issue-1, May 2016

Survey on Dynamic Group Job Scheduling in Grid Computing

Mr. Gopal S. Ade, Dr. S.Y. Amdani

***Abstract*— Grid computing is the novel framework that offers a flexible, secure and high performance computing, on demand for solving high compute-intensive applications with large number of independent jobs. However, user jobs developed for grid might be small and of varying lengths according to their computational needs and other requirements. Certainly, it is a real challenge to design an efficient scheduling strategy to achieve high performance in grid computing. But there exists some grouping based job scheduling strategy that intends to minimize total processing time by reducing overhead time and computation time, and on the other hand maximizing resource utilization than without grouping based scheduling. The purpose of the study is to analyze and achieve better performance by extending the concept of grouping based job scheduling. Therefore, this paper proposes "A Time-Minimization Dynamic Grouping-Based Job Scheduling in Grid Computing" with the objective of minimizing overhead time and computation time, thus reducing overall processing time of jobs. The work is verified through various observations made in different simulated grid environments. The results obtained shows that the proposed grouping-based scheduling algorithm is on average, comparable to, or even better than, other grouping based scheduling algorithms.**

***Index Terms*— Grid computing; Job grouping; Job scheduling.**

I. INTRODUCTION

The word "Grid" refers to systems and applications that integrate resources and services distributed across multiple control domains. Computational grids provide large-scale resource sharing, such as personal computers, clusters, MPPs, Data Base, and online instructions, which may be cross-domain, dynamic and heterogeneous. If one considers the internet as a network of communication, grid computing can be considered a network of computation This enables the users to compute large scale applications in science, engineering and business, by utilizing the increased access to geographically spread and dynamically available processing, scalable, economical and authorized resources to the registered users on demand, hiding most of its underlying details and complexities from the outside world. To realize the full potential of grid computing, grid middleware needs to support various services such as security, uniform access, resource management, job scheduling, application composition, economic computation, and accounting. Since, it is responsible for Various scientificand business organizations tend to have increased number of applications with large number of independent jobs, scheduling of these jobs onto the grid is significantly more difficult and complicated than scheduling applications in traditional supercomputer because of the heterogeneous, Dynamic and diverse nature of the Grid resources.

Fig. 1: Grid Computing

Therefore, optimal scheduling of various jobs onto grid is not easy to attain, since optimal scheduling of heterogeneous jobs in heterogeneous environments is known to be NP-Complete problem. In order to ensure the efficiency and better performance of job scheduling, an effective and near optimal scheduling mechanism has to be developed and implemented to cater the needs of the grid users. In traditional parallel computing system, the communication cost is considered to be insignificant as homogeneous computing nodes are interconnected in a geographically small area network for instance LAN. When a job is submitted to a grid resource for execution, it is transmitted over the networks incurring a communication cost. An application with large number of fine-grained jobs when submitted individually to the grid resources over the networks incurs a communication overhead that is more than the total computation time of each job at the resource. This grouping based job scheduling strategy reduces communication time resulting increase in computation communication ratio (CCR), which encourages distributing grouped jobs for processing on remote resources. Hence, scheduling should be addressed by developing a grouping strategy suitable to both type of grid environment. Grid resource allocation algorithm that must be efficient and effective in reducing the total processing time of jobs. Analyzes related works in the field of parallel and distributed memory system and grid computing systems. The grid system and scheduling components (broker). Dynamic grouping based job scheduling model. Conclusion and future work and lastly, the references.

Mr. Gopal S. Ade, M.E. (Second year), Department of CSE, BNCOE, Pusad
Dr. S.Y. Amdani, Professor & HOD,Department of CSE, BNCOE, Pusad

II. GROUPING STRATEGY

Although Grids have been used extensively for executing intensive jobs, but there are also exist several applications with a large number of lightweight jobs which needs small processing requirements. Sending/receiving each small job individually to/from the resources will increase the total communication time and cost. These applications involve high overhead time and cost in terms of job transmission to and from Grid resources and, job processing at the Grid resources. So there is a need for an efficient job grouping-based scheduling system to dynamically assemble the individual fine-grained jobs of an application into a group of jobs, and send these coarse-grained jobs to the Grid resources. This dynamic grouping should be done based on the processing requirements of each application, Grid resources' availability and their processing capability. The job grouping strategy results in increased performance in terms of low processing time and cost. Job grouping based scheduling also useful for maximum resource utilization and reducing communication overhead time.

1. Basic Job Grouping

The grouping strategy groups the small scaled user jobs into few job groups according to the processing capabilities of available Grid resources. This strategy is based on processing capability (in MIPS), bandwidth (in Mb/s), and memory-size (in Mb) of the available resources. The existing Group Based Job scheduling algorithm basically follows the below given steps: -

A. The job scheduler obtains information about the available resources from the Grid Information Service (GIS).

B. Based on the information, the job scheduling algorithm is used to determine the job Grouping and resource selection for grouped jobs.

C. Jobs are started to put into a group until job processing requirements not equal to resource Processing capability. So During job grouping the following conditions must be satisfied:

Grouped job_ MI<=Resource_ MIPS*Chunk_ size. Where MI (Million Instruction) is job's required computational power, MIPS (Million Instruction per Second) is processing capability of the resources and Chunk_ Size is User defined time

A. This grouping process continues until all the jobs are put in groups.

B. After completion of grouping, the scheduler sends the grouped Jobs to the corresponding resources for computation.

C. The Grid resources compute the received grouped job and send result back to the user.

III. LITERATURE SURVEY:

There has been a lot of research done in the area of Grid Computing. In this section, some of the representative research works on job scheduling in parallel and distributed computing systems and Grid computing environment have been reviewed to explore the relevance of these works. Vishnu Kant Soni, Raksha Sharma, Manoj Kumar Mishra, et.al. suggested "Grouping-Based Job Scheduling Model in Grid Computing" [1]. In their paper Grid computing is a high performance computing environment to solve larger scale computational applications. PayalSinghal, Ravinder Singh and Pinky Rosemarry, et.al. Suggested "An Improved Constraint Based Resource Scheduling Approach Using Job Grouping Strategy in Grid Computing" [2]. In their paper Grid computing is a collection of distributed resources interconnected by networks to provide a unified virtual computing resource view to the user. The constraint based job and resource scheduling algorithm has been proposed. Rabab MohamedEzzat,AmalElsayedAboutabl, Mustafa Sami Mustafa et.al. suggested "Grouping-based Scheduling with Load Balancing for Fine-Grained Jobs in Grid Computing" [3]. In their paper proposes Grid computing is characterized by the existence of a collection of heterogeneous geographically distributed resources that are connected over high speed networks and presents efficient grouping-based scheduling models that group fine-grained jobs to form coarse-grained jobs which are sent for execution on grid resources. Sandeep Kaur, Sukhpreet Kaur et.al. suggested "Survey of Resource and Grouping Based Job Scheduling Algorithm in Grid Computing" [4]. In their paper, various job grouping based scheduling algorithms in grid computing have been surveyed. PanktiDharwa, HarshvardhanMathur et.al. Suggested "Coarse Grained Job Scheduling with Dynamic Strategy in Grid Computing" [5]. Proposed Grid technologies are emerging as the next generation of distributed computing, allowing the aggregation of resources that are geographically distributed across different locations. ThuZar Mon, and Cho Me Mget.al. Suggested "MIPS Group Based Job Scheduling Model for Deploying Applications" [6]. Proposed Different scheduling algorithms have been proposed in every system to achieve his goals. Miss. Hemangi Joshi, Prof. VirajDaxini et.al. Suggested "An Effective Load Balancing Grouping Based Job & Resource Scheduling Algorithms for Fine Grained Jobs in Distributed Environment" [7]. The results obtained shows that the proposed grouping-based scheduling algorithm is on average or comparable to other grouping based scheduling algorithms**.** S. Gokuldev, R. lalithkumaret.al. Suggested "User Deadline Based Job Scheduling in Grid Computing" [8]. Grid computing could be a style of distributed computing that co-ordinates and provides the ability of resource sharing over varied geographical locations. Their paper proposes the job of the bacterial foraging optimization technique for grid resource planning. Dr. D. Manimegalai, S. Gomathi, et.al. Suggested "An Analysis of MIPS Group Based Job Scheduling Algorithm with other Algorithms in Grid Computing" [9]. Two major problems in grid computing applications are, resource management and job scheduling. These problems do occur due to distributed and heterogeneous nature of the resources. Their paper introduces a model in job scheduling in grid computing environments. The results show that the proposed scheduling algorithm efficiently utilizes resources at its best and reduces the processing time of jobs.

IV. NEED OF GRID COMPUTING:

Computational approaches to problem solving have proven their worth in almost everywhere. Computers are used for modeling and simulating complex scientific and engineering problems, diagnosing medical conditions, controlling industrial equipment, forecasting the weather, managing stock portfolios, and many other purposes. Yet, although there are certainly challenging problems that exceed our ability to solve them, computers are still used much less extensively than they

could be. To pick just one example, university researchers make extensive use of computers when studying the impact of changes in land use on biodiversity, but city planners selecting routes for new roads or planning new zoning ordinances do not. Yet it is local decisions such as these that, ultimately, shape our future. There are a variety of reasons for this relative lack of use of computational problem-solving methods, including lack of appropriate education and tools. But one important factor is that the average computing environment remains inadequate for such computationally sophisticated purposes.

V. BENEFITS OF GRID COMPUTING:

Grid computing can provide many benefits not available with traditional computing models:

A.*Better utilization of resources*:Grid computing uses distributed resources more efficiently and delivers more usable computing power. This can decrease time-to-market, allow for innovation, or enable additional testing and simulation for improved product quality. By employing existing resources, grid computing helps protect IT investments, containing costs while providing more capacity.

B.Increased user productivity:By providing transparent access to resources, work can be completed more quickly. Users gain additional productivity as they can focus on design and development rather than wasting valuable time hunting for resources and manually scheduling and managing large numbers of jobs.

C.Scalability:Grids can grow seamlessly over time, allowing many thousands of processors to be integrated into one cluster. Components can be updated independently and additional resources can be added as needed, reducing large one-time expenses.

D.Flexibility: Grid computing provides computing power where it is needed most, helping to better meet dynamically changing workloads. Grids can contain heterogeneous compute nodes, allowing resources to be added and removed as needs dictate.

- ***Advantages:***

a. Maximum resource utilization.
b. Gives better performance in terms of processing time of jobs.
c. Eliminate the communication time of individual small scaled jobs.
d. Reduce the network latency.

- ***Disadvantages:***

a. Does not take dynamic characteristics of the resources into account.
b. Some resources fully occupy and other may be unutilized or idle.
c. Processing load among the available resource is not balanced.
d. Allocation large number of jobs to one resource will increase the processing time.

APPLICATIONS OF GRID COMPUTING:

Grid Computing is being extensively used in various areas like science, business, governments etc. The areas and scope of application is increasing at a very fast pace.

- Energy
- Financial Services
- Manufacturing
- Life Science
- Health Care
- E Telecommunication
- Media
- Collaborative Games
- Government and Education
- Environment
- Astronomy
- Engineering and Design

VI. CONCLUSION:

In order to reduce processing time and utilize grid resources sufficiently, Grouping-Based Job scheduling model has been proposed taking memory constraint into account. Grouping strategy plays a vital role to improve the overall performance of grid computing environment.Grouping of small scale job into group efficiently reduces the processing time of jobs. Grouping strategy based on the processing capability of selected resource. This Can experimental result also shows that the proposed algorithm enhanced the load balancing and for future work It can be introduced more effectively with resource utilization using better scheduling approach.

ACKNOWLEDGMENT:

I would like to thank my supervisor and all of my teachers for their helpful comments in improving our manuscript.

[1] Vishnu Kant Soni, Raksha Sharma, Manoj Kumar Mishra," Grouping-Based Job Scheduling Model in Grid Computing", World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:4, No:5, pp. 961-964,2010.

[2] PayalSinghal, Ravinder Singh and Pinky Rosemarry, "An Improved Constraint Based Resource Scheduling Approach Using Job Grouping Strategy in Grid Computing", International Journal of Grid Computing & Applications (IJGCA) Vol.3, No.4, pp. 33-42, December 2012.

[3] Rabab Mohamed EzzatAmalElsayedAboutablMostafa Sami Mostafa," Grouping-based Scheduling with Load Balancing for Fine-Grained Jobs in Grid Computing", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, pp. 67-75, 2013.

[4] Sandeep Kaur, SukhpreetKaur," Survey of Resource and Grouping Based Job Scheduling Algorithm in Grid Computing", International Journal of Computer Science and Mobile Computing A Monthly Journal of Computer Science and Information Technology IJCSMC, Vol. 2, Issue. 5, pp.214 – 218, May 2013.

[5] PanktiDharwa Harshvardhan Mathur," Coarse Grained Job Scheduling with Dynamic Strategy in Grid Computing", International Journal for Technological Research in Engineering Vol 1, Issue 10, pp.1123-1127, June-2014.

[6] Thu Zar Mon, and Cho Me MeMg," MIPS Group Based Job Scheduling Model for Deploying Applications", International Conference on Advances in Engineering and Technology (ICAET'2014) pp.135-139, March 29-30, 2014 Singapore.

[7] Miss. Hemangi Joshi, Prof. VirajDaxini," Research Article an Effective Load Balancing Grouping Based Job & Resource Scheduling Algorithms for Fine Grained Jobs in Distributed Environment", Hemangi Joshi et al, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.4, pp. 1209-1220, April- 2014.

[8] S.gokuldev,R.lalithkumar," user deadline based job scheduling in grid computing" ijcsns international journal of computer science and network security, vol.15 no.3, pp.62-68, march 2015

[9] Dr. D. Manimegalai, S. Gomathi, "An Analysis of MIPS Group Based Job Scheduling Algorithm with other Algorithms in Grid Computing" IJCSI International Journal of Computer Science Issues,http://ijcsi.org/papers/IJCSI-8-6-3-335-340.pdf. Vol.8, Issue 6, No 3, pp.335-340 November 2011.

[10] Manoj Kumar, Mishra PrithvirajMohanty, G. B. Mund," A Time-minimization Dynamic Job Grouping based Scheduling in Grid Computing", International Journal of Computer Applications (0975 – 8887) Vol 40– No.16, pp.16-25, February 2012

Dr. S. Y. Amdani received his Ph.D (CSE) degree from SGB Amravati university, Amravati and Working as Head and Professor in Deptt. Of CSE B.N.C.O.E., Pusad (India), and life member of Indian Society for Technical Education New Delhi. He has published and presented many research papers in International conferences & journals. He is also Edi-torial Board member of many International Journals.

Gopal Sheshrao Ade, from Pusad and born on 6th may 1987. He did his schooling from JSPM's High school pusad. Completed graduation (BE) in Computer Science and Engg from BNCOE pusad with 65% and Pursuing for ME in Computer Science and Engg with same college, SGBAU University, Amravati, (MS) India. He is willing to work on grid computing. His paper (An efficient technique by using session key to protect relational database) is already published in International Journal of Engineering Research (IJOER).