A Brief Literature Survey on Grid Scheduling Approaches

Saumitra Singh

Senior Auditor

Defence Accounts Department, Prayagraj, India.

Abstract : *Grid is a distributed environment of homogeneous or heterogeneous resources. These distributed resources are frequently used to solve the problems. To solve the problems effectively and efficiently, a good grid scheduling technique has used by which resources are efficiently utilized to meet user demands. However on the user perspective minimization of completion is the main challenge.*

IndexTerms **– Grid, Heterogeneous, Homogeneous, Scheduling,**

I. INTRODUCTION

In distributed environment with the improvement in communication capabilities, the distributed remote resources are now frequently used to solve applications involving large volume of data [1], [2]. This overall system is known as grid computing. A computational grid is a hardware and software infrastructure that provides dependable, consistent, and inexpensive access to high-end computational capabilities. The definition adopted for grid is "A type of parallel and distributed system that enables the sharing, selection, and aggregation of geographically distributed autonomous and heterogeneous resources dynamically at runtime depending on their availability, capability, performance, cost, and users quality-of-service requirements" [3].

II. SYSTEM MODEL

A grid is considered as the combination of multiple layers. In our model the whole system is composed of three layers (Figure 1). The first layer is the user application layer in which the user authentication is done and jobs are submitted to the scheduler by the user. The second layer contains scheduler and Grid Information Services (GIS). The scheduler schedules jobs among various resources after taking resource status information from GIS. The second layer is connected through a Virtual Private Network (VPN) to user. VPN provides additional security and only authorized users can access services. All the resources reside in third layer where user's jobs are executed which are also connected through VPN.
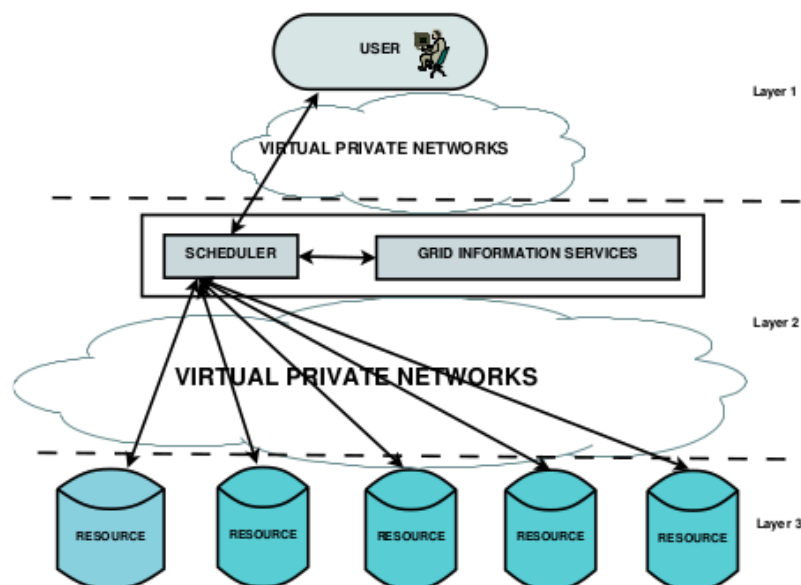

Figure 1: Layered Architecture

III. LITERATURE SURVEY

Resource scheduling for jobs submitted in grid environment had been extensively studied in recent years. Most of the researchers focus on the study of load balancing approaches [6], [7], while others focus on job scheduling in different environments [8], [9], [10]. In job scheduling, few researchers consider independent jobs while others consider the jobs dependency. Recent works in this area consider load balancing approach with job in-dependency and make use of genetic algorithms.

In [11], A. J. S. Santiago proposed an optimization algorithm for assigning jobs to resources and solved the Direct Acyclic Graph (DAG) problem. For both cases CPU idle cycles were used to minimize the execution time. Genetic algorithms were used to

solve the optimization problem. In his approach the communication between resource and scheduler considered to be constant. According to the processing power of resources, the load was distributed. In this way the load on each resource was equally distributed. The resources completed their load on same time.

B. Yagoubi [6] proposed a model for load balancing of tasks to improve its response time. In his model the grid was considered to be of three levels: Intra-site, Intra-cluster, and Intra-grid. In his work he first made a load balance between Intra-site, if it fails to balance then Intra-cluster load balancing is done. If Intra-cluster load balancing also fails then the Intra-grid load balancing strategy was used.

One more approach is proposed by Junwei Cao [7], where intelligent agents were used to decide the work load on resources. In his model three members were present namely, Broker, Coordinator, and Agent. He balanced the grid load at two level-one was local grid load balancing and another for global grid load balancing. The focus of recent works are on new job scheduling strategies which consider both the environments- homogeneous resource environment and heterogeneous resource environment.

In [8], Seung-Yeon You proposed a task scheduling algorithm for heterogeneous resource environment. In his model the tasks are considered as DAG architecture, each level in DAG the jobs are parallel and executed simultaneously on different resources. His algorithm chooses a resource based on its communication delay and execution time.

In [12], Jiayi Zhou proposes a scheduling model by introducing a dynamic resource broker and making use of fuzzy logic in scheduling algorithm. With the proposed algorithm, response time is reduced and throughput is increased. In all the above mentioned works the computing power of the resource is already known, and the nature of tasks/jobs is also known; either they depend on each other or they are independent of each other. Researchers mainly use already known processing power of resources to achieve their goal by performing load balancing.

REFERENCES

[1] I. Foster, C Kesselman. 1999. The Grid: Blueprint for a new computing infrastructure. Morgan Kaufmann Publishers. San Francisco, USA.

[2] R. Buyya, D. Abramson, J.Giddy. 2000. Nimrod/G: An Architecture for a Resource Management and Scheduling System in a Global Computation Grid. International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia 2000). Beijing, China. IEEE Computer Society Press, USA.

[3] M. Baker, R. Buyya and D. Laforenza. 2002. Grids and Grid Technologies for Wide-area Distributed Computing. in J. of Software-Practice Experience. Vol.32, No.15, pp. 1437-1466.

[4] N. Muthuvelu, J. Liu, N. L. Soe, S.r Venugopal, A. Sulistio and R. Buyya. 2005. A Dynamic Job Grouping-Based Scheduling for Deploying Applications with Fine-Grained Tasks on Global Grids. Proceedings of the 3rd Australasian Workshop on Grid Computing and e-Research (AusGrid 2005), Newcastle, Australia. 41-48.

[5] http://www.garudaindia.in/tech research.asp.

[6] Yagoubi, B., and Slimani. Task Load Balancing Strategy for Grid Computing. Journal of Computer Science, pp. 186-194.

[7] J. Cao, D. P. Spooner, S. A. Jarvis, G. R. Nudd. 2005. Grid load balancing using intelligent agents. Future Generation Computer Systems 21 (1) (2005) 135149.

[8] S. You, H. Kim, D. Hwang, S. Kim. 2004. Task Scheduling Algorithm in GRID Considering Heterogeneous Environment. in The 2004 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'04),Monte Carlo Resort, Las Vegas, Nevada, USA. pp. 240-245.

[9] N. T. Loc and S. Elnaffar. 2007. A dynamic scheduling algorithm for divisible loads in grid environments. Journal of Communications, vol. 2, no. 4. pp. 5764.

[10] K. Somasundaram, S. Radhakrishnan. 2009. Task resource allocation in grid using swift scheduler. Jounal of Computers, Communication Control, vol. 4, no. 2. pp. 158-166.

[11] A. J. S. Santiago, A. J. Yuste, J. E. M. Exposito, S. G. Galan, J. M. M. Marin, and S. Bruque. 2009. A Dynamic-Balanced Scheduler for Genetic Algorithms for Grid Computing. 2009. WSEAS Transactions on scheduling algorithms for MEMS-based storage, Computers, Vol.8, No.1.

[12] Zhou, J., Yu, K.M., Chou, Ch.H., Yang, L.A., Luo, Zh.J. 2007. A Dynamic Resource Broker and Fuzzy Logic Based Scheduling Algorithm in Grid Environment. ICAN-NGA, 604613.

[13] Cormen TH, Leiserson CE, Rivest RL. 2001. Introduction to algorithms 2nd edition. MIT and McGraw-Hill Book Company, Boston Massachusetts. cp. 16. 370-403.

[14] Rajkumar Buyya, Manzur Murshed. 2002. GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing. Technical Report, Monash University, Nov. 2001. To appear in the Journal of Concurrency and Computation: Practice and Experience (CCPE). pp. 1-32.

[15] William J. Stewart. 2009. Probability, Markov Chains, Queues, And Simulation: The Mathematical Basis of Performance Modeling. Princeton University Press. cp. 17. 625-646.

[16] Volker Hamscher, Uwe Schewiegelshohn, Achim Streit, Ramin Yahyapour. 2000. Evaluation of job-Scheduling Strategies for Grid Computing. in 1st IEEE/ACM International Workshop on Grid Computing (Grid 2000), Berlin. Lecture Notes in Computer Science (LNCS), Springer, Berlin, Heidelberg, New York. pp. 191-202.

[17] Ammar H. Alhusaini, Viktor K. Prasanna,C.S. Raghavendra. 1999. Unified Resource Scheduling Framework for Heterogeneous Computing Environments. in Proceedings of the Eighth Heterogeneous Computing Workshop, San Juan, Puerto Rico. pp. 156-165.

[18] Cong Liu, Sanjeev Baskiyar and Shuang Li. 2007. A General Distributed Scalable Peer to Peer for Mixed Tasks in Grids. High Performance Computing-HiPC 2007. ISBN:978-3-540-77219-4. 320-330.

[19] KM Yu, ZJ Luo, CH Chou, CK Chen, J Zhou. 2007. A fuzzy neural network based scheduling algorithm for job assignment on computational grid. NBiS 2007. LNCS 4658. pp. 533-542.

[20] V. Subramani, R. Kettimuthu, S. Srinivasan, and P. Sa-paper. Dayappan. 2002. Distributed job scheduling on computational grids using multiple simultaneous requests. In Proceedings of the 11th High Performance Distributed Computing.

[21] http://www.gridbus.org/publications-years.html.