OPTIMIZATION GRID SCHEDULING WITH PRIORITY BASE AND BEES ALGORITHM

MOHAMMED SHIHAB AHMED

UNIVERSITI UTARA MALAYSIA

2014

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

ABSTRAK

Pengaturcaraan grid bergantung kepada perkongsian skala besar di dalam rangkaian yang berhubung dengan sendiri seperti internet. Oleh itu, kebanyakan pengkaji dan para cendekiawan pengaturcaraan grid telah bertumpu kepada jadual tugasan yang juga dianggap sebagai isu-isu NP-Complete. Penyelidikan ini bertujuan untuk mengoptimumkan jadual awal bagi pengaturcaraan grid dengan menggunakan algaritma lebah. Algaritma moden sedia maklum dengan penyelidikan ini. Prosedur yang dicadangkan bermaksud bahawasanya algaritma yang baru dibangunkan boleh melaksanakan jadual tugasan grid sementara ia mengira keutamaan dan tarikh akhir untuk mengurangkan masa yang diperlukan untuk melengkapkan sesuatu tugasan. Purata masa menunggu bagi persekitaran grid boleh dikurangkan dan menerusi pengurangan ini, secara tak langsung dapat menghasilkan peningkatan pemprosesan persekitaran.

Key words: grid, pengaturcaraan, mengoptimumkan, algaritma lebah, masa

ABSTRACT

Grid computing depends upon sharing large-scales in a network that is widely connected within itself such as the Internet. Therefore, many grid computing researchers and scholars have focused on task scheduling, which is considered one of the NP-Complete issues. The main aim of this current research to propose an optimization of the initial scheduler for grid computing using the bees algorithm. Modern algorithms informed this research. The suggested procedure means that a newly developed algorithm can implement the schedule grid task while accounting for priorities and deadlines to decrease the completion time required for the tasks. The average waiting time of the grid environment can be minimized, and this minimization, in turn, creates an increase in the throughput of the environment.

Key words: grid, computing, optimization, bees algorithm, waiting time

Acknowledgement

I would like to express my deep and sincere gratitude to my supervisor, Associate Dr. Massudi bin Mahmuddin for the patience, humble supervision and advice I received from her in the course of this project. The supervision and support that she gave truly helped the progression and smoothness of this research.

Many thanks to my friend, Muhammed Shiban, for supporting me during my studies, and for being the family that I needed during my stay in Malaysia. My research would have not been possible without his help.

Furthermore, I would like to thank Wael Ali and Karena Hiyoshi for enriching me with their knowledge, and for encouraging me and supporting me throughout my research.

Last but not least, I would also like to thank my parents, Shihab Ahmed and Majda Abdul Jabbar. They were my role models and the best example for me to follow in my life. They never stop supporting me through my whole life and without them I would have never been what I am today. I want to thank my siblings, Ahmed, Abdul Jabbar, Amani, and Alia for their constant demonstrations of love and continuous moral support throughout my years of study.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACO	Ant Colony Optimization
BCO	Bee Colony Optimization
EDF	Earliest Deadline first
ERD	Earliest Release Date
FCFS	First Come First Serve
GA	Genetic Algorithm
GIS	Grid information service
GRAM	Grid Resource Allocation and Management
GS	Grid Scheduler
HAT	Hybrid algorithm technique
HAST	Hybrid algorithm search technique
HC	Hill climbing
IT	Information Technology
LJF	Longest Job First
LM	The Launching and Monitoring
LRM	Local Resource Manager
MTTD	Minimum Time to Due Date
NP	Nondeterministic Polynomial time
OGSA	Open Grid Services Architecture
P2P	Peer-To-Peer
PBBS	Priority Based Bee Scheduling

PSO	Particle Swarm Optimization
QOS	Qualities of services
QRC	Qualifying Resource Collection
SA	Simulated annealing
SJF	Shortest Job First
TS	Taboo search
VO	Virtual Organization
WIP	Work-in-process

CHAPTER ONE

INTRODUCTION

Grid computing is a type of computing that depends on sharing a large-scale network that is widely connected within itself such as in the Internet. [1] Researchers have suggested that grid and cluster computing are examples of different ways of starting a distributed system. A distributed system is way of connecting many computers in order to give them the ability to communicate within a computer network. Having multiple computers or workstations in cluster computing joined by local networks allows them to create distributed applications. Due to their limits, fixed-area applications in cluster computing are inflexible [2].This particular disadvantage has led to the suggestion that grid computing could help to solve this problem. Grid computing is built based on combining numerous resources from several geographic locations. This combination of resources from several geographic locations differentiates grid computing from cluster computing and conventional distributed computing. Different requirements and constraints exist for computation grid compared with those in traditional high performance computing systems [3].

True standardization has been developed to meet critical industrial requirements and so that grid computing technology can be enhanced. The Global Grid Forum started in 1998 as an international community and standards organization. The main responsibility of this organization was to develop different standardization activities [4]. Open Grid Services Architecture (OGSA) was established as another standard

The contents of the thesis is for internal user only

REFERENCE

[1] Foster, I., N.R. Jennings, and C. Kesselman. Brain meets brawn: Why grid and agents need each other. in Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1. 2004. IEEE Computer Society.

[2] Anderson, D.P. Boinc: A system for public-resource computing and storage. in Grid Computing, 2004. Proceedings. Fifth IEEE/ACM International Workshop on. 2004. IEEE.

[3] Baker, M., R. Buyya, and D. Laforenza, Grids and Grid technologies for wide-area distributed computing. Software: Practice and Experience, 2002. **32**(15): p. 1437-1466.

[4] Brown, M., et al., The International Grid (iGrid): Empowering global research community networking using high performance international Internet services. INET '99, San Jose, 1998.

[5] Östberg, P.-O. and E. Elmroth, GJMF—a composable service-oriented grid job management framework. Future Generation Computer Systems, 2013. **29**(1): p. 144-157.

[6] Buncic, P. and F. Carminati, A Discussion on Virtualisation in GRID Computing, in From the Web to the Grid and Beyond. 2012, Springer. p. 155-175.

[7] Simion, B., et al., A hybrid algorithm for scheduling workflow applications in grid environments (icpdp), in On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS. 2007, Springer. p. 1331-1348.

[8] Rimal, B.P., et al., Architectural requirements for cloud computing systems: an enterprise cloud approach. Journal of Grid Computing, 2011. **9**(1): p. 3-26.

[9] Prakash, S. and D.P. Vidyarthi, Maximizing availability for task scheduling in computational grid using genetic algorithm. Concurrency and Computation: Practice and Experience, 2014.

[10] Somasundaram, T.S., et al., Semantic-enabled CARE Resource Broker (SeCRB) for managing grid and cloud environment. The Journal of Supercomputing, 2014: p. 1-48.

[11] Beloglazov, A., J. Abawajy, and R. Buyya, Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Generation Computer Systems, 2012. **28**(5): p. 755-768.

[12] Pooranian, Z., et al., Gloa: a new job scheduling algorithm for grid computing. IJIMAI, 2013. **2**(1): p. 59-64.

[13] Zhu, Q. Grid Resource Scheduling Algorithm Based on Improved DCC Strategy. in Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 3. 2014. Springer.

[14] Casanova, H., et al., Mapping applications on volatile resource. International Journal of High Performance Computing Applications, 2014: p. 1094342013518806.

[15] Garg, R. and A.K. Singh, Enhancing the Discrete Particle Swarm Optimization based Workflow Grid Scheduling using Hierarchical Structure. International Journal of Computer Network and Information Security (IJCNIS), 2013. **5**(6): p. 18.

[16] Bhatt, K. and M. Bundele, Review Paper on PSO in workflow scheduling and Cloud Model enhancing Search mechanism in Cloud Computing. IJIET-International Journal of innovations in engineering and technology, 2013. **2**(3).

[17] Azmi, Z.R.M., et al., Performance Comparison of Priority Rule Scheduling Algorithms Using Different Inter Arrival Time Jobs in Grid Environment. International Journal of Grid and Distributed Computing, 2011. **4**(3): p. 61-70.

[18] Abdullah, M. and M. Othman, An Improved Genetic Algorithm for Job Scheduling in Cloud Computing Environment. AWERProcedia Information Technology and Computer Science, 2013. **2**.

[19] Agarwal, A. and P. Kumar, Multidimensional Qos Oriented Task Scheduling In Grid Environments. International Journal of Grid Computing & Applications (IJGCA), 2011. **2**(1): p. 28-37.

[20] Vykoukal, J., M. Wolf, and R. Beck. Does Green IT Matter? Analysis of the Relationship between Green IT and Grid Technology from a Resource-Based View Perspective. in PACIS. 2009.

[21] Bagchi, S. Simulation of grid computing infrastructure: challenges and solutions. in Proceedings of the 37th conference on Winter simulation. 2005. Winter Simulation Conference.

[22] Foster, I. and C. Kesselman, Computational grids, in Vector and Parallel Processing—VECPAR 2000. 2001, Springer. p. 3-37.

[23] Krauter, K., R. Buyya, and M. Maheswaran, A taxonomy and survey of grid resource management systems for distributed computing. Software: Practice and Experience, 2002. **32**(2): p. 135-164.

[24] Buyya, R., Economic-based distributed resource management and scheduling for grid computing. arXiv preprint cs/0204048, 2002.

[25] Buyya, R., S. Chapin, and D. DiNucci, Architectural models for resource management in the grid, in Grid Computing—GRID 2000. 2000, Springer. p. 18-35.

[26] Ferreira, L., et al., Introduction to grid computing with globus. 2003: IBM Corporation, International Technical Support Organization.

[27] Caminero, A., C. Carrión, and B. Caminero. Designing an entity to provide network QoS in a Grid system. in Proc. of the 1st Iberian Grid Infrastructure Conference (IberGrid). 2007.

[28] Dong, F. and S.G. Akl, Scheduling algorithms for grid computing: State of the art and open problems. School of Computing, Queen's University, Kingston, Ontario, 2006.

[29] Naik, V.K., et al. On-line evolutionary resource matching for job scheduling in heterogeneous grid environments. in Parallel and Distributed Systems, 2006. ICPADS 2006. 12th International Conference on. 2006. IEEE.

[30] Szymanski, T.L., the fit of certified disability management specialists'(cdms) knowledge domains with minnesota's qualified rehabilitation consultants'(qrcs') competencies. 2002, university of wisconsin.

[31] Kurowski, K., J. Nabrzyski, and J. Pukacki. User preference driven multiobjective resource management in grid environments. in Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on. 2001. IEEE.

[32] Deng, Y., F. Wang, and A. Ciura, Ant colony optimization inspired resource discovery in P2P Grid systems. The Journal of Supercomputing, 2009. **49**(1): p. 4-21.

[33] Zhou, P., et al., Genetic characterization of Toxoplasma gondii isolates from pigs in China. Journal of Parasitology, 2010. **96**(5): p. 1027-1029.

[34] Abu-Rukba, R.a.O., Decentralized Resource Scheduling in Grid/Cloud Computing. 2013.

[35] Robert, Y. and F. Vivien, Introduction to scheduling. 2010: CRC Press.

[36] Laili, Y., et al., A study of optimal allocation of computing resources in cloud manufacturing systems. The International Journal of Advanced Manufacturing Technology, 2012. **63**(5-8): p. 671-690.

[37] Dongarra, J. and A. Lastovetsky, An overview of heterogeneous high performance and grid computing. Engineering the Grid: Status and Perspective, 2006: p. 1-25.

[38] Zhu, Y., M. Li, and C. Weng. Ant Algorithm with Execution Quality Based Prediction in Grid Scheduling. in ChinaGrid Annual Conference, 2009. ChinaGrid'09. Fourth. 2009. IEEE.

[39] Zhu, Y., A survey on grid scheduling systems. Department of Computer Science, Hong Kong University of science and Technology, 2003.

[40] Korkhov, V.V., Hierarchical resource management in grid computing. 2009.

[41] Gil, Y., et al., Artificial intelligence and grids: Workflow planning and beyond. Intelligent Systems, IEEE, 2004. **19**(1): p. 26-33.

[42] Priya, S.B., M. Prakash, and K. Dhawan. Fault tolerance-genetic algorithm for grid task scheduling using check point. in Grid and Cooperative

Computing, 2007. GCC 2007. Sixth International Conference on. 2007. IEEE.

[43] Dong, F., Workflow scheduling algorithms in the grid. 2009.

[44] Pujari, A.K., Data mining techniques. 2001: Universities press.

[45] Kaur, R., T. Kaur, and H. Kaur, Scheduling in Grid Computing Environment. International Journal, 2013. **3**(6).

[46] Tonellotto, N., R. Yahyapour, and P. Wieder, A proposal for a generic grid scheduling architecture, in Integrated Research in GRID Computing. 2007, Springer. p. 227-239.

[47] Sacerdoti, F.D., et al. Wide area cluster monitoring with ganglia. in Cluster Computing, 2003. Proceedings. 2003 IEEE International Conference on. 2003. IEEE.

[48] Wolski, R., N.T. Spring, and J. Hayes, The network weather service: a distributed resource performance forecasting service for metacomputing. Future Generation Computer Systems, 1999. **15**(5): p. 757-768.

[49] Foster, I., et al., Grid services for distributed system integration. Computer, 2002. **35**(6): p. 37-46.

[50] Ranganathan, K. and I. Foster. Decoupling computation and data scheduling in distributed data-intensive applications. in High Performance Distributed Computing, 2002. HPDC-11 2002. Proceedings. 11th IEEE International Symposium on. 2002. IEEE.

[51] Trout, R.C., Complementary concurrent cooperative multi-processing multi-tasking processing system using shared memories with a minimum of four complementary processors. 1996, Google Patents.

[52] Gupta, A.K. and A.I. Sivakumar, Job shop scheduling techniques in semiconductor manufacturing. The International Journal of Advanced Manufacturing Technology, 2006. **27**(11-12): p. 1163-1169.

[53] Raman, N. and F. Brian Talbot, The job shop tardiness problem: A decomposition approach. European Journal of Operational Research, 1993. **69**(2): p. 187-199.

[54] Fujimoto, R.M., Parallel discrete event simulation. Communications of the ACM, 1990. **33**(10): p. 30-53.

[55] Calheiros, R.N., et al., CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software: Practice and Experience, 2011. **41**(1): p. 23-50.

[56] Buyya, R. and M. Murshed, Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. Concurrency and Computation: Practice and Experience, 2002. **14**(13-15): p. 1175-1220.

[57] Ramos, V., C. Fernandes, and A.C. Rosa, Societal implicit memory and his speed on tracking extrema in dynamic environments using self-regulatory swarms. Journal of Systems Architecture, Farooq M. and Menezes R.(Eds.), special issue on Nature Inspired Applied Systems, Elsevier, Summer, 2006.

[58] Bonchio, M., et al., Bio-inspired oxidations with polyoxometalate catalysts. Journal of Molecular Catalysis A: Chemical, 2006. **251**(1): p. 93-99.

[59] Li, Y., A bio-inspired adaptive job scheduling mechanism on a computational grid. International Journal of Computer Science and Network Security (IJCSNS), 2006. **6**(3): p. 1-7.

[60] Zhou, Y., et al. Data Scheduling Strategy in P2P VoD System Based on Adaptive Genetic Algorithm. in Proceedings of the 2012 International Conference on Information Technology and Software Engineering. 2013. Springer.

[61] Sambridge, M., A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization. Geophysical Journal International, 2014. **196**(1): p. 357-374.

[62] Mashinchi, M.H., M.A. Orgun, and W. Pedrycz, Hybrid optimization with improved tabu search. Applied Soft Computing, 2011. **11**(2): p. 1993-2006.

[63] Martincová, P. and M. Zábovský, Comparison of simulated Grid scheduling algorithms. Systemova Integrace, 2007. **4**: p. 69-75.

[64] Singh, M., S. Sholliya, and P. Gupta, Scheduling in Grid Computing–a Review. 2014.

[65] Lin, J.-N. and H.-Z. Wu, Scheduling in grid computing environment based on genetic algorithm. Journal of computer research and development, 2004. **41**(12): p. 2195-2199.

[66] Gonçalves, J.F., J.J. de Magalhães Mendes, and M.c.G. Resende, A hybrid genetic algorithm for the job shop scheduling problem. European journal of operational research, 2005. **167**(1): p. 77-95.

[67] Kennedy, J.F., J. Kennedy, and R.C. Eberhart, Swarm intelligence. 2001: Morgan Kaufmann.

[68] Zhang, L., et al., A task scheduling algorithm based on pso for grid computing. International Journal of Computational Intelligence Research, 2008. **4**(1): p. 37-43.

[69] Yan, H., et al. An improved ant algorithm for job scheduling in grid computing. in Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on. 2005. IEEE.

[70] Li, J. and W. Zhang. Solution to multi-objective optimization of flow shop problem based on ACO algorithm. in Computational Intelligence and Security, 2006 International Conference on. 2006. IEEE.

[71] Burke, E., et al. An ant algorithm hyperheuristic for the project presentation scheduling problem. in Evolutionary Computation, 2005. The 2005 IEEE Congress on. 2005. IEEE.

[72] Hanani, A., S. Nourossana, and A. Rahmani. Solving the scheduling problem in multi-processor systems with communication cost and precedence using bee colony system. in Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on. 2010. IEEE.

[73] Mousavinasab, Z., R. Entezari-Maleki, and A. Movaghar, A bee colony task scheduling algorithm in computational grids, in Digital Information Processing and Communications. 2011, Springer. p. 200-210.

[74] Wong, L.-P., M.Y.-H. Low, and C.S. Chong. A bee colony optimization algorithm for traveling salesman problem. in Modeling & Simulation, 2008. AICMS 08. Second Asia International Conference on. 2008. IEEE.

[75] Pham, D., et al. The bees algorithm–a novel tool for complex optimisation problems. in Proceedings of the 2nd Virtual International Conference on Intelligent Production Machines and Systems (IPROMS 2006). 2006.

[76] Karaboga, D. and B. Basturk, On the performance of artificial bee colony (ABC) algorithm. Applied soft computing, 2008. **8**(1): p. 687-697.

[77] Zhang, C., D. Ouyang, and J. Ning, An artificial bee colony approach for clustering. Expert Systems with Applications, 2010. **37**(7): p. 4761-4767.

[78] Carretero, J. and F. Xhafa, Use of genetic algorithms for scheduling jobs in large scale grid applications. Technological and Economic Development of Economy, 2006. **12**(1): p. 11-17.

[79] Singh, P., V. Singh, and A. Pandey, Analysis and Comparison of CPU Scheduling Algorithms.

[80] Zhao, L., et al. A Flexible Resource Publishing Framework for Eligible Subsystems Orchestration and Efficient Requests Scheduling. in Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 3. 2014. Springer.

[81] Tang, W., et al., Toward balanced and sustainable job scheduling for production supercomputers. Parallel Computing, 2013. **39**(12): p. 753-768.

[82] Jackson, D., Q. Snell, and M. Clement. Core algorithms of the Maui scheduler. in Job Scheduling Strategies for Parallel Processing. 2001. Springer.