Dynamic Resource Allocation: A comparison of First Price Sealed Bid and Vickrey Auctions

Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao, & William Okello-Odongo*
School of computing and informatics,
University of Nairobi

Bernard Manderick, & Ann Nowé
Computational Modeling Laboratory (COMO),
Free University of Brussels

Resource allocation involves assigning the resources to the resource users. Where the resources and the resource users do not change with time, the resource allocation problem can be solved as a scheduling problem. Where the resources and the resource users change with time then some different allocation mechanisms are needed. In this paper an environment is considered in which the resources emerge and the resource requests also emerge unpredictably as time goes. The resources emerge with different capacities and in the same way the resource requests also emerge from the users with different demands as time goes. In this dynamic environment the resource allocation performance of the first price sealed bid and the Vickrey auctions are explored and compared. The system allocation performance is measured using the number of the emerging resource provisions and the resource requests that get matched. The simulation results show that there is no performance difference between the two mechanisms apart from the fact that the Auctioneer's earnings are higher for the first price sealed bid auction.

Categories and Subject Descriptors: [Artificial Intelligence Applications]:- *Multi-Agent Systems*;
General Terms: Agents, Auctions, Grid Computing, Resource Allocation, Scheduling.

IJCIR Reference Format:

Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao, William Okello-Odongo, Bernard Manderick, and Ann Nowé. Game Theoretic Multi-Agent Systems Scheduler for Parallel Machines. International Journal of Computing and ICT Research, Special Issue Vol. 1, No. 1, pp. 28-34. http://www.ijcir.org/Special-Issuevolume1-number1/article 4.pdf.

---

1. INTRODUCTION

Resource allocation is the process of assigning the resources to the resource users. In a static environment the number of the resources and the number of the resource users do not change. The resource allocation problem can, therefore, be solved as a scheduling problem. The resources can, however, change with time

* Author's Address: Elisha T. O. Opiyo, Erick Ayienga, Katherine Getao, William Okello-Odongo, School of Computing and Informatics, University of Nairobi.
Bernard Manderick, Ann Nowé, Computational Modeling Laboratory, Free University of Brussels

International Journal of Computing and ICT Research, ISSN 1818-1139 (Print), ISSN 1996-1065 (Online),), Special Issue Vol.1, No.1 pp. 28-34, October 2008.

in different ways. The resource capacities can be changing with time. The resources can also change by being available at different times with different capacities. Consider a situation in which there are three computers each with the different types of packages that are used through online access by some ten people in a small organization. At any moment there may be a number of requests for the packages. If there are only three requests then there may be no problem as the requests can be immediately satisfied if no additional conditions are imposed. At some other time there could be, say, five requests. This requires making a decision on whom to allocate the computers immediately and who should wait. With time the number of requests change and the number of computers may also change. This is a dynamic resource allocation scenario. The computers are the resources and the workers are the resource users. This paper presents an exploration of the performance of the allocation process when it is managed by the first price sealed bid, or the Vickrey auctions. It is assumed that the resources do not fail and a resource user is not transferred from one resource to another for the sake of continuity (i.e. no pre-emption).

Auctions involve buying and selling items by offering them for bids, handling the bids, and then allocating the item to the one whose bid is the highest. The goods change hands from the auctioneers who manage the bids and represent the resource owners to the bidders who represent the resource consumers. The *first price sealed bid auctions* are those in which the proposed bids are private in that the participants do not see each other's bids. The highest bidder is allocated the item and that bidder pays the amount that was in the bid. The *Vickrey auctions* are also private but the highest bidder gets the item but pays the amount that was proposed by the second highest bidder [Wooldrige 2002]. *Grid computers* are virtual computers that are constructed out of flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources [Forster et al. 2001]. Although they involve many computing elements that are on the network their individual users think that they are each using a single computer. *Pre-emption* is a notion that is used to describe a situation where a job that is being executed is interrupted and continued at a later time on the same or a different machine. The job may also simply be transferred from one machine to another. *Schedules* are plans for performing work or achieving an objective, specifying the order and allotted time for each activity; *or* a program of events or appointments expected in a given time [Houghton 2000]. *Agents* are autonomous entities in the environment that can sense and react to the environmental changes. *Multi-agent systems* are those systems that are made up of several agents [Wooldrige et al. 2000].

Scheduling has been identified as one of the grid computing challenging issues [Buyya 2002]. Foster et al. [2001] also identify resource sharing as important activity in grid computing. The resources should be assigned to users so that the users can work smoothly and harmoniously. In this paper the dynamic resource allocation process is considered. The resources and the users emerge and leave on their own in an unpredictable manner as time goes on. This is the basis of the dynamism. The grid computing problem can be structured as consisting of two sets one set consists of the resource users and another set consists of the resource providers requiring matching dynamically. Section 3 gives more details on this problem. This paper reports the use of the auction-based techniques for the dynamic resource allocation of the resources in a simplified environment. The dynamism is retained but pre-emption and resource failures are not allowed. The performance measurements are based on the time that the users and the resources take before they are served. Ideally a user wants to get a resource immediately while a resource seeks to be utilized immediately.

This paper continues with a presentation of the simplified resource allocation model and then gives the outline of the allocation algorithm. Related work is considered and the experimental set up is outlined. The summarized results are given, and discussed, and finally the conclusion is made with further areas of research pointed out.

2. THE RESOURCE ALLOCATION MODEL

In this section the auction-based allocation model and an outline of the allocation algorithm are presented. The resource allocation model that is investigated and is reported in this paper is a simplified version of an agent-based grid scheduling model of Opiyo et al. [2007]. In that model, the auctioneer, bank, registrar and service agents manage the resource allocation process as resource agents offer their facilities while the user agents make the requests for the resources. The simplified model that is used in this paper is found in Figure 1.

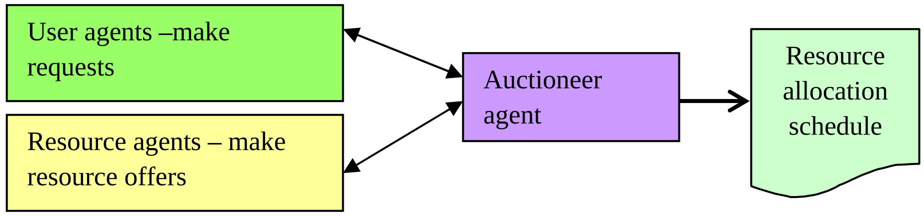

Figure 1. Auction-based resource allocation framework

The core of the resource allocation algorithm is as follows:

WHILE (the system is running)
Form a list of resource users and providers as they emerge
At some pre-determined time intervals {
For each resource compile a list of eligible users
Allocate the resource to most eligible user using auction (sealed bid/Vickrey)
Remove scheduled resources and users from their respective lists
Note important statistics
Release the current allocation schedule
}
ENDWHILE

3. RELATED WORK

In this section the scheduling problem as related to the grid computing context is outlined. The way this problem has been approached is considered. The grid computing resource allocation problem can be stated in the same way as the general scheduling problem. Consider the set of user tasks $\{T_1, T_2, T_3,,,,\ldots,, T_n\}$ and a set of resources $\{R_1, R_2, R_3,,,,\ldots,, R_m\}$. The problem is to find a way to assign resources to the tasks such that some objectives are optimized. In grid computing, however, due to the dynamism, uncertainties, wide distribution and resource autonomies the optimization factor is not pursued much. The grid computing problem has been approached using the techniques that are economic on the one hand and non-economic on the other [Buyya 2000]. The economic approaches use commodity market-based and auction-based techniques [Buyya 2000; Ferguson et al. 1999; Rajkumar et al. 1997; Wolski et al. 2000; Buyya et al. 2005]. The commodity-based approaches involve buyers and sellers where the prices are publicly agreed upon and may be demand driven. The auction-based methods involve the auctioneers and the bidders. The allocations are done by the auctioneer awarding the resources to the most qualified bidders according the rules of the respective auctions. The non-economic approaches to solving the resource allocation problem are responsible for the early grid computing platforms such as Condor. These non-economic allocation methods rely on scheduling techniques that emerged out of distributed operating systems [Berman et al. 2003c; Zoltan and Lazlo 1999; Ammar et al. 1999; Gabrielle et al. 2001; Cao et al. 2001; Tomasz et al. 2003; Walfredo et al. 2004; Adnan 2004; Bucur 2004]. These researchers were mostly concerned with delivering complete grid computing architectures using various scheduling techniques that are based on queues, service discovery or brokering. A non-economic approach that is close to this work is that of Getao and Opiyo [Getao and Opiyo 2007], in which data structures are used in managing the resource allocations process. This paper presents the work that differs from the other previous work in that the auction-based approach is used in the resource allocation. The focus in this paper is on the performance related issues that would arise in the grid computing context. The rate at which the emerging resource requests and resource offers are matched by the dynamic allocation mechanism is the performance matter that is being investigated. The other previous auction-based work has concentrated on the allocation of particular commodities such as the bandwidth but not concerned with the grid computing [Rosu et al. 1997; Ramanathan et al. 1999; Chandra et al. 2002; Ghosh et al. 2004; Georgiadis et al. 2006]. The other previous work involving auctions has also been concerned with the centralization and the decentralization issues. It has involved investigating the nature of the auction mechanisms that are strategy-proof and encourage truthful revelations [Nisan & Ronen 1999; Nisan & Ronen 2000]. The pioneering work of Chunlin and

Layuan [2008] concentrates more on the layered modeling of global optimization in grid computing with focus on the quality of service.

4. THE EXPERIMENT AND THE RESULTS

The objective is to investigate rate at which the resource users and the resource providers are matched by the system's dynamic allocation mechanism as they emerge. Some questions are used to guide the investigation. When a user is looking for a resource how long would the user wait before the resource request is met? Similarly, how long does it take a resource that has been availed for use before it begins to be used? What is the overall waiting time for the system? What does the auctioneer gain? The variables whose values are recorded include user efficiency, resource efficiency, overall efficiency, user waiting time efficiency, resource waiting time efficiency and auctioneer's efficiency. The *user efficiency* is the ratio of the total users whose requests are met to the total number of users expressed as a percentage. The *resource efficiency* is the ratio of the total resource providers whose offers are utilized to the total number of resource providers expressed as a percentage. The *overall efficiency* is the ratio of the total resource users and providers whose needs are met to the total number of users and providers expressed as a percentage. This variable acts as the systems performance metric. The user's waiting time efficiency is computed as: *(Total Length of Time - Cumulative Waiting Time for Users)*100/Total Length of Time*. The resource provider's waiting time efficiency is computed as: *(Total Length of Time - Cumulative Waiting Time for Resources)*100/Total Length of Time*. The *auctioneer's efficiency* is computed as the ratio of total amount earned to potential amount that can be earned expressed as a percentage.

The average values were obtained from all the sessions beginning with 10,000 rounds to 100,000 rounds for each dynamic allocation mechanism. The first price sealed bid auction allocation mechanism yielded the efficiencies above 85% for the user, resource, system, user waiting and resource waiting. The auctioneer efficiency was 100%. The Vickrey auction allocation mechanism yielded the efficiencies above 83% for the user, resource, system, user waiting and resource waiting. The auctioneer efficiency was about 70%. Figure 2 gives a comparison of the performances.

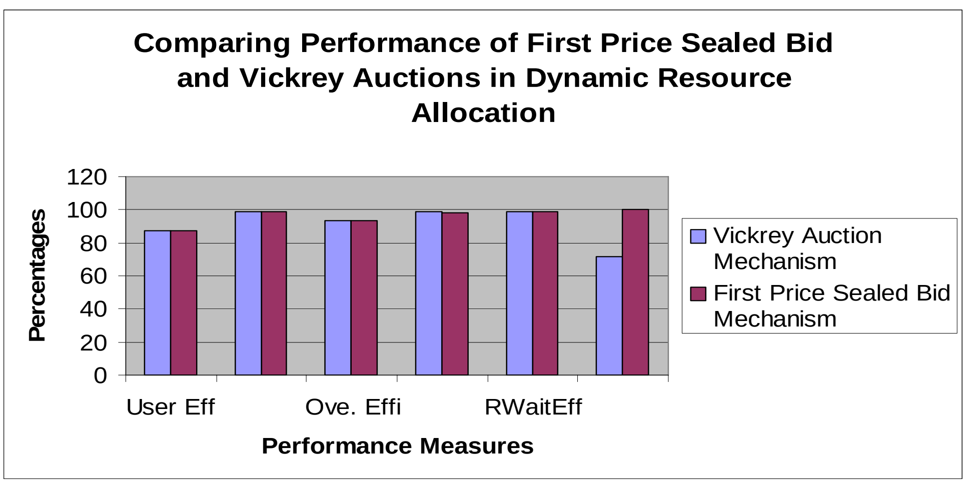

Figure 2: Comparing performances of the auction mechanisms

Legend: User Eff – User Efficiency; Res. Eff- Resource Efficiency; Ove. Effi – Overall Efficiency; UWaitEff – user waiting efficiency; RWaitEff – Resource waiting efficiency; Auc. Eff – Auctioneer Efficiency

5. DISCUSSION

The resource users are cleared at an efficiency rate above 83%. The resource provider's efficiency is above 95% for both mechanisms. The non attainment of 100% is explained by the fact that there will be some requests or offers that cannot be met. Some of the resources may not be used due to the non availability of the suitable resource users. Sometimes some resource users are always bidding lower than the other competing users. The overall performance in both cases is over 80%. In both mechanisms the mean waiting time efficiencies are above 95% that means that the emerging resource requests and resource offers do not wait for a long time before their needs are met. This is an indication of the potential that exists in the use of the auction-based allocation mechanisms in a dynamic environment such as the grid computing.

Figure 2 shows how the performances of the two mechanisms compare. The percentage difference between the corresponding values of the performance measurements are below 3 % except for the auctioneer's earning efficiency. The auctioneer's earning is more for the first price sealed bid allocation mechanism because in it the bidder pays the amount that is on the bid. This enables the auctioneer to collect all the possible revenue. For the Vickrey auction, on the other hand, the bidder pays the value of the second highest bidder which is most likely lower than the winner's bid price. The auctioneer therefore earns less than the possible total revenue from the winning bids. So apart from the auctioneer's earnings, all other performance parameters seem to be similar. The similarity of other performance values is explained by the similarity of the two allocation processes. Both are sealed bids. Both have the same allocation process. In both cases to get an eligible bidder the list of bidders must be in descending order of the bidding amounts. The process of picking the second highest bidder when the auctioneer is collecting payment in case of the Vickrey allocation mechanism does not seem to introduce any waiting overhead. This is because the process of picking payment from the highest bidder and from the second highest bidder need not take different lengths of time.

6. CONCLUSION

The results from these experiments point to the fact that the use of the first price sealed bid and the Vickrey allocation mechanisms in the grid computing environment is promising. Either of them may be used in managing the grid computing resource allocation tasks. Since both mechanisms deliver similar performance capabilities except for the auctioneer's revenue, the circumstances may determine which one to use. Where the resource owners insist on maximizing their revenues the first price sealed bid auction allocation mechanism may be selected. However, if truthful bidding is emphasized then the Vickrey auction allocation mechanism can be used. The auction-based resource allocation mechanisms such as the ones that have been compared can only be implemented where the resource requests are always accompanied by some numeric offers. The numeric offers should represent the valuations that are attached to the resources by those making the requests. These mechanisms can also work only under the circumstances where the resource requests and resource offers are periodically processed particularly to allow the auctioneers to perform the allocations. In computing environments there are periodic checks that enable the events to be recognized for servicing. As an item of further work, investigations can be conducted to find out the effect on performance of pre-emption, uncertainties of the resource providers and of the resource consumers. There is also need to investigate the effect of the length of the time that the auctioneer takes before the allocations are conducted.

7. ACKNOWLEDGMENTS

We are very grateful to the Flemish Inter-University Council: Institutional University Cooperation Programme (VLIR-IUC) of Belgium. They have supported our work. We also thank the VLIR Secretariat at the University of Nairobi, and the Free University of Brussels (VUB), for their support. We are thankful to the anonymous reviewers for helping in the improvement of this paper.

8. REFERENCES

ADNAN FIDA . 2004. *Scheduling Policies for Enterprise Wide Peer-to-Peer Computing*. In Proceedings of the 2003-2004 Grad Symposium. CS Dept, University of Saskatchewan, April 2004. Citeseer.ist.psu.edu/709611.html – last visited February, 2008.

AMMAR H. A., PRASANNA V. K., RAGHAVENDRA C.S. 1999. A Unified Resource Scheduling Framework for Heterogeneous Computing Environments. In *Proc. HCW'99, 8th Heterogeneous Computing Workshop*, pages 156--168, San Juan, Puerto Rico, April 1999.

BERMAN F., FOX G. C., HEY A. J. G. 2003. (Editors). *Grid computing: making the global infrastructure a reality*. Wiley publishers.

BERMAN, F., FOX, G.C., HEY, A.J.G. 2003b. The Grid: Past, Present, Future. In Berman, F., Fox, G.C., Hey, A.J.G.,(eds): *Grid Computing: Making the Global Infrastructure a Reality*. Wiley publishers, 9 – 50.

BERMAN F., WOLSKI R., CASANOVA H., CIRNE W., DAIL H., FAERMAN M., FIGUEIRA S., HAYES J., OBERTELLI G., SCHOPF J., SHAO G., SMALLEN S., SPRING ., SU A., DMITRII Z. 2003c. Adaptive Computing on the Grid Using AppLeS. In *IEEE Transactions On Parallel And Distributed Systems*, VOL. 14, NO. 4, APRIL 2003.

BUCUR ANCA. 2004. Performance Analysis of Processor Co-Allocation in Multicluster Systems. *A PhD Thesis*, Delft University of Technology.

BUYYA, R. 2002. Economic-Based Distributed Resource Management and Scheduling for Grid Computing. *PhD Thesis, Monash University* Melbourne Australia.

BUYYA R., D. ABRAMSON, AND S. VENUGOPAL, 2005. The Grid Economy. In *Proc. IEEE*, vol. 93, no. 3, March. 2005.

CAO J, KERBYSON D. J., NUDD G. R. 2001.Use of Agent-Based Service Discovery for Resource Management in Metacomputing Environment. In *Proc. 7th Int. Euro-Par Conf.*, Manchester, UK, LNCS 2150, 882-886, 2001.

CHANDRA A., W. GONG, AND P. SHENOY. 2002. Dynamic resource allocation for shared data centers using online measurements. *Technical Report TR02-30*, Department of Computer Science, University of Massachusetts, 2002. http://citeseer.ist.psu.edu/chandra03dynamic.html

CHUNLIN LI, LI LAYAN. 2008. Multi-level scheduling for global optimization in grid computing. In Computers and Electrical Engineering, Volume 34, pp. 202-22, Elsevier 2008.

FERGUSON D., C. NIKOLAOU, J. SAIRAMESH, AND Y. YEMINI. 1996. Economic models for allocating resources in computer systems. In Scott Clearwater, editor, *Market-Based Control: A Paradigm for Distributed Resource Allocation*, Scott Clearwater. In World Scientific, Hong Kong, 1996. http://citeseer.ist.psu.edu/ferguson96economic.html.

FOSTER, I., CARL, K., STEVEN, T. 2001. The Anatomy of the Grid: Enabling Scalable Virtual Organizations .2001. Reprint in Berman, F., Fox, G.C., Hey, A.J.G.,(eds): *Grid Computing: Making the Global Infrastructure a Reality*. Wiley publishers (2003) 171 – 191.

GABRIELLE A. ANGULO D., FOSTER I., LANFERMANN G.,LIU C., RADKE T ., SEIDEL E., SHALF J. 2001. *The Cactus Worm: Experiments with Dynamic Resource Discovery and Allocation in a Grid Environment*. www.cs.uchicago.edu/files/tr_authentic/TR-2001-28.pdf - last visited February, 2008.

GETAO KATHERINE, OPIYO E. T. O. 2007. A Grid Service Matrix Model for Resource Scheduling in Grid Computing Environments. In *1st International Conference in Computer Science and Informatics*, 5th – 7th February 2007, Nairobi, Kenya, School of Computing & Informatics, University of Nairobi, Kenya.

GEORGIADIS L,, NEELY M. J., TASSIULASL . 2006. Resource Allocation and Cross-Layer Control in Wireless Networks. In *Foundations and Trends in Networking* Vol. 1, No 1 (2006) 1–144.

GHOSH S., RAJKUMAR R., HANSEN J., LEHOCZKY J. 2004. *Integrated Resource Management and Scheduling with Multi-Resource Constraints*. www.cs.cmu.edu/People/hansen/papers/radar-sched.pdf - last visited February, 2008.

HOUGHTON. 2000. *The American Heritage® Dictionary of the English Language*, Fourth Edition, 2000, Houghton Mifflin Company. education. yahoo.com/reference/dictionary- last visited February 2008.

NG C., D. C. PARKES, AND M. SELTZER. 2003. Strategyproof Computing: Systems Infrastructures for Self-Interested Parties. *Technical report, Harvard University*, 2003. http://citeseer.ist.psu.edu/article/ng03strategyproof.html

NISAN N. RONEN A. 2000. Computationally Feasible VCG Mechanisms. In *ACM Conference on Electronic Commerce 2000*: 242-252

NISAN N., RONEN A . 1999. Algorithmic Mechanism Design. In *STOC 1999*: 129-140.

RAJKUMAR R., C. LEE, J. LEHOCZKY, AND D. SIEWIOREK. 1997. A Resource Allocation Model for QoS Management. In *Proceedings of the IEEE Real-Time Systems Symposium*, December 1997. http://citeseer.ist.psu.edu/rajkumar97resource.html

RAMANATHAN P., K. M. SIVALINGAM, P. AGRAWAL AND S. KISHORE. 1999. Dynamic resource allocation schemes during handoff for mobile multimedia wireless networks. In *IEEE Journal on Selected Areas in Communications*, 17(7), July 1999. http://citeseer.ist.psu.edu/ramanathan99dynamic.html

ROSU D., K. SCHWAN, S. YALAMANCHILI, AND R. JHA. 1997. On Adaptive Resource Allocation for Complex Real-Time Applications. 18th *IEEE Real-Time Systems Symposium,* Dec., 1997. http://citeseer.ist.psu.edu/article/rosu97adaptive.html

SYCARA K., S. ROTH, N. SADEH M. FOX. 1991. Distributed constrained heuristic search. In *IEEE Transactions on Systems, Man, and Cybernetics*, 21:1446--1461, 1991.

TOMASZ W., KURZYNIEC D, DRZEWIECKI D, SUNDERAM V. 2003. Resource Monitoring and Management in Metacomputing Environments. In *Proceedings 10th European PVM/MPI User's Group Meeting*, Venice, Italy, pp. 629-635, September 2003.

WALFREDO C., BRASILEIRO F., COSTA L., PARANHOS D., SANTOS-NETO E., ANDRADE N. 2004. *Scheduling in Bag-of-Task Grids*: The PAUÁ Case. http://citeseer.ist.psu.edu/cis?q=Luiz+Biondi+Neto, walfredo.dsc.ufcg.edu.br/ papers /paula.pdf - last visited February 2008.

WOOLDRIDGE MICHAEL. 2002. *An introduction to multiagent systems*. John Willey & Sons, Ltd.

WOLSKI R.,PLANK J. S., BREVIK J., BRYAN T. 2000. Analyzing Market-based Resource Allocation Strategies for the Computational Grid. In *International Journal of High Performance Computing Applications* ,2001, 258-281.

WOOLDRIDGE M., NICHOLAS R. JENNINGS, DAVID KINNY. 2000. *The Gaia Methodology for Agent-Oriented Analysis and Design, in Autonomous Agents and Multi-Agent Systems*, 3, 285-312. Kluwer Academic Publishers.

ZOLTAN J., LAZLO K. 1999. A Jini-based prototype Metacomputing framework. *Lecture Notes in Computer Science*, Vol. 1900, P. 1171--, 2001.