2009 International Symposium on Computing, Communication, and Control (ISCCC 2009)
Proc .of CSIT vol.1 (2011) © (2011) IACSIT Press, Singapore

Adaptive Task Scheduling Strategy for Economy Based Grid

Babar Nazir [1], Mohd Fadzil Hassan [2] and Halabi Hasbullah [3 +]

Department of Computer and Information Sciences
Universiti Teknologi PETRONAS, Bandar Seri Iskandar,
31750 Tronoh, Perak, Malaysia.

Abstract. An adaptive task scheduling strategy is proposed in an economy based grid to ensure consistent performance, despite of variation in individual resource performance (i.e. meeting user QoS requirement like budget and deadline). The proposed strategy maintains performance index of grid resources, which depicts the resources past performances. The performance index shows resource vulnerability towards variance in their past performance. Furthermore, in our approach the resource performance index based on how an assigned task is completed is dynamically updated based on how assigned task is completed by a particular grid resource. While scheduling job on resource(s), this performance index can be utilized by the broker in applying different heuristic like dummy task allocation and penalty/reward heuristic and/or both. A dummy task is some times sent in order to check the claims of the resource provider. In the presence of the variance in resource performance, the proposed strategy can effectively schedules grid jobs. As such, more jobs can be executed according to the user specified deadline and allotted budget. Hence, help in upholding trustworthiness of grid environment.

Keywords: Economy Based Grid, Grid Task Scheduling, Grid Resource Management, Distributed System.

1. Introduction

The term Grid computing [1-4], was introduced in the early 1990s by Ian Foster and Carl Kesselman. Grid computing is an approach of making the computation power of idle work stations available to remote grid users for the execution of their computation hungry jobs. This was intended to be done in the same pervasive fashion, as we access the electric power grid [5-7]. Grid computing uses idle computational power from different geographical places. It involves runtime aggregation of these resources in the form of Virtual Organization (VO) [8, 10], according to the needs of the job submitted by the grid user*.

The economy based grid [11-13] is a user-centric resource management and job scheduling approach. It offers incentives and profits to resource owners for contribution of their resources. It also gives users a dynamic environment to maximize their gains by relaxing QoS requirements such as budget and deadline. In this way, economy based grid computing provides a competitive environment that satisfies both the parties involved in the grid i.e. resource producers and resource consumers. Over the years, the availability of powerful PC and rapid growth of the Internet has made it possible to visualize economic base grid as the technology for the future.

The motivation of this paper is to present a grid scheduling strategy for economy based grid that can provide efficient QoS service to the user in varying grid resource performance over the time. This enables the economy grid environment to improve user satisfaction by efficiently fulfilling budget and deadline QoS requirement of the user(s). Furthermore, this also minimize the penalty paid [14] by the resource provider. The approach described in the paper provides a proactive measure of not assigning jobs to resource(s) which can not complete the job(s) within deadline and budget. In this way, it enables grid to uphold the faith of the

+ Corresponding author.
E-mail address: (babarnazir@gmail.com, mfadzil_hassan@petronas.com.my, halabi@petronas.com.my).

user by not compromising its QoS requirements, due to varying grid resources performance. The rest of this paper is organized as follows. Section 2 briefly explains different strategies for task scheduling strategy for economy-based Grid. In Section 3, explanation of proposed adaptive task scheduling strategy for economic grid is given. The final section includes conclusions and suggests future work.

2. Related work

In [11-13], economic based grid model is proposed. Economy grid has four main players namely Grid Resource Provider, Grid User, Grid Resource Broker (GRB) and Grid Information Service (GIS). Grid users submit their jobs to the grid resource broker by specifying their QoS requirements i.e. deadline in which users want their jobs to be executed and the budget which users have for the completion of jobs. The Grid Resource Broker consults the Grid Information Service (GIS) for the list of available resources. GRB then schedules user jobs on the best available resource by optimizing user QoS requirement. Result of the job was submitted to user upon successful completion of the job.

In [11], time optimization strategy was proposed, which do optimization on the base of time. List of available resources which can execute the job within user defined deadline and budget, were obtained from GIS. Resources were then sorted based on their processing capability and job was given to the fastest available resource, to ensure that job takes minimum time to execute. This optimization was used in deadline constrained user applications, where user primary demand was to execute the job within the hard deadline even at expense of more budgets but deadline can not be relaxed.

In [15-17], cost optimization strategy was proposed, which do optimization on the base of cost. List of available resources which can execute the job within user defined deadline and budget, were obtained from GIS. Resources were then sorted based on their on their processing cost and job was given to the fastest available resource, to ensure that job takes minimum cost to execute. This optimization was used in budget constrained user applications, where user primary demand, was to execute the job within the allotted budget even at expense of more deadline time, and budget can not be relaxed.

In [13], cost/time optimization strategy was proposed, which do optimization on the base of cost as well as time. List of available resources which can execute the job within user defined deadline and budget, were obtained from GIS. Resources were then sorted based on their on their processing cost. Further resources were sorted based on their processing capability and job was given to the cheapest and fastest available resource, to ensure that job consumes minimum budget to be executed using minimum time of deadline. This optimization was used in Deadline and Budget Constrained (DBC) user applications, where user primary demand was to execute the job within the allotted budget and user was not willing to relax the budget or to extend deadline.

In [18], information asymmetry problem in market grid was addressed. The paper emphasized that this problem would result in dishonest trades and grid user loss. Two mechanisms namely, reputation-aware aggregate mechanism and reputation-aware double auction were proposed. The paper also proposes a job scheduling strategies which use reputation value for each grid resource depicting resource reliability. Job allocation decision was based on heuristic that involves resource price as well as resource reputation.

In [19], incentives for grid resource provider and for grid user were presented by objective functions. A QoS scheduling algorithm is proposed which aims to optimize the performance of each individual grid user and the performance of the grid resources. Using dynamic programming, QoS scheduling optimization was decomposed in two sub optimizations of grid user's cost and grid resource provider's profit.

3. Adaptive Task Scheduling Strategy for Economy Based Grid

This section explains the proposed strategy (see figure1) that enables the system to adapt to varying performance of the resource by using the heuristic of dummy task allocation, using penalty reward heuristic [14] and/or both.

3.1 Proposed Scheduling Strategy Framework

The interaction between different components of economy based grid in the proposed scheduling strategy (see figure1) is as follows:

Grid user submits jobs to the grid resource broker by specifying its characteristics i-e number of tasks, task length, scheduling strategy, deadline and budget etc..

A grid resource is a member of a grid and it offers computing services to grid users. Grid users register themselves to Grid Information Server (GIS) of a grid by specifying the QoS requirements such as the cost of computation, deadline to complete the execution, the number of processors, speed of processing, internal scheduling policy, and time zone. A GIS contains information about all available grid resources with their computing capacity and cost at which they offer their services to grid users. All grid resources that join and leave the grid are monitored by GIS. Whenever a grid broker has jobs to execute, it consults GIS to give information about available grid resource. The Adaptive Performance Schedule Manager (AP Schedule Manager) maintains resource performance index history about grid resources. And updates (increment/decrement) performance index of a grid resource by receiving requests from broker. The Check Resource Schedule Manager (CR Schedule Manager) sends a dummy job to the grid resource to check the real computational capability of the grid resource.

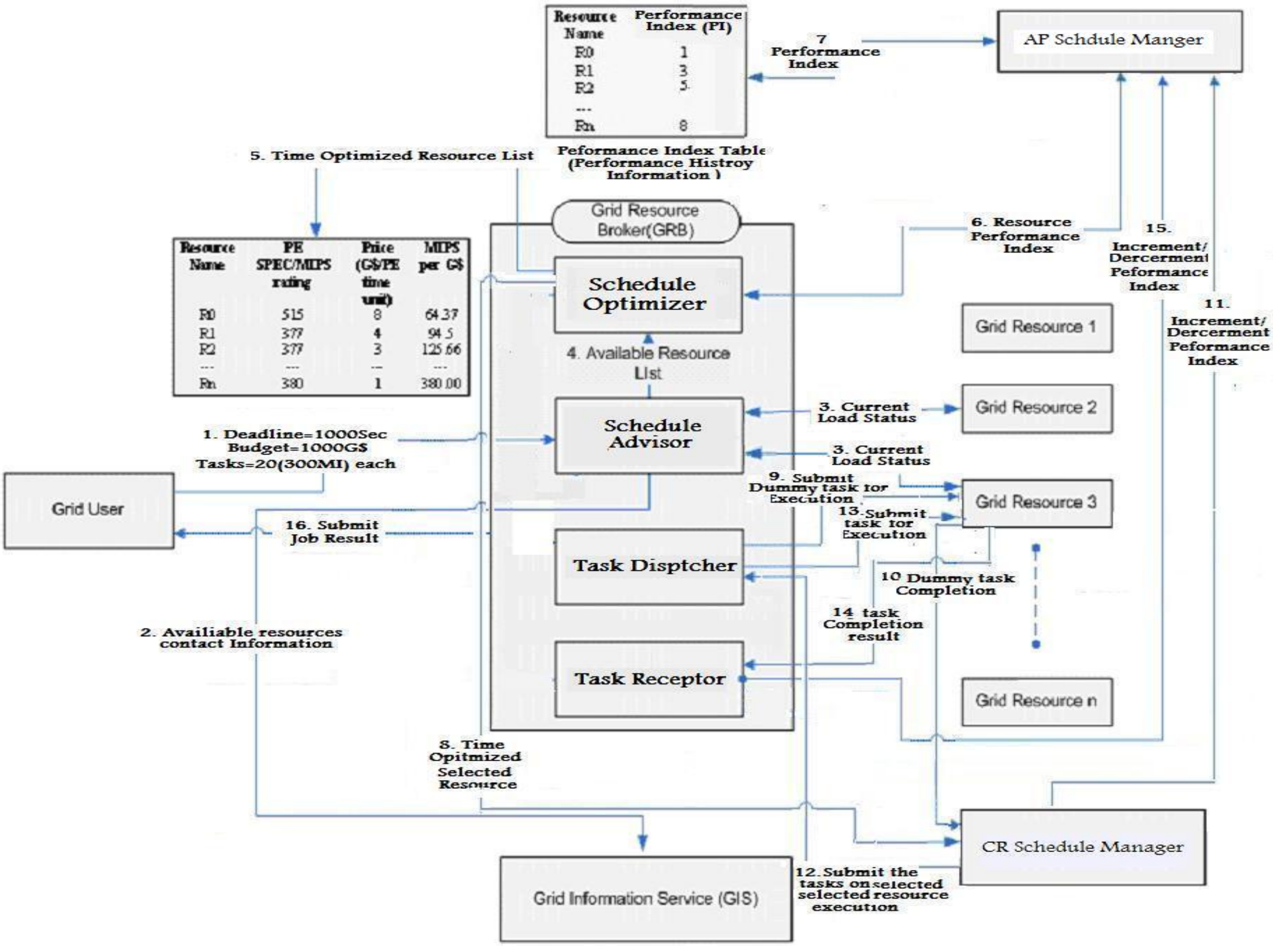

Fig.1. Interaction of different grid components in propose strategy

If the grid resource is able to complete that dummy job according to its mentioned computing capability. Job is dispatched to that resource, otherwise, it is sent to some other resource. When CR Schedule Manager receives the result of dummy task execution, it implements Algorithm 1 to take appropriate decision.

Algorithm 1:

1. IF CR Schedule Manager receives the dummy task completion result from the resource THEN
    a. CR Schedule Manager evaluates the performance of the dummy task according to the computational capability specified by the resource.

i. IF resource performance is inline with its capability
THEN
(1) Task is given to the task dispatcher to assign it to that resource.
ELSE
(1) Send a message to AP Schedule Manager to increment the performance index of resource that exuberate its resource capability.
END IF
END IF
2. EXIT

Grid Resource Broker (GRB) is an important entity of a grid. A grid resource broker is connected to an instance of a user. Each grid job (composed of tasks) is first submitted to its broker, which then schedules the grid job according to the user's scheduling policy. When GRB receives a grid job from a user, Schedule Advisor gets the contact information of available grid resources from the GIS and then requests the resources to send their current work load condition. Based on current work load condition of the resources, Schedule Optimizer applies time optimization strategy on the available resource and prepares a list of resources that can execute the task satisfying deadline and budget constraints based on time optimization strategy. The GRB then gets the performance index of the selected resources of the list from AP Schedule Manager. Depending on the performance index of the resource, the GRB implements Algorithm 2 to take appropriate decisions.

Algorithm 2:

P: Performance index of the selected grid resource

IF 0<= P < 1 THEN
The task is queued to that grid resource.
GO TO Step 6.
END IF
IF 2 <= P < 3THEN
The penalty reward heuristic [14], is used which specify the penalty. The resource have to pay, if it is unable to fulfill job deadline/budget parameters.
GO TO Step 6.
END IF
IF 4 <= P < 5 THEN
The resource is given to the CRManager, to send dummy task on that resource and check its resource capability. If satisfies this test, only then original job is scheduled to that resource.
GO TO Step 6.
END IF
IF 6 <= P < 7 THEN
The resource is given to the CR Schedule Manager, to send dummy task on that resource and check its resource capability.
IF grid resource dummy task execution satisfies this test THEN
Schedule original job to that resource by using penalty/reward task heuristic [14] to specify penalty. Which resource has to pay, if it is unable to fulfill job deadline/budget parameters.
ELSE
GO TO Step 6.
GO TO Step 6.
END IF
IF P >=7 THEN
Remove resource from the available resources list and label it as unavailable resource(i.e. no job is assigned) to that resource until:
It has passed CR Schedule Manager resource capability test, by executing the dummy job. This test is performed after a specific time interval.
IF resource is able to pass that test THEN
Include resource in available resource list. Other wise it has to wait for next time interval.
ELSE
Other wise it has to wait for next time interval.
END IF

Add the job to the unassigned job list and reapply the time optimization algorithm for this task on available resources.

END IF

EXIT

A Task Dispatcher dispatches the tasks one by one to the respective grid resource as listed in the queue by GRB. Task Receptor receives the task execution result from the grid resource where the task is dispatched by the dispatcher. The performance index (P) of a grid resource is updated by AP Schedule Manager using Algorithm 3.

1. IF task receptor receives the task completion result from the resource within the deadline specified and budget allotted.

THEN

a. IF P>=1

THEN

i. Send a message to AP Schedule Manager to decrement the performance index of resource that completes the assigned task.

END IF

b. GOTO Step3

END IF

2. IF task receptor receives the task completion result from the resource after the deadline specified and/or over consuming budget allotted.

a. Task acceptor evaluates the performance of the resource according to the SLA agreement specified. And grid resource is penalized according to the criteria defined in SLA agreements. That penalty is proportional to the time extent to which job is delayed from the actual deadline.

b. IF resource performance index>=7

THEN

i. Send a message to AP Schedule Manager to increment the performance index of resource that can not complete the assigned task.

END IF

END IF

3. EXIT

4. Conclusion

In this paper, we present a adaptive task scheduling strategy in order to ensure consistent performance, despite of variation in individual resource performance in an economy based grid environment. The proposed strategy maintains performance index of grid resources, which shows resource past performance and is depiction of resource vulnerability toward variation in performance in past. Based on this performance index we further used dummy task execution, penalty/reward heuristic or both, while making task allocation decision to the resource.

In the presence of the variance in resource performance, the proposed strategy can effectively schedules grid jobs. And able to execute more jobs successfully, within the specified deadline and allotted budget. It can also improve the overall execution time and minimizes the execution cost of grid jobs.

5. References

[1] I Foster, and C Kesselman, "The Grid: Blueprint for a New Computing Infrastructure", Morgan Kaufmann Publishers, USA, 1999.

[2] I Foster and C Kesselman, “The Grid 2: Blueprint for a New Computing Infrastructure”, Morgan Kaufman Publishers Inc., 2004.

[3] Madhu Chetty and Rajkumar Buyya, "Weaving Computational Grids: How Analogous Are They with Electrical Grids?", Computing in Science and Engineering (CiSE), IEEE Computer Society Press and American Institute of Physics, USA, ISSN 1521-9615, Vol. 4, Issue 4, pp. 61 - 71, July - August 2002.

[4] I Foster, C Kesselman, and S Tueke, “The anatomy of the grid: Enabling scalable virtual organizations”. International Journal of Supercomputing Applications, 2001.

[5] I Foster, C Kesselman, J Nick and S Tuecke, “The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration”, June 22, 2002.

[6] I Foster, "What is the Grid? A Three Point Checklist", GRID Today, July 20, 2002.

[7] Fran Berman, Geoffrey Fox, Anthony J.G. Hey,” Grid Computing: Making the Global Infrastructure a Reality”, Wiley Publishers, 2003

[8] Rajkumar Buyya, "Economic-based Distributed Resource Management and Scheduling for Grid Computing", Ph.D. Paper, Monash University, Melbourne, Australia, April 12, 2002.

[9] Rajkumar Buyya, David Abramson, and Srikumar Venugopal, "The Grid Economy, Special Issue on Grid Computing", Proceedings of the IEEE, Manish Parashar and Craig Lee, IEEE Press, New York, USA, ISSN 0018-9219, Vol. 93, Issue 3, pp. 698 - 714, March 2005.

[10] Manjuka Soysa, Rajkumar Buyya, and Baikunth Nath, "GridEmail: Economically Regulated Internet-based Interpersonal Communications", Advanced Parallel and Distributed Computing: Evaluation, Improvement and Practice, Yuanshun Dai, Yi Pan, Rajive Raje, Nova Science Publishers, USA, pp. 279 - 295, 2006.

[11] Rajkumar Buyya, David Abramson, Jonathan Giddy and Heinz Stockinger, "Economic Models for Resource Management and Scheduling in Grid Computing”, Concurrency and Computation: Practice and Experience (CCPE), Wiley Press, USA, Vol. 14, Issue 13-15, pp. 1507 - 1542, November - December 2002.

[12] Rajkumar Buyya, Manzur Murshed, David Abramson and Srikumar Venugopal, "Scheduling Parameter Sweep Applications on Global Grids: A Deadline and Budget Constrained Cost-Time Optimization Algorithm", Software: Practice and Experience (SPE), Wiley Press, New York, USA, Vol. 35, Issue 5, pp. 491 - 512, April 25, 2005.

[13] Rajkumar Buyya, Manzur Murshed and David Abramson, "A Deadline and Budget Constrained Cost-Time Optimization Algorithm for Scheduling Task Farming Applications on Global Grids", Proceedings of the 2002 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'02), Las Vegas, USA, June 24-27, 2002.

[14] Chee Shin Yeo and Rajkumar Buyya, "Service Level Agreement based Allocation of Cluster Resources: Handling Penalty to Enhance Utility", Proceedings of the 7th IEEE International Conference on Cluster Computing, Cluster 2005, IEEE CS Press, Los Alamitos, CA, USA, Boston, Massachusetts, USA, September 27-30, 2005.

[15] Cost Marcos Dias de Assuncao and Rajkumar Buyya, A Cost-Aware Resource Exchange Mechanism for Load Management across Grids, Proceedings of the 14th IEEE International Conference on Parallel and Distributed Systems (ICPADS 2008, IEEE CS Press, Los Alamitos, CA, USA), Dec. 8-10, 2008, Melbourne, Australia.

[16] Srikumar Venugopal and Rajkumar Buyya, Cost-based Scheduling for Data-Intensive Applications on Global Grids, Proceedings of the 14th IEEE International Symposium on High Performance Distributed Computing (HPDC-14, 2-pages Poster Paper, IEEE CS Press, Los Alamitos, CA, USA), July 24-27, 2005, Research Triangle Park, North Corolina, USA.

[17] Jia Yu, Rajkumar Buyya, and Chen Khong Tham, Cost-based Scheduling of Workflow Applications on Utility Grids, Proceedings of the 1st IEEE International Conference on e-Science and Grid Computing (e-Science 2005, IEEE CS Press, Los Alamitos, CA, USA), Dec. 5-8, 2005, Melbourne, Australia.

[18] Weina Lu, Shoubao Yang, Leitao Guo, Rui Zhang, “Reputation-aware Transaction Mechanisms in Grid Resource Market”, Sixth Intl: conf: on Grid and Cooperative computing(GCC’07), 16-18 Aug, 2007

[19] Chunlin, L. and L. Layuan, "A Pricing Approach For Grid Resource Scheduling With QoS Guarantees " Fundam. Inf. 76 (1-2): 59-73, 2007.