A Grid Resource Scheduling Algorithm Based on the Utility Optimization

Jiang Chen[1], Jian Peng[1,2], and Xiaoyang Cao[3]

[1] School of Computer Science, Sichuan University, Chengdu 610065, China
[2] Department of Computer Science, University of Maryland, MD 20742, USA
[3] Chengdu University of Information Technology, Chengdu 610225, China

Abstract. To solve the problem of heterogeneity of user requirements in grid resource allocation, a grid resource scheduling algorithm based on utility function is proposed by analyzing the relationship between the executing time and cost and the user utility function, the theory of economics is used to solve the optimal problem of the user utility function. The result of experiment shows, when the system finished the same set of gridlets, the algorithm achieves better performance not only in cost than the algorithm based on the time optimization when they spent equal time, but also in time than the algorithm based on the cost optimization on the assumption that they consumed the equal quantity of cost.

Keywords: grid, resource allocation, utility function, optimization.

1 Foreword

Grid computing as a significant emerging realm are mainly used in large-scale sharing of resources and high-performance computing[1-2]. In the grid, because of the distributedness, heterogeneity and dynamic of resources, scheduling and allocation of resources are very complicated[3-4]. Conventional resource scheduling algorithms are from the perspective of provider of resources, using theory of optimization to solve scheduling problems, to decide the quantity of resources which consumers can schedule, but it often faces the restriction of the invariability of the number of the variable, and the operation is centralized, so it does not adapt the demand of load balance, high fault tolerance of modern computing.

From the viewpoint of economics, the market in the economic activity is a resource distribution mechanism based on the independent distribution, that is, participants in the market based on market prices and their own preferences make decision. The market mechanism through the price fluctuation reflect the dynamic change of resource supply and demand condition, through supply and demand balance to achieve optimal allocation of resources, so it has good performance of the dynamic characteristics of grid. When the user's needs involve the factors of time and cost, the market mechanism is suitable to solve the grid resource allocation problem[5]. This paper proposes a grid resource allocation strategy based on the optimization of the utility, considering the factors of cost and time, to produce a reasonable budget under the premise that the grid user can complete its task.

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1355–1362, 2009.

2 Scheduling Algorithm Based on the Time Optimization and Scheduling Algorithm Based on the Cost Optimization

Monash University of Australia Rajkumar Buyya proposes two resource scheduling algorithms, one is based on the time optimization, the other is based on the cost optimization [6-7], and he designed the Gridsim simulator to simulate the scheduling algorithms. In the simulation system, the use of resources is valuable, that is, the cost of different computing power's resource is different, which user has to pay corresponding price to use the corresponding resources. There are two parameters Deadline and Budget in the task which the user submits. The Deadline parameter denotes that the duty must be completed before this deadline, if there are no resources to meet the condition, then the system gives up the execution of the task; the Budget parameter denotes that the total price which must be paid to complete the task which can not exceed the budget, otherwise the system gives up the execution of the task.

In this way, in the process of task execution, the user's QoS requirements are static[8-10], either based on the optimal time or based on the optimal cost, it can not reflect that the QoS requirements are dynamic in the process of task execution. This paper according to these two parameters constructs a utility function based on the deadline and the budget. The algorithm reflects that the QoS requirements are dynamic, and ensures that users can get the biggest benefits from the task scheduling, thus improves the utilization of grid resources and meet the QoS requirement at the same time.

3 A Grid Resource Scheduling Algorithm Based on the Utility Optimization

3.1 The Algorithm's Economic Model

In the grid environment, the resources provider obtains the benefit through the resources transfer, and the resources user completes the large-scale work which is unable to complete on its own resources through purchasing resources. Such model lets us very naturally think that we can manage grid resources through the economic model in real life.

The use of economic model can realize optimized allocation strategy of distributed resources through the market behavior, which is the main reason for the introduction of economic model. The market in the economic activity is a resource distribution mechanism based on independent distribution, that is, participants in the market based on market prices and their own preferences make decision, and the grid resource distribution needs to realize the similar independent decision-making exactly. The Grid is a heterogeneous, dynamic distributed environment, the use of resources and supplies are constantly changing. Through the introduction of economics, especially the market mechanisms based on price, in which the price fluctuation reflects the dynamic change of resource supply and demand condition, achieving optimal allocation of resource through supply and demand balance can greatly represent the dynamic characteristics of grid.

The grid resource scheduling algorithm based on utility optimization consists of three parts, from bottom-up followed by the resources layer, proxy layer and user layer.

1. The resources layer consists of one or more resources, they are the resource providers. They enter the grid resources market through being registered by the resources agent. The registration information includes CPU type, CPU quantity, memory size, operating system and its version, and so on.
2. Proxy layer includes three kinds of agents: resource agent, user agent and information service agent. The resource agent represents the grid resources, the user agent represents the grid users, the information service agent is the place that the resource provider and the resource consumer carries on the transaction, namely grid resource market. When one computing resource joins into the computing grid, its corresponding resource agent registers its information into the information service agent; Correspondently, when one computing resource withdraws the computation grid, its corresponding resources agent withdraws its information from the information service agent.
3. The user layer is the grid users or the applications, they are the resources consumers. They propose their own demand to the grid resources market through the user agent. The demand includes CPU type, CPU quantity, memory size, operating system and its version, deadline, budget, and so on.

3.2 The Algorithm's Utility Function

The grid resource allocation method should be based on the specific needs of the user to choose a suitable computing resource to perform the task. Utility represents the degree of user satisfaction obtained from receiving a service or occupying a resource. Cobb-Douglas utility function of the economics reflect the balance between different model variables, so this paper introduces this utility function to describe the heterogeneity between user needs, reflecting the grid user's preferences and objectives.

For any two commodities S and R, the utility function is

$$U(S,R) = (S)^{\theta}(R)^{1-\theta} \quad (1)$$

And a is a real number in the range of [0,1], S, R respectively denote the quantity of two kind of commodity. The function describes commodity S's degree of preference relatively to commodity R.

In general, the utility function is monotonous, but considering the request for minimum of the execution time and cost, we introduce a utility function equivalent to (1), and conduct the modeling of the grid resource scheduling with this function, computing the utility of time relative to cost.

$$U(S,R) = a\ln(S) + (1-a)\ln(R) \quad (2)$$

$S \leqslant Deadline$, $R \leqslant Budget$.

The meaning of utility function (2) is: grid users spent R grid currency in S seconds to complete the work.

3.3 Utility Optimization

The complete competitive market mechanism is based on the price, which when the resources supplies are more than the demand, the price declines, stimulating consumption; when the resources supplies are less than the demand, the price increases, restraining consumption, until the system achieves a balanced state, which the total demand is equal to the aggregate supply, the price of resources at this time is the equilibrium price. Under this equilibrium state, each resources user can satisfy the maximum benefit, at the same time the benefit of resources provider is maximum, which the total utility of system is maximum.

Economic theory has proved that, when the price is equal to marginal rate of substitution, the price is the equilibrium price. The allocation of resources at this time can achieve Pareto optimal.

$$\frac{\partial U/\partial S}{\partial U/\partial R} = \frac{aR}{(1-a)S} = P \tag{3}$$

The user demand function is based on the principal of maximizing the user utility to determine the best resource allocation plan, we can get the user demand function from (3):

$$S = \max(\frac{aR}{(1-a)P}, 0) \tag{4}$$

3.4 Algorithm Design

From 3.2 and 3.3 sections, the key to solve grid resources scheduling problem through the market mechanism is to get equilibrium price P. In particular, the core problem needed to be solved is how to get equilibrium price P when many grid users apply resources in the grid resource market, that is, the optimal allocation of resources.

To get equilibrium price P, the grid resources market use following iterative algorithm:

Step 1. The grid resource market accepts the resource application from the grid user, and the user specify the minimum resource requirement in the application;

Step 2. According to the principle of first-come-first-serve, determine that the sum of minimum resource requirement can't exceed the processing capability of the grid resource during within a time frame of time length T;

Step 3. Execute step 4-13 in each T time section;

Step 4. Take the grid resources' current price P as the initial price, determine price adjustment rate V, iteration-terminating parameter R;

Step 5. Return the price P to the grid user;

Step 6. Grid user obtains the optimal resource demand amount under the price P according to formula (4), the optimal resources demand subtracts the processing capability of grid resources to obtain total excess demand Z;

Step 7. If $|Z| \leqslant R$, then the current price P is the equilibrium price, distribute resources according to the optimal resources demand amount, this resource distribution finished, return to step 1;

Step 8. Get new price P according to formula $P = P + ZV$;

Step 9. If the P is not positive, then adjust the rate V according to the formula $V = V/2$, return to Step 8;

Step 10. Return the price P to the grid user;

Step 11. Grid user obtains optimal resources demand amount under the price P according to formula (4), the optimal resources demand subtracts the processing capability of grid resources to obtain total excess demand Z';

Step 12. If $|Z'| \geqslant |Z|$, then adjust the rate V according to the formula $V = V/2$, return to Step 8;

Step 13. Adjust Z to Z', return to step 7.

Algorithm flow chart as shown in Fig. 1:

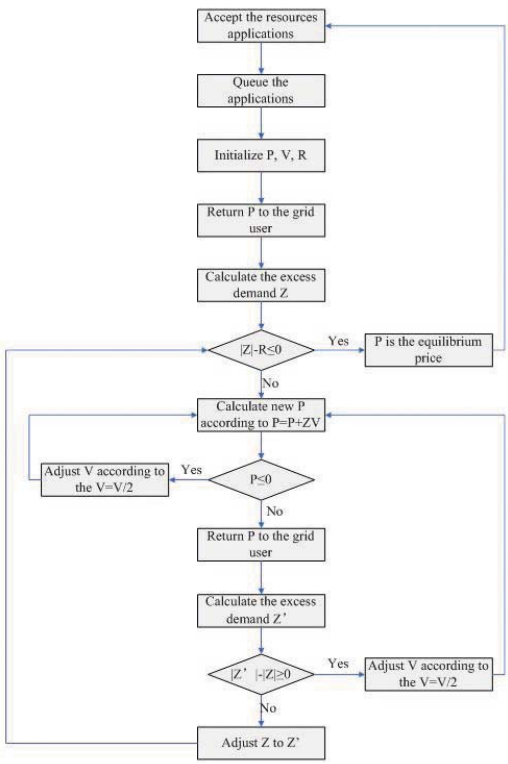

Fig. 1. The flow chart of the utility optimization algorithm

The basic idea of the algorithm is based on the Tatonnement process of economics: when the resources supplies are greater than the demand, namely Z is less than zero, the price declines; Otherwise, when the resources supplies are less than the demand, namely Z is greater than zero, the price increases, and the price fluctuation speed is proportional to the Z's absolute value. The economic theory has already proved that the Tatonnement process converges to the equilibrium price, and the convergence rate is very quick [11].

4 Simulation Experiment

4.1 The Condition of Simulation Experiment

The simulation is run on the grid simulator GridSim[12], comparing the scheduling algorithm based on utility optimization proposed by this paper to the scheduling algorithm based on the time optimization and the scheduling algorithm based on the cost optimization proposed by Rajkumar Buyya. The configuration of computer running the simulation is as follows: Inter (R) Celeron (R) CPU2.40GHz, 1G Memory, Windows Server 2003, JDK1.4.2. We establish simulation environment through Visual Modeler, create a simulation system for the six heterogeneous computing resources, fix the deadline of execution time is 300 units of time, the budget starts from 32000 to 35000, and the length of stride is 500. The user submits ten gridlets.

4.2 The Simulation Experiment Result and Its Analysis

In the simulation experiment, three different scheduling algorithms separately complete the computing task of ten gridlets. We compare the performance of the three algorithms from the cost, time and the number of gridlets completed. The results are in Fig. 2, Fig. 3, Fig. 4.

From the simulation result figures we can know: Overall, the scheduling algorithm based on the cost optimization is less than the scheduling algorithm based on the time optimization in the actual cost, the scheduling algorithm based on the time optimization is less than the scheduling algorithm based on the cost optimization in the actual time, but the scheduling algorithm based on utility

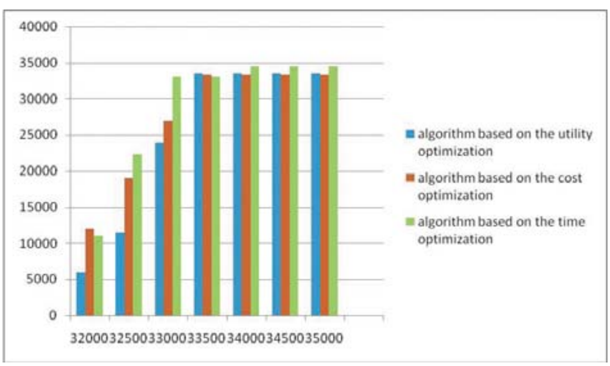

Fig. 2. The actual cost of three grid scheduling algorithms

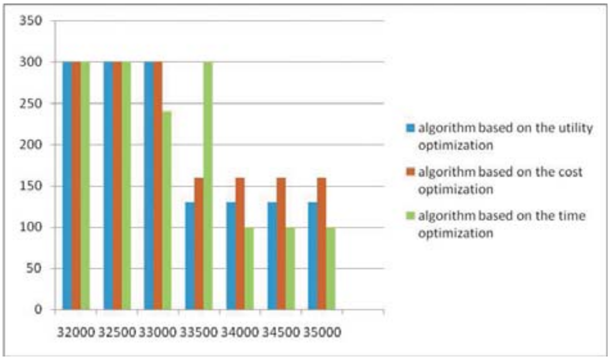

Fig. 3. The actual time of three grid scheduling algorithms

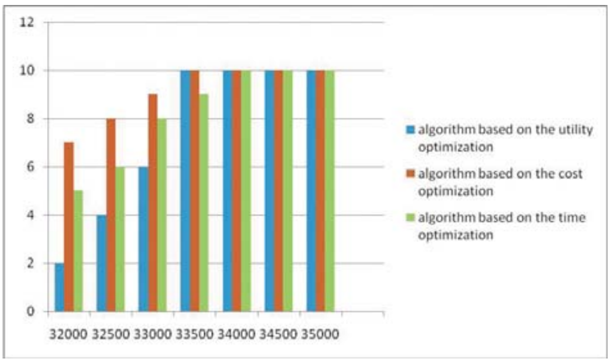

Fig. 4. The actual number of gridlets completed of three grid scheduling algorithms

optimization, because of considering dynamic change of QoS requirements, considering the factors of cost and time through the utility function, reflecting the heterogeneity of user requirements, achieves that, under the condition of completing the same quantity of gridlet grid task, when the time consumed is the same, the cost of this algorithm is less than the scheduling algorithm based on the time optimization, and, when the cost is the same, the time of this algorithm is less than scheduling algorithm based on the cost optimization.

5 Concluding Remarks

The grid resource scheduling is an important research topic in grid computing area. This paper models the grid resource scheduling with economics method, considers the factors of cost and time, proposes a grid resource scheduling algorithm based on utility function, which reflects the heterogeneity of user requirements through the utility function, and optimizes the algorithm with the theory of economics, namely the equilibrium point of market transaction is the optimal allocation of grid resources. The results show that, the scheduling algorithm based on utility optimization has better performance than the scheduling algorithm based on the time optimization and the scheduling algorithm based on the optimization of cost, and can better meet the dynamic changes of the user QoS requirements.

What the present simulation experiment considered is the single-user, multi-task environment, in order to make the simulation experiment reflects the grid environment better, the next step is to construct a utility function which better reflect the heterogeneity of user requirements in multi-user, multi-task environment, and evaluate the performance of the scheduling algorithm based on utility optimization further.

References

1. Fox, F., Gannon, D.: Computational Grids. IEEE Computational Science and Engineering, 74–77 (2001)
2. Foster, I.: The Grid: Computing Without Bonds. Science American, 78–85 (2003)
3. Krauter, K., Buyya, R., Maheswaran, M.: A Taxonomy and Survey of Grid Resource Management System. Software: Practice and Experience, 135–164 (2002)
4. Buyya, R., Abramson, D., Giddy, J.: Economic Models for Resource Management and Scheduling in Grid Computing. Concurrency and Computation: Practice and Experience, 1507 (2002)
5. Wolski, R., Plank, J.S., Brevik, J.: Analyzing Market-based Resource Allocation Strategies for the Computational Grid. International Journal of High Performance Computing Applications (2001)
6. Bemsten, J., Spets, T.: CPE WAN Management Protocol. DSL Home Technical Working Group (2004)
7. Stark, B.: LAN-Side DSL CPE Configuration Specification. DSL Home Technical Working Group (2004)
8. He, X., Sun, X., Gregor, V.L.: QoS guided Min-min heuristic for grid task scheduling. Journal of Computer Science and Technology, 442–451 (2003)
9. Yu, J., Buyya, R., Tham, C.K.: QoS-based Scheduling of Workflow Applications on Service Grids. In: Proceedings of the 1st IEEE International Conference on e-Science and Grid Computing, pp. 1–9 (2005)
10. Dogan, A., Ozguner, F.: On QoS-based scheduling of a meta-task with multiple QoS demands in heterogeneous computing. In: The 16th Int'l Parallel and Distributed Processing Symposium, Fort Lauderdale, FL (2002)
11. Varian, H.R.: Microeconomic Analysis. W.W.Norton&Company, New York (1992)
12. Murshed, M., Buyya, R., Abramson, D.: GridSim: A Toolkit for the Modeling and Simulation of Global Grids. Technical Report, Monash-CSSE (2001)