Research an Intelligent Schedule Measure in Grid Network System

Huanjiong Zhang

(Hangzhou Dianzi University, Hangzhou, 310018)

Abstract—The asset management and schedule measure in Grid computing system are the key technologies of describing the network resources, network management and task assignment. The intelligent schedule measure in Grid Internet system is researched, the essential conditions of realizing intelligent schedule measure are discussed, one intelligent schedule measure is pointed out, and the example prove the advantage of this schedule measure.

Keywords—Grid Computing; Grid Network; Intelligent Schedule Measure

1. Introduction

Grid computing can integrate the different computing resources that distribute different positions and are with different equipment types to become a unified virtual super-strong computer, it also can realize the very grand-scale computing function. Not only can grid computing do it best to utilize the existing computer resources, but also it can express the advanced performances of the grid computing system, so research of grid computing becomes a very hot branch of hot points in these researching fields.

There are many research achievements of grid computing [1-3], especially, some larger-scale research institutes depend on their favorable conditions to build their experimental grid networks with different types computers and infrastructures, some examples of experimental researching achievements are taken as follows: one case is that the computing ability of experimental grid network with 55 thousand different personal computers is stronger than that of 10 advanced super computers; another case is the giant advanced performance computing item of TeraGrid[4], and so on. But, up to now, the time about the research and application of grid computing is only in the primary stage, there is long distance to the real universality and practicality of grid network. So to deeply research the key technologies of grid network system, to use the creative methods that are different from the traditional methods of network linkage to solve the difficult problems which include the problems of resource management and monitor, system mutual-operation, software compatibility, system security, system communication, intelligent schedule measure, and so on, become the most important and urgent tasks.

The schedule measure technology of grid network is one of key technologies, it has developed from the traditional schedule measure technology of distributional systems with the parallel proceeding, so it also shows the general characters of ordinary parallel proceeding systems [5-7], but grid network system has the particular constructs and functions, the schedule measure of grid network also appears the distinct individual characters. On the aspect of increasing utilization ratio of system resource, the schedule measure technology can avoid the non-essential loss, and enhance the operative velocity, it also make a key role of enhancing the validity and reliability of the grid network system, besides the above, schedule measure can directly express the agility and active characters of the system.

In this paper, intelligent schedule measure technology is studied, the essential conditions of realizing intelligent schedule measure are summarized, furthermore, a feasible intelligent schedule measure scheme is pointed out, this new scheme is adaptive to the heterogeneous environment, the main idea of this measure scheme is to define a intelligent comprehensive cost function that can reflect the dynamically tasks allocation to network nodes, when the solution of the comprehensive cost function tend to the minimum, the validity of task allocation in grid network can tend to the highest.

2. Conditions of Realizing Intelligent Schedule Measure

Dynamical and agile schedule measure can make the resource deployment match the target requirements of computation, it means that the scale and performance of virtual computer of grid computing is adaptive to the computational requirement of network user, then it can avoid wasting the network resource, and also can avoid the situation that grid computing can not implement the concrete
computational requirements because the scale of the grid computing network system is too small. Though intelligent schedule measure is regard as the tool that belongs to the middle layer of network system, it must be supported by many constructs and protocols in the basic layer of network system, besides all of them, it also accepts the support of other key technologies, the support conditions of intelligent schedule measure in grid computing system are summarized, they at least include the following conditions:

Grid network must be supported by the powerful network mutually-linkage, the nodes of grid network should express agile and active under the hyper-construct condition, because the constructs and systems of different computing resource in grid network are different, to unify these different computing resource as a virtual super-power computer, one necessary condition is that nodes have the grand powerful mutually linkage, these nodes are the computing providers and the receivers of grid computing tasks, they consist of the high terminal servers, data books and some kinds of terminal equipment with computing functions, based on the protocols of communication, route and so on, then all of these equipment can be created to unify a grand-scale grid computing platform, and great scale programs that can adapt to the running of grid computing system are exploited, during this proceeding, the nodes are entrusted some self-authorities, such as process power of computation, security, check and correction wrong codes.

To realize intelligent schedule measure is also supported by the technologies of parallel computing and distributed computing [8], because the existing computational resources are from different geometrical positions, the mode, computational power, disposition and many others are different with one another, only do the distributed and parallel technologies be utilized, the whole system can be unified and intelligent schedule measure can be realized.

What intelligent schedule measure is successfully realized also depend on the good cooperative power, security and opening character. In new Pervasive/Grid architecture, the highly cooperative power of nodes is a basic guarantee to successfully make schedule measure, and in order to easily add additive services and additive business, the architecture and function must be enough open, so the opening character of the network system is also an essential condition; the confidentiality and security of the network are not only the essential condition of information transmission in grid network, but also the condition of network management and intelligent schedule measure.

3. A Scheme of Intelligent Schedule Measure

The intelligent schedule measure really is a non-determined dynamical task allocation under the heterogeneous environment. The optimal schedule measure appears the following aspects: the resource utilization ratio of system is the highest, the time of proceeding and communication is the shortest, the confidentiality, security and reliability of the system is the best, and so on, so the intelligent schedule measure proceeding can be regarded as that parallel computing makes the optimal policy with multiple indexes under the determined conditions in heterogeneous environment. In theory, this scheme can be transformed to find the minimum solution of the cost function, but the schedule measure usually is a NP-perfect problem, the optimal solution is difficult to be obtained, in the real engineering field, the schedule measure about elicitation method sub-optimal algorithm is researched and applied, it was fully discussed in many literatures [9][10].

The complete process of the intelligent schedule measure can be divided into three parts, in the first part, based on the heterogeneous environment in grid network, a resource function is defined, the variables and parameters of the resource function correspond to the performances of independent computers and their computational powers in heterogeneous environment. It is easy to be denoted as $G(m_1(k_1), m_2(k_2), \ldots, m_1(k_1), m_i(k_i))$, here, $m_1(k_1), m_2(k_2)$ and $m_i(k_i)(i = 1, 2, \ldots, I)$ express the resource character index of the first, second and $i^{th}$ processors in resource function, this resource function can shield the original heterogeneous environment with the mapping proceeding. In the second part, the system of grid network pre-proceeds the computational environment which the user requires, the total task is cut apart many sub-module tasks, all of these sub-tasks with grain requirement are coded and
become independent sub-module tasks, it is also called as task allocation function \( F(n_i, n_2, \ldots, n_i, \ldots, n_j) \), \( n_i \) is the \( i \)th sub-task module. In the third part, to make the mapping from the resource function to task allocation function, that is that resource function matches the task allocation function, to assume the relationship between resource function and task allocation function is denoted \( \Phi(\cdot) \), so the above description is rewritten:

\[
\begin{align*}
G(m_1(k_i), \ldots, m_i(k_i), \ldots, m_j(k_i)) \\
= \Phi\{F(n_i, n_2, \ldots, n_i, \ldots, n_j)\}
\end{align*}
\]

\( i \rightarrow m_i(k_i) \)

In here, the sub-task module \( n_i \) also matches \( m_i(k_i) \) of the independent proceeding machine, then the task requirement can adapt to the independent proceeding machine, and the schedule measure is also finished.

Accepting this kind of schedule measure, the main reason is that it can shield the concrete heterogeneous character of processors through defining resource function, it also overcomes the restriction that is from the heterogeneous environment. A reasonable and reliable cost function can be defined with two performance indexes, one performance index expresses the total time tends to the shortest, another performance index expresses the resource utilization ratio tends to the highest. If the intelligent schedule measure can be satisfied with these two aspects requirement, this scheme measure is regard as the intelligent schedule measure.

In fact, the situation that these two condition are satisfied in the same time hardly appears, so the comprehensive cost function must be designed in, a simple method of designing it is to make a linear combination function with adding weight coefficient to every variable which is mapped to the corresponding performance index, it can be concretely written:

\[
\Phi(\eta, t) = \alpha(1 - \eta) + (1 - \alpha) \sum_{i=1}^{l} t_i
\]

Here, \( \alpha \) is a weight coefficient, and \( \alpha \in [0,1] \); \( \eta \) is the resource utilization ratio, \( \eta \in [0,1] \); \( t_i \) is the total time of the pre-proceeding and communication of \( i \)th sub-module, so to choose the optimal schedule measure transforms to find the solution of \( \min \Phi(\eta, t) \), furthermore, ordinary mathematical model of the optimal schedule measure can be described:

\[
\min_{\text{conditions}} \{\Phi(\lambda_1, \lambda_2, \ldots, \lambda_k)\}
\]

Here \( \Phi(\lambda_1, \lambda_2, \ldots, \lambda_k) \) is the comprehensive cost function that is based on the task requirement of grid computing, \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are variables of \( k \) numbers corresponding performance indexes. If the comprehensive cost function is designed in the linear combination, it can be written:

\[
\Phi(\lambda_1, \ldots, \lambda_k) = \sum_{i=1}^{k} \alpha_i \lambda_i, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1
\]

\( \alpha_i \) is weight coefficient too, and \( i = 1, 2, \ldots, k \).

The intellectual ability of schedule measure also embodies the mutual influence between the network system and network user, when the comprehensive cost function is designed, the expectation and asking of grid network user can also express, especially, the choice of performance indexes and the number of parameters are defined with the requirement of user, the weigh coefficients is decided with the special requirement of user, In a word, this intelligent schedule measure is an optimal scheme that can base on the special requirement of user, it does the best to appear the personality and agility of grid computing network system.

In order to deeply explain this intelligent schedule measure, a hypothetical example is taken, because of the restraining from the experimental conditions, the example is relatively simple. The comprehensive cost function is adopted form (2), one performance index that is denoted variable \( \eta \) is the system utilization ratio, another performance index is the pre-proceeding total time, when weight coefficient \( \alpha \)
and $\eta$ are determined, based on finding the minimum solution of form (2), the schedule measure is easy to be obtained, furthermore, the schedule measure is also affected by the weight coefficient $\alpha$, when $\alpha$ changes, the concrete schedule measure also changed.

The summarized main steps of intelligent schedule measure are as follows:

Step 1: the user of grid network points out the objective computing tasks through the corresponding node.

Step 2: based on the objective computing task, the task scale is defined, and the pre-proceeding about cutting apart the task, denoting and coding sub-tasks is made, then the task allocation function $F(n_i, n_2, \cdots, n_i, \cdots, n_j)$ is obtained, in here, $n_i (i = 1, 2, \cdots, j)$ is a task allocation sub-module.

Step 3: based on the asking of user, a comprehensive cost function with multi-objective indexes and multi-coefficients are designed in.

Step 4: to call the resource function $G(m_1(k_1), m_2(k_2), \cdots, m_i(k_i), \cdots, m_j(k_j))$, and to build a map to the task allocation function, every map relation is to point the value in the comprehensive cost function.

Step 5: to find the minimum solution of the comprehensive cost function, then the coefficients are determined, and every subsidiary task module is allocated to the corresponding processor.

4. Conclusion

Intelligent schedule measure in grid computing system is one kind kernel technology, the intelligent schedule measure technology can enhance the utilization ratio of resource, it also can avoid the extra loss, furthermore, it is to enhance the validity and reliability of grid computing system, it expresses the advantages of this new computing technology.

Under heterogeneous condition, this new intelligent schedule measure is transformed to obtain the minimum solution of the comprehensive cost function with multiply performance indexes, it is easy to be realized and it can enhance the performances of grid network system, the simple example also proves that this new intelligent schedule measure is very useful.

Reference

[1] Douglos. F and Foster. I, The Grid Grows Up (J) . IEEE Internet Computing (2003)

[2] Hong xuehai, Xu zhuoqun, Ding wenkui. Summarization of the technology and application of Grid Computing, journal of Computer Science(J),2003,30(8),pp.1-5.

[3] Foster,L. The Grid: A new infrastructure for 21st century science (J) . Physics Today, 55(2). 2002, pp42-47.

[4] Hesham El-Rewinei, and Mostafa Abd-El-Barr, Advanced computer architecture and parallel processing[M]. Electrical Industry Press House, Beijing Dec.2005, pp.127.

[5] Gabow. H. An almost linear algorithm for two-processor scheduling (J) . Journal of ACM,29(3), 1982, pp766-780.

[6] Lewis,T. and El-rewini, H.Parallax A tool for parallel program scheduling. IEEE parallel and Distributed Technology (J) . Systems and Applications, 1(2) 1993,pp62-72.

[7] Lo,V. Heuristic algorithms for task assignment in linear array networks(J), IEEE Transactions on computers,41(7), 1992, pp877-880.

[8] Duncan,R. A survey of parallel computer architectures (J) , Computer,23(2), 1990,pp5-16.

[9] Chen,s. et al A Selection Theory and Methodology for Heterogeneous Supercomputing. Proc. Workshop on Heterogeneous Processing (M), IEEE CS Press, Los Alamitos, CS, Order No.3532-02, 1993, pp15-22.

[10] Freund,R. and Siegel,H. Heterogeneous Processing (J) , IEEE computer, 26, 1993,pp13-17.