

Available online at <http://www.meecspress.net/ijwmt>

Study on Replica Strategy in the Grid Environment

Rui-qiong Zhong^{a,b,*1}, Yun-jia Xi^{c,*2}, De-zhi Han^{b,*3}

^a School of Management, Jinan University, Guangzhou, China

^b School of Informatics, Guangdong University of Foreign Studies, Guangzhou, China

^c School of Educational Technology and Communication, Guangdong Polytechnic Normal University, Guangzhou, China

Abstract

An effective replica management method can save bandwidth, reduce latency, balance load, and improve system reliability. A comprehensive scheme is introduced to manage replica for the grid environment. The schemes of replica creation and placement based on the frequency of visits and storage space, replica replacement based on the replica value, and replica selection based on comprehensive performance are proposed. Our scheme can achieve optimization of data distribution and replication, improve the efficiency of data visits in the grid environment.

Index Terms: Grid; Replica; Value; Frequency

© 2011 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science

1. Introduction

With the development of electronic information technology, the data grid has become an important field of computer science and technology research. Wide-area data sharing and integration of computing power have been provided with data grid technology for users. It is imperative to provide data replicas in the grid environment due to the distribution of large amounts of data. The access latency and bandwidth consumption can be reduced through the replica technique. By creating multiple replicas of the same data, the load balance and reliability of the whole system can be improved.

In this paper, the replicas selection, creation, placement and replacement methods in the grid environment have been deeply studied, and the corresponding schemes and algorithms have been put forward. A certain quality of service for the grid can be ensured due to sufficient replicas. Meanwhile, the characteristics of the grid and how to make the performance of the whole system best has been comprehensively taken into consideration. The schemes about replicas management, such as the replica creation based on the frequency of visits inside the domain, the replica placement based on network bandwidth between domains to select a domain, the replica replacement in accordance with its value and storage space, and the replica node selection

* Corresponding author:

E-mail address: ¹zrq@jnu.edu.cn; ²10986661@qq.com; ³han_dezhi88@tom.com

according to the comprehensive performance of nodes are to be proposed. Our replica value model includes three factors: the recent frequency of visits to the replica, the later access times which is calculated by prediction, and the trust evaluation on the replica.

The related work is to be discussed and the characteristics of this work is to be analyzed in the section 2; the selection strategy of the node based on comprehensive performance norm is to be proposed in the section 3 ; the replica creation and placement method based on the frequency of visits and node storage space is to be presented in section 4; the replica replacement strategy based on the value and storage space is to be put forward in the section 5 ; finally, the conclusion is to be raised.

2. Related Work

A typical data replication strategy needs to solve the following questions: which one to duplicate, the replica number and the position. In recent years, study on replica location has been conducted by the domestic and foreign experts. The well-known grid development tools Globus provides a kind of replica directory service to support the replica location. Although it has been applied in many projects, the centralized directory approach it used can't support dynamic expansion of the system, and its reliability is not high.

The replication strategy based on the frequency of visits is proposed in [1], which calculates the frequency of local visits and places replicas to the node with larger node degree, but it doesn't consider the system resources like storage space. The replica placement method based on p-median is presented [2], by means of solving the minimization problem of the objective function to determine the replica location, but the differences between inside-domain and the outside-domain and the specific algorithm are not taken into account. The replica mechanism combined with information security is proposed [3], and the mathematical model determining the replica number is presented. The model comprehensively considers the economic interests and reputation of the service provider party, gives a simplified treatment on the optimization problem of one or two goals, and determines the number of replicas and the optimal limits by the way of numerical computation and analysis.

During the data replication process, if there is not enough storage space, we need to use the replica replacement algorithm. The most used strategies include the recently least used algorithm (LRU), the least frequency used algorithm (LFU). The two replication algorithms of Simple Bottom Up (SBU) and Aggregate Bottom Up (ABU) for Multi-tier data grids is put forward [4]. To minimize the time of the visit to data and the network load, replicas of the data should be created and spread from the root center to regional centers, or even to national centers. However, the strategies in this work are applicable only to multi-tiered grids. The replication strategy is defined as the map of the ratio of the frequency of visits to data replica [5], and three different replication strategies are discussed, so as to find out that the square root replication is the optimal under the standard of the average searching length. The algorithm called LALW [7] is presented where replica is dynamically created based on the weights, but the placement is only in the site level.

Previous work in this area mainly focuses on one aspect of data replicas creation, selection and replacement. Meanwhile, whether the performance of the whole system is optimal hasn't been considered after replication. In addition, the replica strategy suitable for the grid environment is still relatively less, and specific algorithm and implementation is rarely presented. In this paper, the scheme of replica management is based on the hierarchy structure in which nodes in the grid are connected in the form of unstructured peer to peer network inside the domain. Nodes are divided into three categories: ordinary node, inside-domain super node and inter-domain super node, the node generally referred to is the ordinary node. The network structure of the grid model in this paper is as shown in Fig.1.

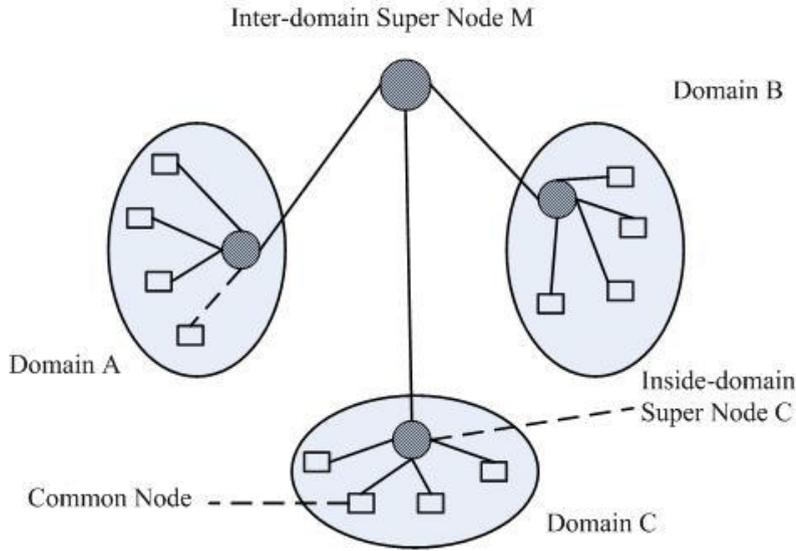


Fig. 1. Network structure of this grid model

3. Replica Selection Strategy

When the system is in actual operation, the available resources of the network link is to be affected with a lot of uncertain factors, and the service near users is often not able to provide satisfactory service quality [8]. Therefore, we must make the dynamic evaluation and selection on the replica. Through the grid information service, we can get the dynamic information of all the resources in grid environment, such as available CPU time, memory and bandwidth of node. The inter-domain super node is responsible for detecting the average network bandwidth of each domain, which is as a standard of selecting a domain. In order to select from multiple replicas in the domain, we define a comprehensive norm M for each node. With the dynamics of the grid, in order to optimize the efficiency of visits to data, reduce the access latency, and improve the system's load balance, the main target of the replica selection is to maximize the whole system performance.

In our model, four factors which have effects on replica selection are considered. They are network bandwidth, CPU load, I/O state, credit. Thus, the comprehensive norm has been improved according to [10]. It is defined as followed:

$$M_{ij} = \alpha \times P_{ij}^{BW} + \beta \times P_j^{CPU} + \gamma \times P_j^{I/O} + \theta \times Cre_j \quad (1)$$

Where M_{ij} shows the comprehensive performance of the node j from node i 's perspective; P_{ij}^{BW} shows the percentage of bandwidth available from node i to node j , the value of which is current bandwidth divided by the preset theoretical bandwidth; P_j^{CPU} shows the percentage of CPU idle states of the node j ; $P_j^{I/O}$ shows the percentage of I/O idle states of the node j ; Cre_j shows the credit and reliability of the node j , and it is given from other nodes according to the behaviors of the node j and the available resources. The Cre_j value

can be obtained from the following (8). $0 \leq \alpha, \beta, \gamma, \theta \leq 1, \alpha + \beta + \gamma + \theta = 1$, each of them shows the importance of each resource, and is measurement value. Sort by comprehensive performance, the node with the highest comprehensive performance is selected as the inside-domain super node. The inside-domain super node is responsible for calculating and updating the value M of the node inside the domain. The node with the comprehensive value M greater than the threshold of the node which user expects, are often to be selected.

4. Replica Creation Strategy

The cost factors of the node to create a replica include node properties, the network bandwidth, etc. When the inside-domain user is to visit the replicas in this domain, here only consider the storage space of nodes for the time being. The larger the storage space, the greater the superiority of the replica creation on it is. When the node needn't frequently replace the local replicas, the time and bandwidth consuming can be saved. Therefore, when selecting the node where to place replicas, node's remaining storage space should be considered.

On the basis of the calculating the frequency of visits and the number of replicas [4], we make some improvements. Each super node within the domain is responsible for calculating the required replicas of the requested popular files, while the inter-domain super node is in charge of the inter-domain message passing and getting the network bandwidth of each domain.

First of all, time is divided into continuous time slots, and at each time slice the number of visits to the replica is counted. At the end of time slice, the two items whose <replica ID> are the same are merged in the record, and the total number of visits to each replica in this time slice is calculated. The visit to information in different time slots has different weights of time, and the weight of the latest time slice is the highest.

(1) The frequency of visits $Fre(f)$ to replica f is calculated as the followed:

$$Fre(f) = \frac{\sum_{\tau=1}^{L(\tau)} (a_{f(\tau)} \times 2^{-\Delta\tau})}{L(\tau)} \quad (2)$$

Where $f \in F$, F indicates a collection of different replica objects that have been visited, $L(\tau)$ shows the number of the passed time segment, $a_{f(\tau)}$ shows the number of visits that the replica f in the period τ , $2^{-\Delta\tau}$ shows the time weight of the visited number $a_{f(\tau)}$.

(2) The replica to be replicated is selected. Firstly the average frequency of visits to replicas collection is calculated, and the calculation formula is as follows:

$$Fre(F) = \frac{\sum_{\forall f \in F} Fre(f)}{N(F)} \quad (3)$$

Where $N(F)$ is the number of the replicas in the F . Next, the replica whose frequency of the visit is greater than or equal to the average frequency is selected, and the number of replicas for replication is calculated, the corresponding calculation formula is as follows:

$$N(f) = \left[\frac{Fre(f)}{Fre(F)} \right]_{\text{Rounding}}, Fre(f) \geq Fre(F) \quad (4)$$

Here on this basis of [4], and according to the frequency of visits to the replica and node resource status, how to create replicas and where to place them are to be studied. First, set the requested file list at a certain time is $\{f_1, f_2, \dots, f_n\}$, the process is as follows:

(3) The inter-domain super node M sends a query message to each inside-domain super node A, B and C, to determine which one is the popular file and get their information.

(4) After the node A, B, C respectively receives query information, according to the (2),(3) and(4) to determine the popular file and the number of the required replicas.

(5) As is shown in Fig.1, after the inter-domain super node M gets the information about popular files from node A, B and C, according to the total number of the replicas of each domain requiring, the replicas in the requesting list will be sorted in descending order. For example, the replicas f_1, f_2, f_3 are popular files (namely the frequency of visits is greater than the average frequency of visits in its own domain)in domain A,B,C, and the required number of replicas is respectively (3,2,0), (2,8,5) and (0,1,6), then the sequence is B, C and A. The replicas are duplicated to each domain according to that sequence. Before the replication of the inter-domain replicas, we have to consider the whole system performance. Because the inter-domain network bandwidth may become the system bottleneck, before the replicas transfer between domains, the inter-domain super node is responsible for collecting the network bandwidth between domains. If the bandwidth between the two domains is less than a given threshold, the replication isn't conducted.

(6) In each domain, according to the replica selection strategy, the inside-domain super node selects the node with better performance of comprehensive norm to place the replica. Here we suppose the property of resources is only storage space, and it is as follows:

$$M_j = \alpha_1 \times \frac{er_{ji}}{r_1} \quad (5)$$

Where er_{j1} shows the remaining storage space of the node j , r_1 shows the required storage space for replication, $\alpha_1=1$. For each replica that is to duplicate, if there is node j and satisfies $M_j \geq H$ (H is the given storage space threshold), this node will be selected to place the replica. If there is no such node j , it will turn to the algorithm of replica replacement.

5. Replica Replacement Strategy

In order to replace those unstable and the inefficient replicas, effectively improve the efficiency of visits to data replica, on the basis of [6] and [9], a replacement algorithm based on the value of replica is presented. As the node's dynamics and a large amount of data, the cost of producing a replica is relatively high, so unnecessary replica deletion should be avoided. Therefore, before a replica is deleted, by means of the value model to replace replica and whether the node storage space can meet the requirements after replacement should all be determined. The value model can be described in the following mathematical formula:

$$v = \{F, V, V_{TH}\} \quad (6)$$

Where $F = \{f_1, f_2, \dots, f_m\}$, which is a collection of replicas, $V = \{v_1, v_2, \dots, v_m\}$, which is the value collection of replicas. The higher the value V of a replica, the higher the utilization rate of that is or more stable is, so the greater the probability of retaining the replica is. V_{TH} is threshold, and the replica whose value is higher than the threshold is the replica with high value. We use the function of replica value $V(f, m, n)$ to denote the value of the replica, as is shown in the following formula:

$$V(f, m, n) = Fre_m(f) + \sum_{i=1}^n P_i(f) + Trust_m(f) \quad (7)$$

Where $Fre_m(f)$ indicates the average frequency of visits to the replica f in the previous m requests, and it can be obtained by the (2) and (3). $P_i(f)$ shows the probability that the replica f receives the visited request at the i^{th} file request. So $\sum_{i=1}^n P_i(f)$ indicates the access times that the f will be requested in the next n file requests. It is calculated based on the previous m requests and the Spatial Locality. Spatial Locality can be explained that if the file f has been visited last time, then the file whose content has a greater similarity is more likely to be visited next. $Trust_m(f)$ is the comprehensive trust evaluation on the replica f during the last m requests. The definition is as follows:

$$Trust_m(f) = \begin{cases} \frac{\sum_{t=1}^k \omega_T(x_t) \times f(x_t)}{k}, & k \neq 0; \\ 0, & k = 0 \end{cases} \quad (8)$$

Where k shows the number of the visits to the replica f in the last m requests; $f(x_t)$ is the trust score of the replica f evaluated by the user. The faster the access speed and the better the replica quality, the higher the score is. On the contrary, if the speed is slower, or there is quality problem, or there exists Virus and Trojan, then the score will be lower. Each score value is in $[-1, 1]$. The time weight of scoring is as the following formula (9). The more recent the time, the higher the weight is. The weight of the latest score is 1.

$$\omega_T(x_t) = e^{-\Delta t} \quad (9)$$

Suppose the file lists on a node is $Local[] = \{L_1, L_2, \dots, L_n\}$, and all of the files have been sorted in ascending order. The replica f is to be duplicated from the other node to this node. $FreeStorage$ shows the remaining storage space; $Size$ shows the size of the replica. The algorithm of the replica replacement is as followed.

Algorithm: Replica replacement based on value

```

1: if ( $FreeStorage + \sum Size(Local[]) < Size(f)$ ) then
2:     return
3: else
4:     variable i and j initialization
5:      $Size = Size(Local[0])$ 
6: end if
7: repeat
8:      $A[j] = Local[i]$ 
9:      $Size = Size + Size(Local[i])$ 
10:     $i++$  and  $j++$ 
11: until ( $i > n$ ) || ( $Local[i].value > f.value$ ) || ( $FreeStorage + Size \geq Size(f)$ )
12: if ( $FreeStorage + \sum Size(Local[]) \geq Size(f)$ ) then
13:     Delete  $A[0], \dots, A[j-1]$  from Local []
14:     Replicate the replica  $f$  to Local []
15: end if

```

6. Conclusion

The creation, replacement and selection of the replicas are the important functions to realize the replica management in the grid environment. This paper carries out the research on the three aspects, which consider the characteristics of the grid and the global optimization of system performance, and the corresponding schemes are proposed. They are the replica creation and placement scheme based on the frequency of visits and node storage space, the replica replacement scheme based on the value and storage space, and the replica selection scheme based on the comprehensive performance norm. In the definition of the replica value, the previous frequency of visits, the predicted frequency of visits, and the integrated trust value of the replica are considered. Next, a comprehensive analysis, experiment and simulation on the performance of the method will be conducted.

Acknowledgements

This work is supported by National Natural Science Foundation of P. R. China under the Grant NO 61070154, Guangdong Natural Science Foundation of P. R. China under the Grant NO 20092A008 and Guangzhou scientific and technological project of P. R. China under the Grant NO 200922-D081.

References

- [1] YANG Xiao-tao, DENG Su, HUANG Hong-bin. An Access Frequency Based Replication Strategy for Unstructured P2P Systems. MICRO ELECTRONICS & COMPUTER, October 2008.(in Chinese)
- [2] SU Cheng, ZHANG Hong, BI Fang-ming, YANG Shou-bao. Algorithm for Replica Placement Based on p2median Location in Grid Environment. Journal of China University of Mining & Technology, 2009,38(5). (in Chinese)
- [3] LI Jing, CHEN Shu-yu, WU Chang-ze. Model of data replication strategy based on security in grid. Computer Application, October 2006. (in Chinese)
- [4] Tang, M., Lee, B., Tang, X., and Yeo, C. The impact of data replication on job scheduling performance in the Data Grid. Future Generation Computing System 22, 3 (Feb. 2006).
- [5] E. Cohen, S. Shenker. Replication strategies in unstructured peer to peer networks[J]. Proc.of ACM SIGCOMM,Califarnia, 2002.
- [6] YAN Xiao-dong, XU Hui-min, XU Ya-jing, L IU Lian-xi ,CHEN Bao-feng. A Data Replica Replacement Algorithm Based on Value Model. Journal of Beijing University of Posts and Telecommunications, August 2006. (in Chinese)
- [7] Ruay-Shiung Chang, Hui-Ping Chang, Yun-Ting Wang: A dynamic weighted data replication strategy in data grids. IEEE Conference.
- [8] LI Chun-Hong, GU Tie-Cheng, LI Wen-Zhong, LU Sang-Lu, CHEN Dao-Xu. A Grid-Based Multimedia Content Distribution Framework. Computer Science in China,2005,32:105-108. (in Chinese)
- [9] Tian Tian, Junzhou Luo, and Zhiang Wu. A Replica Replacement Algorithm Based on Value-Cost Prediction[J]. CSCWD 2007, LNCS 5236, pp. 365–373, 2008. (in Chinese)
- [10] Chao-Tung Yang, I-Hsien Yang, Chun-Hsiang Chen, Shih-Yu Wang. Implementation of a Dynamic Adjustment Mechanism with Efficient Replica Selection in Data Grid Environments. Proceedings of the 2006 ACM symposium ,SAC'06, April, 23-27