http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 6 (7): 1090-1094, 2007
ISSN 1812-5638
© 2007 Asian Network for Scientific Information

The Grid Resource Discovery Method Based on Hierarchical Model

[1]Yulan Yin, [2]Huanqing Cui and [2]Xin Chen
[1]Department of Mathematics and Physics, Anhui University of Science and Technology,
Huainan Anhui, 232001, China
[2]College of Information Science and Technology, Shandong University of Science and Technology,
Qingdao Shandong, 266510, China

Abstract: Grid is considered as next generation network and resource discovery is a basic component of Grid resource management. This study presents a hierarchical resource discovery method to make use of attractive features of lumped and exhaustive methods. In this study, the resources in Grid are setup in a 3-layer resource organizational model including Physical Network Layer, Resource Information Layer and Index Information Layer. A resource discovery method based on the above hierarchical model is provided. The simulation result proves that the method discussed is of fine extensibility and performance. Furthermore, it can discover a required resource concurrently to improve the efficiency.

Key words: Resource discovery, resource organization, virtual organization, hierarchical model

INTRODUCTION

With the development of Internet and intelligence, it becomes more and more important to integrate varieties of types of resources in Internet. It is the Grid technology that is thought of as the important method to achieve this goal (Zhiwei *et al*., 2004). On the base of resource sharing and integration, the abilities of discovering, allocating, negotiating, monitoring and managing network-accessible resources are essential to achieve various end-to-end or global qualities of services. Thus grid resource discovery becomes a basic component of Grid resource management, which is the core of Grid and provides a transparent global resource view for the resource schedule and application.

Compared with the resources in traditional distributed and cluster systems, the grid resources are much more heterogeneous and wide-area distributed, which make the resource discovery efficiency affect the Grid performance in a larger degree. Therefore, a new kind of resource discovery method should be adapted to replace the centralized method from both the efficiency and scalability.

Globus (Xing *et al*., 2005) uses MDS which is LDAP based and hierarchical. GRIS provides information on resources. GIIS accumulates information from GRIS and GIIS of lower level. The aggregate directory services provide a local resource view related to VO, but it is not from a global resource point and the information is passively updated. Web service uses UDDI (Ellahi and Kechadi, 2004) which is a centralized global directory server to form a global resource view, but the information synchronization among lots of sites will consume network bandwidth. Vega Grid (Fangpeng *et al*., 2003) uses VIRD which is SD-RT based and encapsulates the physical resource into services. It can shield the physical information such as the location of resource. However, the discovery efficiency decreases when SD-RT extends to the worldwide Grid.

Grid computing and P2P computing models share more features in common than we generally recognize (Mastroianni *et al*., 2005). According to different implementation, the P2P resource discovery can be classified into three kinds: Centralized Directory of resource information based (such as Napster (Shedon *et al*., 1995) Flooding based (such as Gnutella (Marzolla *et al*. 2005) and Distributed Hash Table based (Sylvia and Scot, 2002). There are still many improved methods based on them, such as provided a decentralized resource discovery method (Zhu *et al*., 2004).

Generally, the existing resource discovery mechanism can be divided into two kinds (Xing *et al*., 2005): one is centralized mechanism and the other is distributed. The former is structurally simple and of high security, while it often causes single-point failure and performance bottleneck as well as weakly scalability. On the contrary, the latter has nice scalability and reliability but lower security.

Corresponding Author: Yulan Yin, Department of Mathematics and Physics, Anhui University of Science and Technology, Huainan Anhui, 232001, China

Obviously, neither centralized nor distributed resource discovery mechanism are ideal, thus an eclectic scheme based on hierarchical model is presented in this study.

HIERARCHICAL RESOURCE ORGANIZATION MODEL

The basis of resource discovery is resource organization. This section describes the hierarchical resource organization model thoroughly.

The hierarchical resource organization model is composed of three layers as shown in Fig. 1. The bottom layer called Physical Network Layer includes many kinds of Resource Nodes. The links between resource nodes are physical links in Internet. Each registered resource node has some Resource Information Nodes in the middle layer-Resource Information Layer. All the resource information nodes are split into a number of local organizations named Virtual Organizations, which adopt star topology structure with a Super Node in the center. All the super nodes form the top layer-Index Information Layer, which adopts ring or di-ring topology structure.

A node of the resource information layer, which is only used to store the detail and accurate resource information of a resource node, is only a logical one. All resource node registered in Grid may share several types of resources, each of which is associated with a resource information node. Therefore, the relationship between resource information nodes and resource nodes is n: 1, which means a resource information node has only one corresponding resource node while a resource node has a number of mapping resource information nodes.

The resource provider should set up different sharing priorities for the shared resources and the resource with the highest priority is called the Local Interest of a resource node. The local interest can improve the discovery efficiency by avoiding information redundancy.

The partition of virtual organization is based on the IP addresses and on resources sharing of nodes. The resource nodes with neighboring IP addresses and with the same sharing resources fall into one virtual organization. Figure 2 illustrates the virtual organization of Resource Type_a, where node A is the super node. Nodes A to G are resource information nodes whose corresponding resource nodes share resource Type_a. The local interest A, B, C and D is of Type_a, while the other nodes are not of Type_a.

Meanwhile, the count--the number of nodes in a virtual organization-- should be limited in a certain range to make sure that its scale is moderate, thus both upper-bound and lower-bound of the count are set. The virtual organization whose count exceeds the upper-bound should be split into two or more virtual

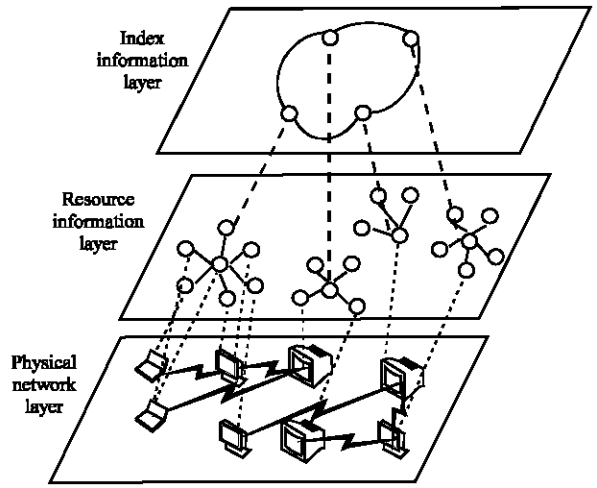

Fig. 1: The 3-layer hierarchical resource organization model

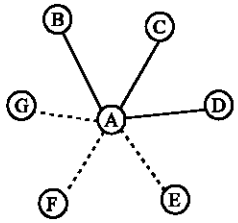

Fig. 2: The virtual organization of Resource Type_a

Name	Type	Total Cnt	T/F	Max Cnt

Name: Virtual orgnization name.
Type: Resource type.
Total Cnt: Total count of resource type in virtual Name.
T/F: If there exists some nodes whose local intersts are Type, T/F = T, else T/F = F.
Max Cnt: If T/F = T, it is the max count of resource Type, among the node with local interset Type, else it is the max count of resource Type among the nodes sharing resource Type

Fig. 3: Information base formats

organizations, whereas the one whose count is less than the lower-bound should merge with others. Each super node of a virtual organization should be the steadiest, the steadiest node is the node with longer live cycle and enough memory to store Information Bases and Adjacent Lists, so that the resource information will seldom be transferred.

The super nodes refresh Information Bases and Adjacent Lists, which store the general resource information of virtual organizations. The formats of information base and adjacent list are shown in Fig. 3 and 4, respectively.

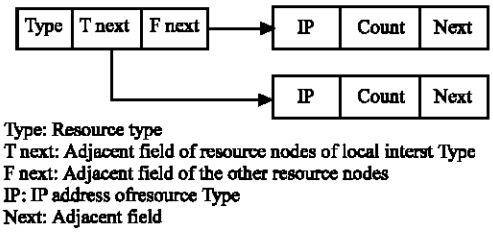

Fig. 4: Adjacent list formats

Note that the adjacent list is in descending order by the value of Count field. The information base describes the general situation of a virtual organization, thus the existing probability of a required resource can be decided; moreover, the adjacent list organized in accordance with the resource type depicts all the resources in a virtual organization and it is. Before locating a resource node to discover the required resource, both of them improve the discovery efficiency.

HIERARCHICAL RESOURCE DISCOVERY

Based on the above resource organizing model, the resource discovery method can be given.

Discovery method: The following four phases consist of the discovery procedure: submission, sorting, searching and locating.

- **Submission phase:** If a resource of Type_a is needed, the user lookups the local area to find out whether there exists the needed resource. If not, the user submits his requires for required resource of type a to index information layer.
- **Sorting phase:** According to the submitted resource information, order the super nodes by resource information in information bases whose Type = Type_a. Compute the value of *w* by formula (1) and sort the super nodes into descending order.

$$w = \begin{cases} p_1 \times \text{TotalCnt} + (1 - p_1) \times \text{MaxCnt} \\ \qquad\qquad T/F = F \\ p_2 \times \text{TotalCnt} + (1 - p_2) \times \text{MaxCnt} \\ \qquad\qquad T/F = T \end{cases} \quad 0 < p1 < p2 < 1 \quad (1)$$

Where p_1 and p_2 are empirical values which can be obtained by experiments.

- **Searching phase:** According to the type of the needed resource, for example of Type_a as above mentioned, scan the related adjacent lists whose Type = Type_a to find out proper nodes which may provide the required resources. In advance, the related adjacent lists is divided into two parts - part-interest and non-part-interest-and is arranged orderly in the way that the head of the items in adjacent lists with a certain number of the resource with Type = Type_a is put ahead of those items with a smaller number of the same kind of resources.

  Several virtual organizations can be looked through concurrently because of their independency.
- According to the result of searching phase, locate the proper resource nodes or return to phase 3.
- Locating phase: scan the detail resource information of the resource information node to ensure whether there is the required resource. If there is, provide the user with IP address or other information of the chosen resource node in the physical network layer. If there is not, return back to phase 3 to search another virtual organization.

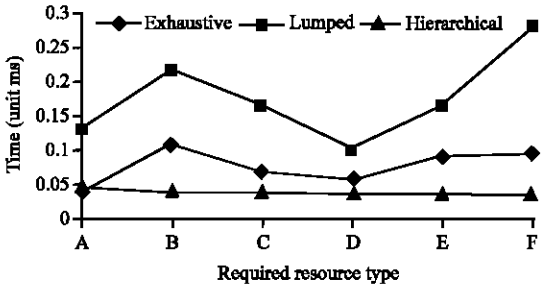

Fig. 5: Results with 100 resource nodes

Simulation analysis: We arrived out a series of simulation to test the effectiveness of this hierarchical resource discovery model and to determine the value of p_1 and p_2 in formula (1). Considering the simplicity, it is supposed that the resource information during discovering phase is static and only one resource requirement is submit one time. The simulation results are shown in Fig. 5-7. Note that the three figures are shown the best results of hierarchical methods which adopt $p_1 = 0.4$ and $p_2 = 0.2$.

Compared with an exhaustive and lumped discovery method in the same situation, the simulation results prove the hierarchical method has the following advantages:

- **Time:** The hierarchical method is shorter than the other two methods as a whole, but sometimes it maybe longer because of the different required resource.
- **Stability:** The hierarchical method is more stable than the other methods. It is almost a beeline while the others are carve lines.

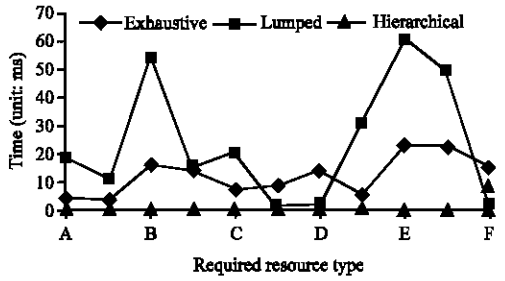

Fig. 6: Results with 500 resource nodes

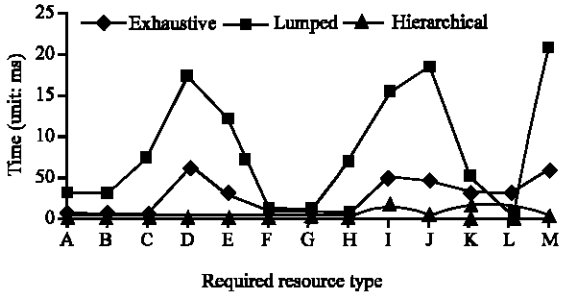

Fig. 7: Results with 1000 resource nodes

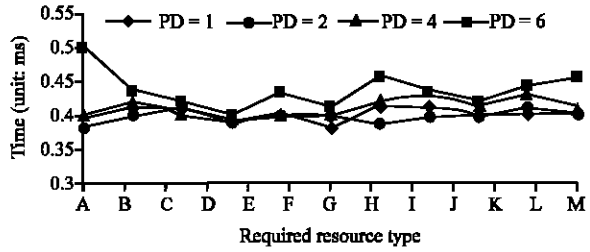

Fig. 8: Parallel discovery results with 1000 resource nodes

Furthermore, the hierarchical resource discovery method can be preceded on several virtual organizations at the same time. The number of virtual organizations being looked up is called Parallelism Degree (PD). The parallel discovery results illustrated in Fig. 8 shows that the shortest time can be obtained when PD = 4.

CONCLUSION

Being a basic component of Grid resource management, resource discovery is the core of Grid. Both exhaustive and lumped methods have some disadvantages, this paper presents a compromise method called hierarchical method.

This method organizes the resource nodes from Physical Network Layer and stores resource information of each resource node in Resource Information Nodes separated into Virtual Organizations which are in the middle layer named Resource Information Layer. One of the steadiest nodes is chosen to store the Information Bases and Adjacent Lists of corresponding virtual organizations, named Super Nodes, forming the Index Information Layer.

When a resource requirement is submitted to the Grid, the required resource information firstly passes to the index information layer. Then, the virtual organizations are sorted by their w value figured out via formula (1). Further, scan each virtual organization according to the descending sort order to find out the possible resource nodes sharing the required resource. Finally, step into the resource information layer to decide whether these resource nodes could provide the required resource.

Since this method is a compromise between lumped and exhaustive methods, it has both high scalability and efficiency and it does not have performance bottleneck and single-point failure of lumped method, as well as blindly discover of exhaustive method. Moreover, it can be carried out on several virtual organizations concurrently.

ACKNOWLEDGMENTS

This study is supported by Mountain Tai Scholar Project of Shandong Province, Peoples Republic of China.

REFERENCES

Ellahi, T.N. and M.T. Kechadi, 2004. Distributed resource discovery in wide area grid environments LNCS, Springer-Verlag Berlin Heidelberg. J. ICCS., 2004, 3038: 210-217.

Fangpeng, D., G. Yili and L. Wei, 2003. Research on resource discovery mechanisms in grids. J. Comput. Res. Develop., 40: 1749-1754.

Marzolla, M., M. Mordacchini and S. Orlando, 2005. Resource discovery in a dynamic grid environment (J). Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA'05). IEEE., pp: 1529-4188/05.

Mastroianni, C., D. Talia and O. Verta, 2000. A Super-Peer Model for Building Resource Discovery Services in Grids: Design and Simulation Analysis, Sloot, P.M.A. *et al.* (Eds.), EGC. LNCS., 3470: 132-143.

Shedon, M.A., A. Duda, R. Weiss and D.K. Gifford, 1995. Discover: A resource discovery system based on content routing (J). Computer Networks and ISDN System, 27: 953-972.

Sylvia, R., S. Scott and S. Ion, 2002. Routing algorithms for DHTS: Some Open questions. In: Proc. The 1st Int. Workshop on Peer-to-Peer Systems (IPTPS'02), Heidelberg: Springer-Verlag.

Xing, L., X. Weidong, X. Lei, L. Jian and T. Jiuyang, 2005. Grid resource discovery mechanism based on compound topology. Eng. Applic. Comput., 9: 132-136.

Zhiwei, X., F. Baiming and L. Wei, 2004. Grid Computing Technology. Beijing: Publishing House of Electronics Industry.

Zhu, C., Z. Liu and W. Zhang, 2004. Decentralized Grid resource discovery based on resource information community. J. Grid Computing, Springer (2005), 2: 261-277.