

The Grid Resource Discovery Model based on RWSRF

Jianhu Zhu^a, Fang Cheng^b

School of Chongqing University of Posts and Telecommunications, Chongqing 400065, China

^azhujianhu2008@163.com, ^bChengfang@cqupt.edu.cn

Keywords: Grid, P2P, RESTful Web Service.

Abstract. Resource discovery is an important part of grid computer technology research. However, due to the extensiveness and heterogeneity of resource distribution, it makes research discovery more challenging. In this paper, a kind of grid discovery model based on RWSRF (RESTful Web Service Resource Frame) is put forward. This model is a combination of P2P and RESTful Web Service technologies, adopting the combination of centralized and distributed frame. It would release resource provided by grid nodes in RESTful Web Service. The experiment shows that the model has improved greatly in terms of efficiency indicators compared with the traditional method.

Introduction

The grid is a kind of distributed computing environment, which integrates the available resources such as the computer, memory and database on different geographical location by the Internet, to share resources and information [1]. This way of resource sharing has advantages of low cost and high efficiency. But it will make the discovery of resources be more and more complex, with the growing size of the grid and the growing number of the grid environment resources.

Resources in the grid have the characteristics of heterogeneous, widely distributed and dynamic change. Currently the way of the grid resource discovery is roughly divided into two types: centralized and distributed. The approach of centralized resource discovery completes resource publishing and discovery, mainly depending on serving node in the center. A typical service of centralized resource discovery is DNS, as well as the Globus' MDS. The mechanism centralized resource discovery is very efficient to manage and easy to query, however, is poor to expand, vulnerable to external attacks, and a single point of failure problem due to Relying too much on the central node. So the centralized way found bottlenecks, can't effectively play a role in a dynamic distributed environment. All the status information of the mechanism of the distributed resource discovery is on each segment peer network nodes, and each node is autonomy, as well as there is no central node. The way overcomes this centralized defect well, and adapts to the dynamic and wide-area environment. However, to find the resources needed from a large number of distributed network resources, it will result in a lot of network traffic. With increasing the number of nodes, the efficiency of the approach will find more and more low.

For the grid resource discovery, researchers have proposed many different methods. [2] proposed a semantic annotation of service behavior grid discovery method based on Web Service. Although recall and precision are improved in a certain extent by the method, the method can't using simple and direct to perform resources. [3] proposed a framework for distributed resources, which can solve the problem of heterogeneous resources, but the implementation is difficult and costly.

Considering the above observations, this paper proposes a RWSRF grid resource discovery model, which combines the centralized and distributed to build an architecture model by the blend of P2P and RESTful Web Service technology, in the grid environment. In a certain extent, the model avoids the defects of the centralized and distributed.

Key technologies for RWSRF

RESTful Web Service Technology. RESTful Web Service [4] Dr. Roy Fielding presented in his dissertation, is consistent with Web service architecture of REST (Representational State Transfer)

style. RESTful Web service as a resource-oriented architecture (Resource Oriented Architecture, ROA), can abstract all things to be resources URI identified on the Web, then display them to an external by these URIs, and one URI often corresponds to a resource. Using the same interface, the users and resource providers can operate resources with the methods of PSOT, GET, PUT, DELETE, meanwhile, ensuring the dynamic nature of grid resources. RESTful Web Service allows the use of more lightweight description format, such as JSON. RESTful Web service itself is a cross-platform, distributed architecture, and is more efficient, simple and easy to be used, which integrates with the grid environment well.

P2P. The network architecture of P2P (Peer-to-Peer) [5] is completely different with a traditional Client / Server or multi-server network. In P2P network, anyone that can access the machines can be a node (peer), or even a computer application above. Two nodes in the network can be connected, and communicate with each other, without relying on a centralized server. Furthermore, P2P network has high availability and fault tolerance, which has been applied in the current epidemic distributed computing well.

Resource Discovery Model Based on RWSRF

Based on RWSRF, entire grid resource discovery model incorporates a RESTful Web Service technology and the thought of P2P, which combines distributed and centralized. The model structure is shown in Figure 1.

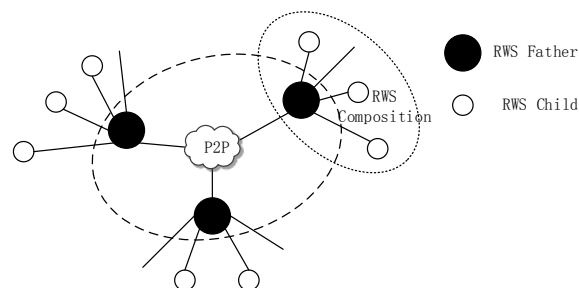


Fig. 1 Resource discovery model based on RWSRF

In the grid system, RWS Fathers (RESTful Web Service Fathers) are at the central server, which publish resource type of the RWS Childs (RESTful Web Service Childs), provide for nodes to retrieve resource information with each other. The RWS Childs keep the basic information of resources for external to retrieve and call in the form of interfaces and parameters. RWS Fathers and RWS Childs consist of RWS Compositions. The combination is a centralized resource discovery resource model, and resources in the RWS Childs are often in the same category, which are conducive to manage and maintain resources, and can reduce the loss of traffic. Nodes in RWS Compositions connect each other via a P2P network. All peer nodes are equivalent and can access each other.

RWS Father and Child description

SOAP-based Web services use WSDLs (Web Service Description Languages) to describe the operations, however RESTful Web Services typically use non-normative documents that people can read to describe, for example, Web pages [6]. There is no equivalent standard with WSDLs. WADL SUN Microsystems proposed uses the idea of WSDL to describe Web applications. WADL defines the way to generate URIs for the operations, which uses XSD (XML Schema) to describe the input and output information. However, WADL only provides descriptions in the syntax level, and, semantic annotation will not work.

For WADL, hRESTS provides a lightweight mechanism to describe Web services. HRESTS describes the text information of RESTful Web services, in the form of HTML (Hypertext Markup Language) pages. These descriptions can either make people to understand, but also allow the machine to identify.

But hRESTS is only able to provide the description for Web service in syntax level and can't relate to the semantic level, which will seriously affect the recall and precision of service, so we introduce

MicroWSMO micro formats in this paper. The MicroWSMO is hRESTS extension on the basis of SAWSDL, and provides four semantic descriptions for RESTful Web services. Therefore, this paper adopted the service description as shown in figure 2.

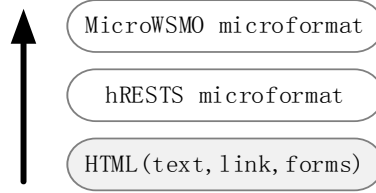


Fig. 2 The hierarchical structure of the service description

Correlation Algorithm

RESTful Web Service Matching Algorithm. The most fundamental in calculating the similarity of RESTful Web Service is the calculation of similarity between the parameter concepts. This paper uses the distance between ontology concepts to calculate the similarity of between the concepts of the parameters [7] distance. The following formula:

$$Sim(C_1, C_2) = 1 - \sqrt{\frac{1}{2} a * Dis(C_1, C_2)}$$

$Sim(C_1, C_2)$ Is the similarity between the two concepts C_1, C_2 $Dis(C_1, C_2)$ is the shortest distance between two concepts C_1, C_2 in the classification tree and the sum of the weights of edges, the formula is:

$$Dis(C_1, C_2) = \sum_{i=1}^n weight_i$$

T is an adjustable parameter. $a = \frac{Dep(C_2)}{Dep(C_1) + Dep(C_2)}$ Is the depth of $Dep(C)$ in the classification tree,

whose size is the distance to the root node, i.e. $Dep(C) = \sum_{j=1}^n 1$.? N is the shortest distance to the node of the root, when $Dep(C) = 0$, C is the root node.

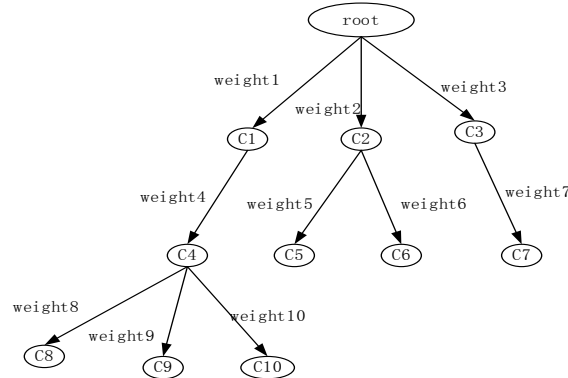


Fig. 3 Ontology concept classification tree

RESTful Web Service described by the language in Figure 3 [8] generates the corresponding semantic concepts, of course, which requires the participation of domain ontology library. But, for domain ontology library this article does not go into too much. HTML document of services, hRESTS and MicroWSMO describe a set of parameters. Because we only concern inputs and outputs here, so the service can be defined as $S_p = (I_p, O_p)$, the request can be defined as $S_r = (I_r, O_r)$, the similarity between the two is:

$$Sim(S_p, S_r) = \sum_{i=1}^n \alpha Sim(I_p, O_p) + \beta Sim(O_p, O_r)$$

$Sim(I_p, O_p)$ And $Sim(O_p, O_r)$ are respectively Similarity between the service and request input, output. α, β are weights, and $\alpha + \beta = 1$.

Resource Discovery Algorithm. In this model, each RWS Father in RWS Composite includes brief information of its RWS Childs. So we only need to use flooding algorithm to find RWS Child, which reduces the flow loss and improves the search efficiency of the system. Specific algorithm for pseudo code is as follows:

Input: the request condition set R, satisfaction require set ε

Output: request result set P

Begin

In a set of RWS Composition, matching the RWS Father to search for qualified R result set p;

If some resource information in RWS Father of RWS Composition meets the conditions R, match its RWS Childs which meets that the similarity is greater than ε , then output P_1 of meeting conditions;

Else

According to the Strategy, send request condition R, wait for the results set P_{p2p} to return;

If timeout or return results

$P = P_1 \cup P_{p2p}$;

End If

End If

Return P;

End Begin

Simulation and Analysis

To assess of the grid resource discovery model based on RWSRF, we set up a test environment which includes 13 computers whose CPU is Intel i3-3220, size of Memory is 4GB, and operating system is Windows XP. Using C # to program and building RESTful Web Services based on WCF in Microsoft Visual Studio 2010, and database is SQL server 2008. Resources were divided into 4 categories and the number was over 3200. The model included a total of 13 RWS Fathers and 34 RWS Childs. Internal used local area network (LAN) in the laboratory, and computer IP set was 172.24.54.11 ~ 172.24.54.23.

Due to the lack of the related field ontology library, so this paper designed a simple domain ontology library for the experiment after referencing to the development principle of domain ontology library and the dictionary, which avoided subjective factor. In order to be able to verify the efficiency of the model have an advantage than traditional method, the test only used the GET method.

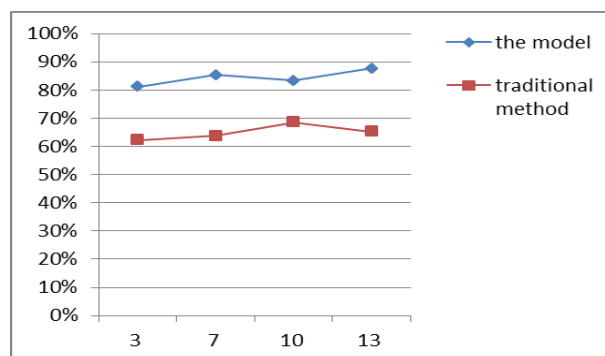


Fig. 4 Recall ratio

The horizontal axis of figure 4 and 5 is the number of computers, and both figures show that this model is compared with the traditional model in efficiency indicators. As shown in Figure 4 this model is higher than traditional method in terms of recall ratio, and close to the traditional methods in terms of precision, because the model in this paper is based on semantic to match and discover resources, however traditional methods are often based on keywords and syntax. The simulation experiment shows that the model has improved greatly than the traditional methods in terms of efficiency indicators.

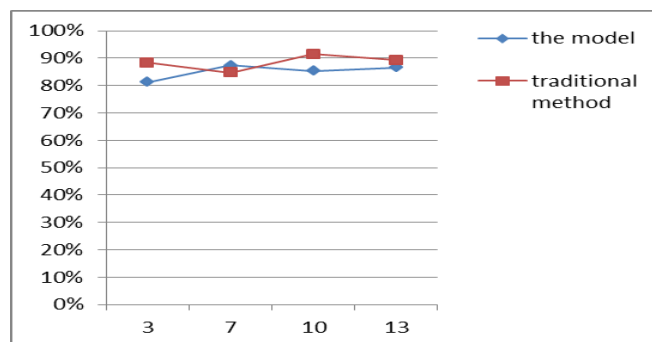


Fig. 5 Precision ratio

Summary

In this paper, with P2P and RESTful Web Service technology, we propose a grid resource discovery model based on RWSRF, which is a collection of centralized and distributed and reduce defects of both in some extent. However, only the efficiency indicators of the model had compared with the traditional ways in simulation of the paper, which didn't compare in response time, traffic loss. In the future the model will need to be improved, and make further experiments in response time and traffic loss to verify the superiority of the model.

Acknowledgments

Foundation Item: Major Science and Technology Projects, Next-generation Broadband Wireless Mobile Communication Network (2012ZX03005008).

References

- [1] Iamnitchi A, Foster I. On fully decentralized resource discovery in grid environments [C] // International Workshop on Grid Computing. 2001: p.51-62.
- [2] Yao Qing, Chen Xingyuan. Service Discovery Algorithm Using Semantic Technology in Grid Environment [J]. Computer Science. 2012 (39) p.54-57.
- [3] Charily Pittaras, Mattijs Ghijsen, Adianto Wibisono, Paola Grosso, Jeroen van der Ham, Cees de Last. Semantic Distributed Resource Discovery for Multiple Resource Providers[C]// Eighth International Conference on Semantics, Knowledge and Grids.2012: p.225-228.
- [4] Weijun Qin, Qiang Li, Limin Sun, Hong song Zhu, Yan Liu. Rest Thing: A Restful Web Service Infrastructure for Mash-up Physical and Web Resources[C]// Ninth IEEE/IFIP International Conference on Embedded and Ubiquitous Computing.2011: p.197-204.
- [5] Kopeck J, Vitvar T, Fensel D. MICROWSMO: SEMANTIC DESCRIPTION OF RESTFUL SERVICES [EB/OL]. (2008-1) [2014-1-13]. Http: //wsmo.org /TR/d38/ v0.1/2 0080219/ d38v0 120080219. Pdf.
- [6] Xia Zhao, Enjie Liu and Gordon J. Clap worthy. A Two-Stage RESTful Web Service Composition Method Based on Linear Logic [C] // Ninth IEEE European Conference on Web Services. 2011: p.39-46.
- [7] XU Dezhi, ZHENG Chun-hui, K. Passi, Concept semantic similarity research based on SUMO. Computer Applications [J], 2006, p. 180-183.
- [8] Gehao Lu, Tengfei Wang, Guojin Zhang, Shijin Li. Semantic Web Services Discovery Based on Domain Ontology.