A Scalable Method for Efficient Grid Resource Discovery

Yan Zhang[1], Yan Jia[1], Xiaobin Huang[2], Bin Zhou[1], and Jian Gu[1]

[1] School of Computer Science, National University of Defense Technology, Changsha,410073 China
Jane325@tom.com
[2] Department of Information Engineering, Air Force Radar Academy, Wuhan, 430019 ,China
Hxbtougao@gmail.com

Abstract. How to discover resource rapidly and exactly in distributed and heterogeneous grid environment is a key problem that affects grid computing performance. In this paper, the P2P method is used to improve scalability of resource discovery mechanism, and a decentralized resource discovery method with well scalability is presented, this method uses binary tree to manage data, each node in grid is responsible for managing a part of resource information. Experimental result shows that the method presented in this paper can resolve many problems that exist in centralized mechanism, such as poor scalability, heavy load on resource information server and single point failure.

Keywords: Grid Computing, Resource Discovery, P2P, Binary Tree.

1 Introduction

The grid [1] is a large scale distributed computing environment supporting scientific applications that require high throughput computation. The computing resources in the grid are workstations and supercomputers distributed across the Internet. Resources are contributed to the system by a number of different autonomous entities. Applications have different requirements for the number and capabilities of the resources to complete their execution. So, how to look up resources rapidly and exactly in grid environment is important to the grid performance.

The early grid resource discovery methods are mainly based on centralized mechanism. As grid size grows rapidly, we should decentralize resource discovery methods to avoid performance bottlenecks. P2P [2] systems have been proven to be highly scalable and highly available which are indicated by the popularity of file sharing applications among users. So using P2P principle to design a well scalable grid resource discovery method will be a good idea.

In this paper, we use binary tree to manage resource information. Each node in the grid manages a part of resource information and maintains a routing table in order to dispatch queries to neighbor nodes. Resource providers publish their resource information to corresponding node based on some rules. Therefore, each node in the grid can acts as resource information server, which not only avoids putting too much load on one resource information server, but also resolves the problem of poor scalability existing in centralized discovery methods.

Y. Luo (Ed.): CDVE 2007, LNCS 4674, pp. 97–103, 2007.

The rest of this paper is organized as follows. Section 2 presents the related work. Section 3 introduces how to allocate, manage and publish resource information. Section 4 discusses how to process resource query and gives the resource search algorithm. Section 5 gives our grid prototype and presents experimental results. Finally, section 6 presents the conclusion.

2 Related Work

A lot of work has been done on the grid resource discovery mechanism. Early grid resource discovery is based on centralized mechanism, Globus Toolkit [3], Condor [4] and Legion [5] are the excellent examples. The MDS-4 (Monitoring and Discovery Services) of Globus Toolkit provides a Web Service Resource Framework (WSRF) [6] compliant implementation of the Index Service, as well as novel mechanisms for delivering notifications in the presence of events that match a set of specified rules (Trigger Service). Matchmaker in Condor uses a centre server to match the attributes in the user's specification and those in the service providers' declaration. Such approach has a single point of failure and scales poorly. In Legion, Collections, the information database, are populated with resource description. The scheduler queries the Collection and finds proper resource for applications. A few global Collections will prohibit the scalability of the system.

Much work has been done to improve the scalability of resource discovery methods. A typical solution is offered by P2P models. In [7], information nodes are organized into a flat unstructured network and various request-forwarding policies are studied. This approach suffers from higher numbers of required hops to resolve a query compared to our approach and provides no lookup guarantees, i.e., an unsuccessful lookup does not necessarily mean that no resource meeting the requirements because a suitable peer was simply not reached. A DHT based service discovery approach is studied in [8]. Another study [9] organizes information nodes into hierarchical topologies to reduce redundant messages. An information node grouping technique is also studied in [10], in which information nodes are randomly grouped together. In [11], Talia et al. propose a P2P architecture for resource discovery in OGSA-compliant Grids, within each VO (Virtual Organization), a hierarchy of Index Services provides information about local resources. Discovery messages are routed across Peer Services using a modified Gnutella protocol. In this method, the centralized mechanism is used when searching resources within one VO. All these methods are not totally decentralized, and some of them also adopt centralized mechanism, whereas our method is a fully distributed one by making each node in the grid as information server.

3 Management and Publication of Resource Information

3.1 Management of Resource Information

The resources provided by nodes have feature of diversity, they can be data, service or computation resources. The resources described in this paper are computation resources. This kind of resource commonly has the following attributes: CPU_GHz