

Ontology-Based Semantic Method for Service Modeling in Grid

Bin Cheng^{1,2*}, Xingang Wang^{1,3}, and Weiqin Tong¹

¹School of Computer Engineering and Science; Shanghai University; Shanghai 200072; China

²College of Mathematics, Physics and Information Engineering; Zhejiang Normal University; Zhejiang Jinhua 321000; China

³College of Information Engineering, Zhejiang University of Technology, Hangzhou 310014, China

cb@shu.edu.cn, wxg@zjut.edu.cn, wqtong@mail.shu.edu.cn

Abstract. Grid is a newly developed technology for complex system with large-scale resource sharing, wide-area communication, multi-institutional collaboration, etc. A service-based approach for Grid can improve the extensibility and interoperability. In this paper, Service Oriented Grid Hierarchical Architecture is proposed within the OGSA framework, which gives a new approach to build the Grid. The need of semantic component in Grid to discover and describe the resources is analyzed and a Star Model of Ontology is introduced to describe services semantically. A Grid Service Matchmaking Semantic method is presented on the base of the ontology.

Key words: Grid service, Ontology, Grid

1 Introduction

Grid is a newly developed technology for complex system, which has been widely adopted in scientific and technical computing. It enable us to share, exchange, discover, select and aggregate geographical or Internet-wide distributed heterogeneous resources- such as sensors, computers, databases, visualization devices and scientific instruments. In Grid where resources are generally owned by different people, obtaining and managing these resources is not a simple task for communities or organizations with varied administration policies and capabilities. And the resource sharing and the problem solving in dynamic multi-institutional virtual organizations are difficult[2].

Grid technologies have been evolving toward an Open Grid Services Architecture (OGSA) in which a Grid provides an extensible set of services that virtual organizations can aggregate in various ways[1]. However, OGSA has not indicated how to build the Grid architecture and the platform directly based on service. A novel hierarchical structure of grid is proposed in this paper which can

be described as an extension of the current Grid. It is a service oriented architecture in which entities provide services to one another under various forms of contract. The services are given well defined meaning, better enabling computers and people to work in cooperation semantically under ontology.

The rest of the paper is organized as follows: Section 2 proposed the service-oriented hierarchical structure of Grid. The critical technologies of the service layer are discussed in Section 3, such as the Grid Ontology, the service discovering, the capability evaluating and the service composing. Section 4 concludes the paper highlighting the advantages and future scope of this research work.

2 Service Oriented Grid Hierarchical Architecture

Open Grid Service Architecture (OGSA)[1], oriented Grid Service (GS), discussed in Global Grid Forum (GGF) has been evolving since it was proposed in the early 2002. According to [1], the service abstraction may be used to specify the access to computational resources, storage resources, and networks, in a unified way. But it does not discuss how to model the Grid architecture based on service. We present a novel approach to integrate Web service and Grid computing by researching the Grid functional model and OGSA[4].

First, Grid is concerned with coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations. Protocols are the building block of the function sharing and cooperation. And the service mechanism is the approach to achieve the protocols. So the service is an ideal form of resource abstraction.

Second, the Service Oriented Architecture (SOA) is introduced to manage the resources in the virtual organizations (VOs). SOA builds a uniform environment of service computing, which is composed of the resource service, the common service and the application service, in the VO layer of Grid. And the resource sharing transforms to a mapping from an abstract resource to a resource service. The cooperation becomes the interaction among different service entities. The job scheduling converts into the lifecycle management of service entities, such as creation, maintenance and negotiation.

Third, the method of workflow modeling based on Web service emphasizes the ability of the need description and studies the description language of Grid job. It works for adapting the request of application and the environmental change.

Last, the mechanism of the Grid Service security and the combination of QoS and OGSA are set up, and implements the credible Grid environment based on safe Grid Service and satisfies the request of Grid application.

The Service Oriented Grid Hierarchical Architecture (Fig 1) introduces a method to build a Grid system based on service semantically and analyzes the relationship of every layer. It describes further details of the Grid functional model based on the Grid Service technology.

- (1) Resource Layer

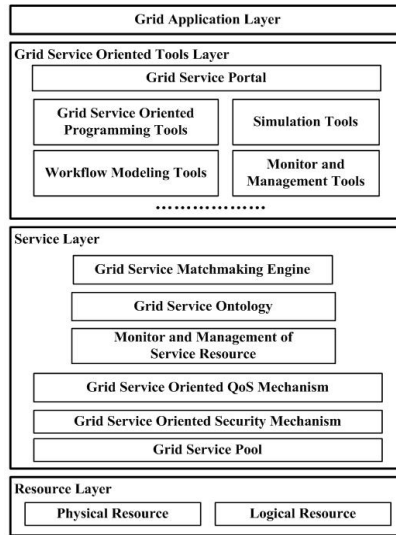


Fig. 1. Service Oriented Grid Hierarchical Architecture

Resources are the infrastructure of the Grid computing and the execution of the scientific task finally[7]. Resources not only include the physical resources such as sensors, computers, databases, scientific instruments, etc., but also the logical resources such as network bandwidth, software, application service, etc. So the resource layer highlights the distributed, autonomous and heterogeneous characteristic of the Grid environment.

(2) Service Layer

The service layer focuses on solving the resource sharing and cooperation for supporting the service oriented development, assignment, processing and testing of Grid application. The detailed components of the Services layer are presented as follows:

a) Grid Service Pool

It includes all the elementary services or complex services, which are registered, and builds the environment of the Grid Service, e.g. the uniform description of service.

b) Grid Service Oriented Security Mechanism

It offers such function as the security support, such as identification, authorization, access control, secure communication, etc. to the Grid Service.

c) Grid Service Oriented QoS Mechanism

It provides the QoS guarantee to the VOs directly and the QoS negotiation to the services. It can satisfy the nonfunctional request, e.g. service capability, reliability, management, etc.

d) Monitor and Management of Service Resource

It can analyze the service request, break apart the computing task, cooperate and schedule the subtask, offer the information service, etc. It also can

provide some common functions such as instantiation of the service, internal status maintenance, lifecycle negotiation and so on.

e) Grid Service Ontology

It is consisted of a global ontology and several local ontologies. Every local ontology records the local services' concepts and the relationships of these concepts. If it is difficult to solve a problem by one ontology, a global mapping mechanism is used to integrate the semantic information of other ontologies. It is helpful to improve the accurate service matchmaking and support the user or service to maintain and search the local ontologies.

f) Grid Service Matchmaking Engine

It supports the service discovering, the capability evaluating, the queuing strategy and the service composing semantically. The mapping from the task to the service is under the ontology. It coordinates the interaction among the services and embodies the cooperative solution of multiple services.

(3) Grid Service

Oriented Tools Layer It offers the user interface and the uniform access interface, e.g. Grid Service portal. There are programming model, test and simulation tools, monitoring and management tools and workflow modeling tools in this layer. All the tools and API can simplify the development, assignment, test and management of the Grid application.

(4) Grid Application Layer

This layer is based on service. These applications are not limited to scientific computing and mass data processing.

3 Critical Technologies of Grid Service Layer

The Service Layer is the core layer of the system in the Fig 1. The heterogeneous and distributed services, which are viewed as the standard software components, can be discovered, issued, composed and scheduled by the uniform programming environment[2][10]. The protocol group of Web service makes the open, scalable and standard infrastructure implement and interact among the services[6]. We introduce the domain ontology to improve the ability of automatic discovery and accurate location and solve the semantic heterogeneity. The global ontology and the local ontologies are integrated to share the semantic information. The service is described by OWL-S. UDDI and WS-Inspection are extended semantically to realize the service discovery and dynamic binding. We implement the service scheduling based on SOAP and its security extension. The system supports the cooperation in security based on WS-Inspection specification. And the service composing is realized under ontology technology. Compared with the 5-layer Grid system[1], the service is viewed as the unified resources in Fig 1. The novel architecture presented in the paper combines with the previous work of Web service. It is built in the autonomous and credible Grid Service environment and solves the resource connection, sharing and assembling.

3.1 Grid Ontology Architecture

The concept of Ontology, which originates from philosophy, has been widely employed by several research communities[8][9]. The use of ontology for the explication of implicit and hidden knowledge is a possible approach to overcome the problem of semantic heterogeneity. The interoperability and the semantic information integration are the key application of ontology[11][5]. An ontology provides semantics by defining concepts and properties, and by describing axioms. It includes machine-interpretable definitions of basic concepts in the domain and their relations. And it has features of domain-specificity, standardization, and evolution.

Now the research on distributed, heterogeneous Grid Service has begun to exploit ontologies in order to support semantic interoperability, because lots of Grid Service providers issue their Grid Service in a distributed environment. The shared definitions and understanding are required to discover a Grid Service [13]. It will be convenient to combine the simple Grid Services into a complex Grid Service and evaluate the capability of the Grid Service based on the ontology[14][15].

Ontology is built in the following ways and we choose the last one to model the ontology, which is called the Star Model of Ontology (Fig 2c)[8]:

(1) Single ontology approach (Fig 2a): a typical example of this approach is SIMS. This approach is susceptible to change in the information sources which can affect the conceptualization of the domain represented in the ontology.

(2) Multiple ontology approach (Fig 2b): the OBSERVER system is an example of this approach. In practice the inter-ontology mapping is very difficult to define, because of the many semantic heterogeneity problems which may occur.

(3) Hybrid ontology approach (Fig 2c): a combination of the two preceding approaches is used. A local ontology is built for each source schema, which is not mapped to other local ontologies, but to a global shared ontology. It avoids the disadvantages of single ontology or multiple ontology approaches.

Definition 1 (Ontology) Ontology is built on OWL(web ontology language). It is expressed as a 4-tuple: $O=\{C, R, I, A\}$, where C is a finite set of concepts; R is a finite set of relations; I is a set of instances; A is a set of axioms, expressed in a logical language over which is $T=C \cup R \cup I$, and it can be used to infer knowledge from an existing one.

Definition 2 (Grid Service) Grid Service is a special Web service. It is expressed as a 4-tuple: $GS=\{CA, UA, IN, OUT\}$, where CA is a set of common attributes that all Grid Services have, e.g. service name, service providers, edition, URI; UA is a set of unique attributes that a service has individually; IN is a set of parameters, describing the input interface; OUT is a set of parameters, describing the output interface.

We can describe the semantic character of an essential service by binding the interface and the service attributes. The service provider and the service requester can be ensured to call the Grid Service based on semantic.

Definition 3 (Service associating) Given two Grid Services $GS1$ and $GS2$, $IN1 \subseteq OUT2$ indicates that $GS1$ is associated with $GS2$, expressed as $GS1 \leftarrow GS2$.

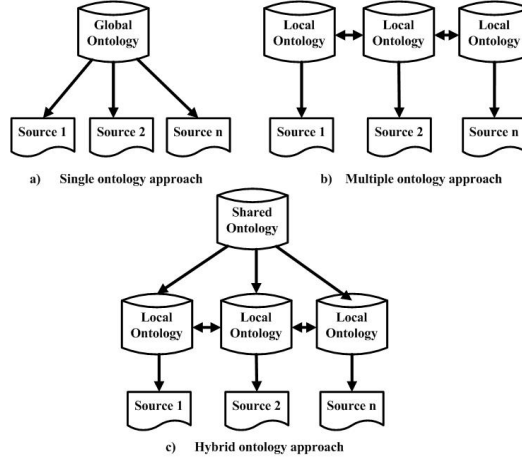


Fig. 2. The three possible ways for using ontologies

The symbol \subseteq means the set IN1 covered by the set OUT2. And the $SD(GS1, GS2)$ is the degree of the association between GS1 and GS2 based on semantic.

Definition 4 (Ontology mapping) Ontology mapping is the task of finding semantic relationships between entities (i.e. concept, attribute, and relation) of two ontologies. The ontology mapping system is a 3-tuple $OM=\{G, L, M\}$

(1) G is the global conceptual schema, expressed over an alphabet Ag. It defines the global semantics and provides the global view for users.

(2) L is the local conceptual schema, expressed over an alphabet As. L defines the local semantics for data source.

(3) M is the mapping between G and S, constituted by a set of assertions which define the relationship between global conceptual schema and local conceptual schema.

The similarity of two entities among different ontologies or between the registered and requested concepts is defined as a similarity function: $\text{sim}(e_i, e_j) \in [0, 1]$ (0 means e_i is different from e_j , 1 means they are synonymous). If there are more than one matching candidates in several ontologies for e_i , the one with the highest similarity is selected as its matched entity.

Since the ontologies evolve as time goes by, the global ontology and the local ontologies for the Grid Services should have a flexible infrastructure that has an ability to reflect the changes in ontologies[8][12].

A flexible ontology management approach is proposed for discovery and description of Grid Service capabilities supporting ontology evolution whose goal is to enhance the interoperability among Grid Services. In this approach, concepts and descriptions in an ontology are defined independently, and they are connected by relationships. In addition, the relationships are updated based on real-time evaluations of ontology users in order to flexibly support ontology evolution. A bottom-up ontology evolution means such environment that allows

ontology users to evaluate impact factors of concepts in an ontology and that results of the evaluation are reflected to the modification of the ontology. So the ontology management framework not only enables semantic discovery and description of a Grid Service capability but also supports a bottom-up ontology evolution based on the evaluations.

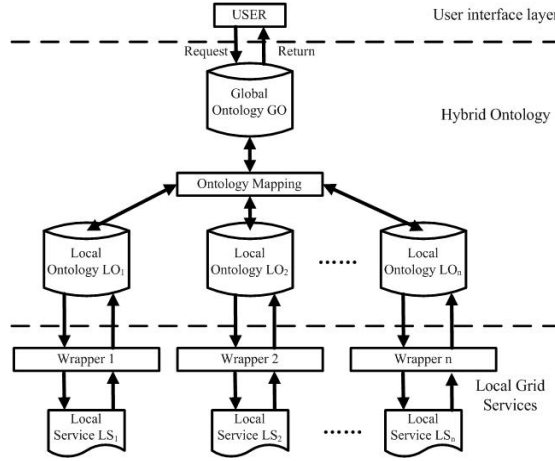


Fig. 3. Framework of the Star Model of Ontology

3.2 Grid Service Matchmaking Semantically

The matchmaking approach proposed in this paper adds semantics to the Grid Service concepts through OWL-S. The service matchmaking algorithm infers knowledge from the description to discover closely related services. We use user-defined weighted QoS factors to evaluate the capability and obtain the suitable services, then combine these elementary services as a complex one that meets the user’s requirements.

The matchmaking approach has 3 modules, which are discovering module, evaluating module and composing module, based on semantics.

(1) Discovering module

We use the semantic similarity between the requested capability and that of the registered ones to discover the service initially. The semantic similarity is determined among concepts from different ontologies directly, named SIM. $SIM = \sum (NCS, ACS, SRS) = sim(e1, e2)$, where NCS is the similarity in the name of concept, ACS is the similarity in the attribute of concept, SRS is the similarity of semantic relation and sim is defined in Definition 4. Semantic relation includes two kinds of relationships. They are hierarchical relation and in-hierarchical relation. The service discovering module can eliminate irrelevant services from being compared and improve the accuracy of the final results.

(2) Evaluating module

The chief criterion for evaluating whether the Grid Service is practicable is the Quality of Service(QoS). Users may have different QoS requirements with respect to the service requested. So we use weighted QoS factors to evaluate the capability. Six important factors are extracted to describe the QoS needed by a service requester, which are CPU cycles(CPU), network bandwidth(NB), memory space(MS), system reliability(SR), packets lost(PL), I/O bandwidth(IOB). These factors can be computed, e.g. $NB=1-(RequiredNB/AvailableNB)$. RequiredNB is the minimum network bandwidth and AvailableNB is the network bandwidth available at the service provider. Therefore, the value of RequiredNB/AvailableNB must be low. It means the service provider can offer a good QoS for bandwidth.

The parameters required by the service request such as RequiredNB are obtained from the service requester. The requester specifies the priorities of every QoS parameters, i.e. weights (W). Then computes overall QoS requested by the user as shown:

$$AQoS = \sum (w1 * CPU, w2 * NB, w3 * MS, w4 * SR, w5 * PL, w6 * IOB); \quad (1)$$

The user can choose the single service or the complex service which has the maximum value of AQoS.

(3) Composing module

The ability to compose services (and applications) based on currently available services, current context, and dynamically defined objectives and goals is critical[3][14]. While the existing systems do address many aspects of composition, they do not completely address the challenges of dynamic service composition based on semantic. We present a dynamic composition model which is to autonomically synthesize the Grid Services from the pool of available services, which are chosen after service discovery and QoS evaluation, to satisfy dynamically defined composition objectives, policies and constraints.

Definition 5 (Directed Constraint Graph)

Directed constraint graph (DCG) is a directed chart $G=\langle V,E \rangle$. And $V=\{GS1, GS2, \dots, GS_n\}$ is the set of the nodes, which represent Grid Services, in the graph. E is the set of directed edges. Every $(v_i, v_j) \in E$ indicates that there is a directed association between v_i and v_j . And the value of the edge is SD. The first node of the graph is the element IN of the request service. And the last node is the element OUT of the request service.

So the service composing can be described in the form of DCG through service associating analysis. And choose one pathway, which can satisfy the user's request and the constraints of the directed graph, to be the composed service. This model is based on the semantic similarity of concepts and the semantic relativity of the service under ontology. It can not only meet the function of service request, but also has high efficiency.

For example, GS_r is the Grid Service of the request service(Fig 4). We find the GS₁, GS₂ and GS₃ are associated with GS_r by computing SD. So there are

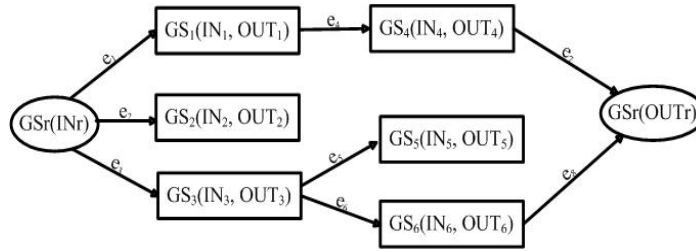


Fig. 4. Directed Constraint Graph

e_1 , e_2 and e_3 to indicate the association and the values are SDs. The other nodes and edges of the DCG are analogized in this way. Then it is found that there are 2 pathways ($e_1 \rightarrow e_4 \rightarrow e_7$, $e_3 \rightarrow e_6 \rightarrow e_8$) that can reach the end point $GSr(OUTr)$, which is the request out. We choose one of them that has high $\sum AGoS$. $\sum AGoS$ is the sum of all nodes' $AGoS$ in the pathway.

4 Conclusions and Future Work

In this paper Service Oriented Grid Hierarchical Architecture is proposed within the OGSA framework. Its supporting environment and the critical technologies are introduced. The ontology technique is used to solve the problem of describing and discovering Grid Services semantically in a heterogeneous grid environment. Then we can discover the suitable Grid Service quickly and accurately and compose the elementary service into the complex service under ontology. The service composing is a method that finds a suitable pathway in a directed graph. We discuss the problems associated with managing ontologies in Grid and present the mapping mechanism to share the semantic information in distributed environment. There are still many problems that need to be solved, such as learn local ontology from data source, semi-automatic ontology mapping, similarity computing optimization, validity checking on Grid Service composition, etc. Future work will investigate these issues more deeply.

References

1. Ian Foster, Carl Kesselman, "Grid services for distributed system integration", Computer, v 35, n 6, June, 2002, p 37-46
2. Kuo-Ming Chao1, Muhammad Younas1, " Analysis of grid service composition with BPEL4WS" 18th International Conference on Advanced Information Networking and Applications, AINA 2004 V1, p 284-289(2004)
3. Shen Derong, Yu Ge, etc. "Heterogeneity resolution based on ontology in web services composition" , in Proceedings of the IEEE International Conference on E-Commerce Technology for Dynamic E-Business, CEC-East 2004, Proceedings of the IEEE International Conference on E-Commerce Technology for Dynamic E-Business, CEC-East 2004, p 274-277(2004)

4. Hu Chun-Ming, Huai Jin-Peng, "Web Service-based Grid Architecture and Its Supporting Environment", *Journal of Software*, vol 15, No.7, p 1064-1073(2004)
5. Pan Le-Yun, Liu Xiao-qiang, "Evaluation of Multistrategy Classifiers for Heterogeneous Ontology Matching On the Semantic Web", *Journal of Dong Hua University*, v 22, n 2, April, p 55-61(2005)
6. Deqing Zou, Weizhong Qiang. "A formal general framework and service access model for service grid", *Proceedings - 10th IEEE International Conference on Engineering of Complex Computer Systems*, p 349-356(2005)
7. May Wolfgang, Alferes Jose Julio, Amador Ricardo. "An ontology- and resources-based approach to evolution and reactivity in the semantic web", *Lecture Notes in Computer Science*, v 3761 LNCS, OTM Confederated International Conferences, p 1553-1570(2005)
8. A.M. Pernas, M.A.R Dantas, "Using ontology for description of grid resources" *Proceedings - 19th International Symposium*, p 223-229(2005)
9. Hongsuda Tangmunarunkit, Stefan Decker, "Ontology-based Resource Matching in the Grid[EB/OL]".<http://epicenter.usc.edu/docs/iswc03.pdf>(2006)
10. Ludwig Simone A. , Reyhani S.M.S. "Semantic approach to service discovery in a Grid environment", *Web Semantics*, v4, n1, p1-13(2006)
11. L Ling, H Yu-jin, etc. "Semantic-based clustering method to build domain ontology from multiple heterogeneous knowledge sources", *Journal of Donghua University (English Edition)*, v 23, n 2, p 1-7(2006)
12. Hu Changjun, Zhang Xiaoming, etc. "Ontology-Based Semantic Integration Method for Domain-Specific Scientific Data", *Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, p 772-777(2007)
13. Pastore, Serena, "Introducing semantic technologies in web services-based application discovery within grid environments", *Fourth European Conference on Universal Multiservice Networks*, p 22-31(2007)
14. Yang X iandi, He N ing, "Ontology based approach of semantic information integration", *Journal of Southeast University (English Edition)*, v23, n3, September, p 338-342(2007)
15. Sudha Rajagopal, Thamarai Selvi S, Semantic grid service discovery approach using clustering of service ontologies, *IEEE Region 10 Conference*, p 414-553(2007)