Comparative Analysis of Resource Discovery Approaches in Grid Computing

Anju Sharma, Seema Bawa
Computer Science & Engineering Department
Thapar University, Patiala
{anjusharma, seema}@thapar.edu

***Abstract*- Grid technologies enable the sharing of a wide variety of distributed resources. To utilize these resources, effective Resource Management systems are needed. Resource Management system performs resource discovery to obtain information about the available resources. However, the complex and dynamic nature of grid resources make sharing and discovery, a challenging issue. Resource Discovery is initiated by a network application to find suitable resources with in the Grid. Resource Discovery process is critical for efficient resource allocation and management. For making the Resource Discovery more efficient and reliable large numbers of Approaches are there. This paper analyses some of existing Approaches for Resource Discovery, which can search for the preferred resources quickly and efficiently (return the correct results quickly and reduce network complexity) in Grid computing. Finally a qualitative comparison between these Approaches based on the factors that affect Grid Resource Discovery process, has been done and results are presented.**

***Index Terms* - Grid computing, Resource Discovery, Grid Resource Management**

I. INTRODUCTION

Grid computing provides scalable, secure and extremely high performance mechanism for discovering and access to remote computing resources in a seamless manner [1]. Huge number and the heterogeneous nature of grid computing resources make the resource management a significantly challenging task. Resource management is the process of managing available resources and system workloads accordingly. It is the manner in which resources are allocated, assigned, authorized, assured, accounted and authenticated.

Various scenarios of resource management include resource discovery, resource monitoring, resource inventories, resource provisioning, fault isolation, autonomic capabilities and service level management activities. Resource Broker, a core part of the grid system, heads the whole hierarchy of maintaining all the service information of the system. Broker performs various processes like discovering, scheduling, co-allocating and evaluating. Out of these resource discovery is the first and the foremost important process. Broker has to discover the resources before implementing other grid resource management scenarios. Resource Management system performs resource discovery to obtain information about the available resources for the particular job.

The discovery of resources that user can access through grid information server(s), negotiates with resources or their agents using middleware services, scheduling, and deployment is done by broker. Discovery in the Grid environment becomes complex as the resources are geographically distributed, heterogeneous in nature and owned by different individuals and organizations each having their own resource management policies and different access and cost models [4]. There have been many projects that have designed and implemented the resource discovery with a variety of algorithms and Approaches.

The motivation behind this paper is to explore the resource discovery mechanism, which is suitable with the Grid Environments. That is, the resource discovery should find out the preferred resources quickly and return the result back to one who requests in the manner time. Consequently, we collect the information about the previous work related resource discovery approaches. Although this paper does not propose new successful resource discovery mechanism, we expect this survey could make us know about the exist facility that could lead us to achieve our goal more or less.

II. RESOURCE DISCOVERY APPROACHES

Resource Discovery is systematic process of determining which grid resource is the best candidate to complete a job with following trade-offs.

- In shortest amount of time
- With most efficient use of resources
- At minimum cost

Resource discovery is challenging issue for efficient deployment of a grid system. Dynamic availability and heterogeneous nature make it challenging task. There are various Approaches for resource discovery in grid environments. The base for all these Approaches is query and agent based resource discovery.

In a Query based Approach discovery is the process in which resource information is queried for resource availability. Most of the contemporary grid systems follow this Approach [2]. Query based system are further characterized depending on whether the query is executed against a distributed database or a centralized database. Resource Discovery can be categorized in two forms that is Query Based Resource Discovery and Agent Based Resource Discovery.

On the other hand, Agent based Approach send active code fragments across machines in the Grid that are interpreted locally on each machine. Agents can also

passively monitor and can distribute resource information either periodically or in response to another agent. The Agent is a software entity with intelligence, autonomy and response which can interact with its surrounding and execute task on behalf of its user. The major difference between a query based Approach and an agent based Approach is that agent based systems allow the agent to control the query process and make resource discovery decisions based on its own internal logic rather than residing upon a fixed function query engine. Most agent systems are based on an underlying mobile code environment like Java. Agent based resource discovery is inherently distributed [3].

Resource discovery Approaches like Peer-to-Peer Approach, Ontology Description-Based Approach, Routing Transferring Model-Based Approach, Parameter-Based Approach, Quality of Service (QoS) Approach and Request Forwarding Approach have been considered in this paper.

A. Peer-to-Peer Approach

Iamnitchi discussed Peer-to-Peer resource discovery in detail. He proposed Peer-to-Peer resource discovery architecture for a large collection of resources [5]. This decentralized resource discovery architecture could lessen huge administrative burden as well as it can also provide very effective search-performance result. Different resource discovery problems in a large distributed resource-sharing environment especially in a grid environment are discussed [6]. In this document, author identified four different architectural components called "Membership protocol", "Overlay construction", "Preprocessing", and "Request processing". Four environment parameter factors are also identified, which dominate the performance and design strategies for a resource discovery solution are "Resource information distribution and density", "Resource information dynamism", "Request popularity distribution", "Peer participation". A brief description of different resource discovery Approaches in Peer-to-Peer networking is described [6]. The authors claimed that using four axes framework; it is possible to design any resource discovery architecture in a grid. A general-purpose query support enabled "Unified Peer-to-Peer Database Framework (UPDF)" for a large distributed system has been proposed. UPDF can be identified as a Peer-to-Peer database framework for a general purpose query support which is unified because it supports arbitrary query languages, random node topologies, different data types, different query response modes, different neighbor selection policies for expressing specific applications [7].

B. Ontology Description- Based Approach

Ontology refers to a description of a service (resource), a semantic service discovery framework in a grid environment has been proposed in [9]. Ontology enhances the interoperability between virtual organizations. They proposed a service matchmaking mechanism based on ontology knowledge and they claimed that this matchmaking framework can provide a better service discovery and also can provide close matches. The main idea behind this Approach is the advertisement of the resources. In this Approach, service provider registers its service description into the service registry database. When a Grid application sends a request to service directory, matchmaker returns the matches to the service requester. Requester chooses the best resource based on the specific need.

C. Routing Transferring Model-Based Approach

A resource discovery technique called Routing-Transferring Model is proposed [10]. This model consists of three basic components - resource requester, resource router and resource provider. The provider sends the resource information to a router and router stores that information in a router table. After that, when requester sends a request to the router, router checks its routing table for an appropriate resource provider and after finding that entry router forwards that request to the service provider or another router. The authors formalized this model and they analyzed the complexity of Shortest Distance Routing Transferring (SD-RT) algorithm based on this formalization. They claimed that resource discovery time depends on topology and they also showed that SD-RT could locate a resource in the shortest time, if the topology and distribution of resources are explicit. They examined their proposed model in Vega Grid project and their experiment shows that higher frequency and more location of resources can reduce the resource discovery time.

D. Parameter-Based Approach

Different Approaches for resource discovery in a grid system are described in [11]. A new concept "Grid potential" is proposed in this paper, which encapsulates the processing capabilities of different resources in a large network. The authors also proposed an algorithm called "Data Dissemination Algorithm". This algorithm follows swamping Approach [13] for message distribution. When a message comes to a node, that message gets validated. The validation process depends on three types of dissemination, universal awareness that permits all incoming messages, neighborhood awareness that allows messages from a certain distance, and distinctive awareness, which discards messages if it finds out that the less Grid potentiality at the local node in remote node, is less than that of the requestor node. The authors also measured the performance of "universal awareness", "neighborhood awareness", and "distinctive awareness" dissemination schemes. The authors claimed that universal Approach is more expensive in terms of message complexity than that of neighborhood and distinctive Approach. The authors also claimed that this new class of dissemination could reduce the communication overhead during the resource discovery.

E. Quality of Service (QoS)-based Approach

An algorithm to discover the occasionally available resources in a multimedia environment is proposed in [12]. In this paper, different policies for a QoS based resource discovery service for a given graph theoretic Approach. A generalized version of Discovering

Intermittently Available Resources (DIAR) algorithm based on occasionally available resources is presented in [12]. The performance of QoS policies based on different time-map strategies in a centralized system is evaluated. Various QoS parameters include processor runtime, storage capacity, network bandwidth and many more. On the basis of these parameters QoS guarantees the best behaviour of grid. Through the experiment they found out randomized placement strategies and increased server storage can facilitate better performance to discover a particular resource.

F. Request Forwarding Approach

According to Iamnitchi following four-request forwarding Approaches are identified.

a. Random Walk Approach

In this Approach, to forward the request, the node is chosen randomly.

b. Learning-Based Approach

A request is forwarded to a node who answered similar request before. If no similar answer is found, the request is forwarded to a randomly chosen node.

c. Best-Neighbor Approach

The number of received answer is recorded without recording the type of requests. The request is forwarded to that node which answered highest number of requests. This Approach to identical to learning-based Approach except when no similar answer is found, request is forwarded to the best neighbor. According to [5], the resource discovery mechanism in an "emulated" grid, which is a large grid network (for this case up to 5000 peers) based on the assumption that every peer provides at least one resource is analyzed. The measured performance evaluation of a simple resource discovery technique is based on "request propagation".

d. Learning-Based + Best-Neighbor Approach

This Approach is identical to learning-based Approach except when no similar answer is found. Request is forwarded to the best neighbor.

III. COMPARISON

Comparison between the Resource Discovery Approaches based upon various features including scalability, reliability, adaptability and manageability has been done and placed in Table I. Large numbers of parameters are to be taken into account for dealing with complexity of grid resource discovery as it becomes complex with the increase in size of grid. For dealing with the large global grids, Peer-to-Peer is the best Approach to be used, as it uses the graph theoretic Approach to achieve scalability and manageability. De-Centralized resource discovery Approach can also be used for large grids, but is limited to certain extent because of its dependent and managing units.

However, Ontology based Approach uses the central broker for matchmaking, which reduces its scalability. A new concept of Grid Potential helps the parameter Approach to achieve the efficient way to maintain the current status of resources. QoS Approach helps the clients with the reliable feedback about the expected behaviour of the system as a whole. It is observed that Routing Transferring Approach is the fastest way to discover the resources for small grids as it uses the shortest distance routing table for matchmaking. Request Forwarding Approach is not much used today because of its nonstandard nature.

IV. CONCLUSION

Resource Discovery is the process of finding the satisfactory resources according to the user's request, including resource description, resource organization, resource lookup and resource selection. In this paper we analyzed various grid resource discovery Approaches. We compare all these Approaches on the basis of performance factors like scalability, reliability, adaptability and manageability. On the basis of this comparison we have an idea about choosing an appropriate approach to discover a particular resource. Our assumption said that peer-to peer approach has succeeded in the world of resource discovery. We expect the resource discovery concepts of peer-to-peer may be adapted one service using in grid systems. However it is not easy to recognize the grid system with peer-to-peer system, so we should explore the new model continuously which can be describe the components which are sufficient for both systems.

ACKNOLEDGMENTS

We are grateful to Grid Group i.e. "TUGrid" actively working at the Centre of Excellence in Grid Computing (CoEGC) at Thapar University, Patiala (India) for providing us motivation and suitable platform for carrying out this work. The brainstorming sessions and discussions held during weekly meetings had been extremely helpful in this endeavor.

REFERENCES

[1] J. Joseph and C. Fellenstein, "Grid Computing", IBM Press Pearson Education, pp. 5.

[2] Chaitanya Kandagatla, "Survey and Taxonomy of Grid Resource Management Systems", University of Texas, Austin http://www.cs.utexas.edu/users/browne/cs395f2003/projects/KandagatlaReport.pdf

[3] Klaus Krauter, Rajkumar Buyya and Muthucumaru Maheswaran, "A taxonomy and survey of grid resource management systems for distributed computing", Copyright 2001 John Wiley & Sons, Ltd. 17 September 2001.

[4] Rajkumar Buyya, Steve Chapin, and David DiNucci, "Architectural Models for Resource Management in the Grid", Springer Verlag LNCS Series, Germany, Dec. 17, 2000, Bangalore, India.

[5] A. Iamnitchi and I. Foster, **"**On Fully Decentralized Resource Discovery in Grid Environments", IEEE International Workshop on Grid Computing, Denver, CO, 2001.

TABLE I.
QUALITATIVE COMPARISON FOR ABOVE APPROACH

Approach/ Property	Peer-to-Peer	Ontology Description	Routing Transferring	Parameter	QoS	Request Forwarding
Base Approach	agent/query	agent	query	agent	query	agent/query
Scalability	More scalable as it uses the four axes framework	Limited scalability centralized broker	Uses routing protocols for scalability hence scalable	Scalable due to grid potential concept used by it	Uses different time map strategies in centralized system to increase the scalability	Nodes are chosen randomly which makes it scalable
Reliability	Based on graph theory so reliability increases	Failures are detected as soon as they occurs so more reliable	Quite reliable as it uses the routing concept	Reliable as we can add or delete a node from anywhere	Considers parameters like network bandwidth, required CPU, storage capacity etc. that make it less reliable	Random walk Approach make it reliable in case the resources are equally distributed
Adaptability	Multiple platforms environment make it more adaptive	Can be made adaptable by providing manager information about different platforms	Routing table is used to make records of different platforms.	Adaptable due to universal, network and distractive awareness parameters	Depends upon the Service Level Agreement (SLA) sign with user for providing adaptability	Using Best neighbour Approach adaptability is easy
Manageability	Complex architecture hence difficult to manage.	Quite easy to manage as a lot of its working is dependent on single node.	Management is easy due to SDRT algorithms as it deals with different topologies	Manage the consistency by using the data dissemination algorithms	Uses algorithm like DIAR for the resource discovery	Better Management can be achieved by combining its two sub Approaches.
Complexity	O(log n)	O(log log n)	Θ(n)1/2	Ω (n)	Ω(log2n)	Θ(n)
Development Agent	DHT	OWL, RDF	C	C, Java	Any Description Language	C
Algorithm Used	Swamping	Matchmaking, Gang matchmaking	SDRT, Routing	Dissemination	DIAR	Request Forwarding Algorithm
Client Server	Each Node is Client as well as server	Yes	Yes	Yes, as it is location independent	Yes	Yes
Extensibility	Easily Extendable	Easily	Complex	Easily extended due to the distinctive awareness	Easily as well as complex	Easily Extendable

[6] A. Iamnitchi, I. Foster, Daniel C. Nurmi, "A Peer-to-Peer Approach to Resource Discovery in Grid Environments", Proceeding of the 11th Symposium on High Performance Distributed Computing, Edinburgh, UK, 2002.

[7] W. Hoschek, "A Unified Peer-to-Peer Database Framework for Scalable Service and Resource Discovery", Proceeding of Third International Workshop on Grid Computing: GRID 2002, Baltimore, MD, Springer, 2002, pp. 126-144.

[8] O. Rana et al., "Resource discovery for dynamic clusters in computational grids", IEEE Proceeding of 15th International Parallel and Distributed Processing Symposium, pp. 759 -767, 2001.

[9] S. Ludwig, P. Santen, "A Grid Service Discovery Matchmaker based on Ontology Description", Euroweb 2002 — The Web and the GRID: from e-science to e-business, 2002.

[10] W. Li, Z. Xu, F. Dong, J. Zhang, "Grid Resource Discovery Based on a Routing-Transferring Model", Proceeding of Third International Workshop on Grid Computing: GRID 2002, Baltimore, MD. , Springer, 2002, pp. 145-156.

[11] M. Maheswaran and K. Krauter, "A Parameter-based Approach to resource discovery in Grid computing systems", 1st IEEE/ACM International Workshop on Grid Computing, 2000.

[12] Yun Huang and Nalini Venkatasubramanian, "QoS-based Resource Discovery in Intermittently Available Environments", Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing HPDC-11 2002 (HPDC'02) 1082-8907/02, 2002 IEEE.

[13] M. Harchol-Balter, T. Leighton, D. Lewin, "Resource Discovery in Distributed Networks", 18th ACM Symposium on Principles of Distributed Computing, 1999.

Anju Sharma, 2005 holds M.Sc. (I.T.) degree form Punjab University Chandigarh. She is pursuing Ph.D from the Thapar University, Patiala, India**.** Her areas of interests include Grid Computing, Resource Management, Resource Discovery and Peer-to-Peer Networks.

Seema Bawa, 1970 holds M.Tech (Computer Science) degree form IIT Khargpur and Ph.D. from Thapar University, Patiala. She is currently working as Professor and Head with Computer Science and Engineering Department. Her areas of interests include Parallel and Distributed Computing, Grid Computing, Security in Grids and Network Management. Prof. Bawa is member of IEEE, ACM and Computer Society of India.