Reliable Resource Discovery Approaches for Grid Environments

Mahdi MollaMotalebi, Raheleh Maghami, Abdul Samad Ismail, and
Aboamama Atahar Ahmed

Abstract—Resource discovery is one of the most important services affecting grid performance. Because of inherent characteristics of grid environments such as heterogeneity, dynamicity and scalability, usually grid resource discovery approaches deal with some challenges to be efficient. One of the most important challenges is reliability and stability of resource discovery in grid environment. The main approaches of resource discovery in grid environments include centralized, hierarchical, structured/unstructured peer-to-peer, and super-peer/cluster. This paper investigates the challenges of above grid resource discovery approaches related to reliability and stability of the system. Also, the main techniques used to tackle with these challenges in different approaches are presented.

Index Terms—Grid, Resource Discovery, Reliability, Stability

1 Introduction

GRIDS aggregate different resources to handle complex applications efficiently. It is formed by collaborating several heterogeneous resources to achieve higher computational capacity [1], [2] . Many of scientific applications require a huge amount of processing and storage capacity much more than available facilities. The earlier systems such as parallel systems and clusters are different from grids by some special features of the grid such as heterogeneity, dynamicity and scalability of the participants and resources. A grid interconnects storage systems, scientific instruments, operating system services, and any computing infrastructure to share resources such as CPU, memory, and software applications to handle data intensive applications [3]. It is distinguished from traditional distributed systems by scope, computation size and heterogeneity. Also, it is different from cluster computing in a few aspects. In a cluster, one centralized resource manager allocates the resources and all machines cooperate as a single unified machine, while in grids, machines have their own resource managers. Another difference is that resources of the grid are heterogeneous but cluster's resources usually are homogeneous. Moreover, the number of cluster's resources is usually static and they are placed near to each other, but in grid environment, almost often they are dynamic and distributed geographically. After all it is noteworthy that a cluster can be itself a resource of the grid environment.

From the management point of view, grids can be in the form of departmental, enterprise and global. Departmental grids are in a single location and with single administration authorities with the aim of sharing idle resources of a department to do large computing applications. Also, their participants are typically similar. Enterprise grids include several departmental grids within an enterprise. They contain more varied participants with more users and different geographical locations. Users are aware of common goals and data sharing is usually well controlled. In global grids, there are several organizations located across the globe and provide resources to be used by each others. Organizations have their own goals and authorities to make decisions for resource managements [4], [5], [6]. In this paper, we address the global grids.

There exist some approaches to index and search the resources in grids including centralized, hierarchical, structured/unstructured peer-to-peer, and super-

- *M. MollaMotalebi is with the Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310, Skudai, Malaysia.*
  Email: Motalebi@gmail.com
- *R. Maghami is with the Faculty of Electrical, Computer and IT Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran.*
  Email: Maghamy@Qiau.ac.ir
- *A. S. Ismail is with the Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310, Skudai, Malaysia.*
  Email: Abdsamad@utm.my
- *A. Atahar Ahmed is with the Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310, Skudai, Malaysia.*
  Email: Aboamama@utm.my

peer/cluster approaches. Obviously with regard to characteristics of global grids such as heterogeneity, dynamicity and scalability, the resource discovery in these systems face with many challenges such as latency, reliability, quality of results, efficiency, maintenance cost and so on. This paper focuses on reliability in different resource discovery approaches in global grid environments.

2 Related work

Resource discovery is the process of locating resources required by a grid application efficiently. The required resources must be available for grid applications on demand, otherwise the application may fail. So, many recent grid studies focus on resource discovery and its related issues. Grid resource discovery methods attempt to find resources on demand and concurrently keeping the system reliable and stable. Resource discovery in grid can be classified into four main categories as centralized, hierarchical, structured/unstructured peer-to-peer, and super-peer/cluster approaches.

In centralized methods such as Globus MDS [17], resource information are gathered and indexed on one or more central servers [8]. Then all queries are handled by these central servers such that the central servers are highly important and busy. In this approach, occurring failure or low efficiency on central servers can affect the whole grid system significantly especially in a large scale environment [8], [9], [10]. Because of the unadaptable nature of centralized approaches in large scales, they are not appropriate to be used in today's grids environments.

In hierarchical methods [8], the load is distributed on more nodes typically as a tree structure. Though this approach tries to index the grid resource information in more nodes as hierarchical [10], but still it has the weakness of unbalanced load in different levels of the hierarchy so that nodes located in higher levels are highly probable to deal with huge amounts of processing, storage and network loads and subsequently they can be as bottleneck points in the grid [43].

Many of resource discovery methods follow the peer-to-peer (or decentralized) approach which distributes the resource information on different nodes on the network. Peer-to-peer architecture may be unstructured or structured. Unstructured peer-to-peer methods typically use flooding approaches which limits the grid scalability due to large number of messages; on the other hand they are load-balanced and adapted to dynamic environments [11]. Structured peer-to-peers are organized by specific criteria and topologies [12] by using some rigid structures such as distributed hash tables to index resource information. They are able to find requested resources quickly but with high maintenance costs especially in dynamic conditions [3], [7].

Super-peer and cluster based methods attempt to get together the advantages of centralized (e.g. quickness) and peer-to-peer (e.g. load-balance) approaches concurrently. Here, a set of nodes are handled by cluster heads (super-peers) as centralized, however the cluster heads themselves connect to each other and behave as peer-to-peer. These methods still suffer from bottleneck and single point of failure in cluster heads especially in high dynamic conditions [13], [14], [15], [7].

3 Reliability challenges in different resource discovery approaches

Grids may be in a small or large scale, but most often they tend to be in large-scale in terms of the number of resources and geographical distribution [7]. Also their participants and resources may be static or dynamic, but the grids nature of multi organizational resource sharing necessitates them to be dynamic. On the other hand, grid applications must access to their required resources on demand, otherwise the jobs may not be executed [16]. Regarding to the main aim of grid systems which is handling complex applications by sharing resources, availability of resources is a vital issue. So any resource discovery approach must be adapted with the heterogeneity, scalability and dynamicity of grid environments to keep the whole system reliable and stable.

From point of view of scalability and bottleneck, grid resource discovery methods in centralized approach typically suffer from potential bottleneck in central servers especially when the network is in a large scale. For example, Globus MDS [17] utilizes LDAP servers to maintain resource information and locating them which can be a point of bottleneck. As illustrated in **Error! Reference source not found.**, hierarchical methods [43], [18], [19], [10], [21] need enough amount of storage to store grid resources information which it limits the scalability of these methods.

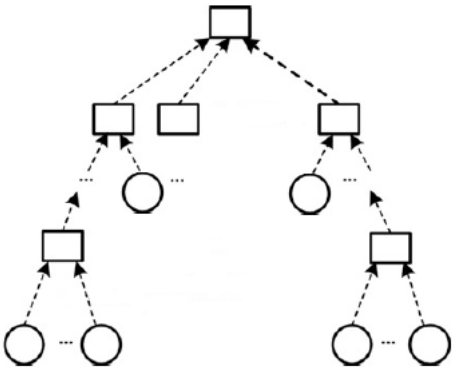

Fig. 1 Hierarchical information system schema, Higher level nodes suffer from burdensome strorage/processing loads

Decentralized (peer-to-peer) resource discovery methods inherently are not dealing with bottleneck in large-scale conditions, but some other obstacles limit their scalability. For example, unstructured peer-to-peer resource discovery methods such as Gnutella [22] send a large number of messages during resource dissemination and search. It will be worse when the network size grows [16] because numerous messages can saturate the whole grid and decrease the grid efficiency dramatically. Many of

unstructured peer-to-peer methods exploit some limiting factors to control the number of messages, but it causes another problems such as false positive error [7], [8] which is disability of traversing enough nodes in the grid network due to use of limiting factors such as Time-to-Live (TTL) for query message forwarding. So when a resource exists in the network, the resource discovery process may report lack of matched results due to disability to traverse all the resource owners in the whole grid network. Also in super-peer based methods, occurring bottleneck in index nodes is probable especially in large-scale grids. So, as illustrated in Fig. 2, redundant super-peer nodes are exploited to tackle this problem, however using redundant super-peers can increase the processing/data loads due to resource information replications.

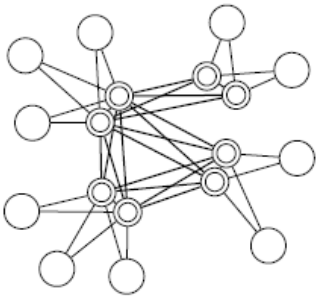

Fig. 2. Increasing the reliability by using redundant super-peer architecture, double rings are super peers and normal rings are resource owners

Super-peer based methods are more scalable than unstructured peer-to-peers [16]. Structured peer-to-peers [23], [24], [25], [26], [27], [28], [29] exploit distributed hash tables (DHT) and they are adapted to large scales due to low number of messages transferred during resource search and dissemination processes [16]. Furthermore, there are some techniques try to make nodes or resource grouping to tackle the scalability of the grid [30], [31], [32], [33], [34], [38].

The second aspect affecting on reliability of grid resource discovery is stability (or fault-tolerance) of grid nodes and their resources. Because of dynamic nature of grid environments, each node may join to or leave from the network at any time and even without any notice. So after node leaving, the index of resource discovery system will be unreliable because some resources will be refered even though they are unavailable. Also if the leaving node is an index node, then some information about grid resources will be lost, so while some resources are available but they may be inaccessible. To tackle this problem, some methods utilize backup and/or redundant peers [16] for important index nodes to keep the system reliable. Redundant peers will be updated by replication once the information of index nodes change. But these update tasks can raise the overhead of bandwidth and processing loads [35], [8]. Therefore some methods apply index updates as periodic [13], [36], [37], however it may decrease the resource information freshness and subsequently resource discovery accuracy. Moreover such systems could not support the dynamic attribute queries [7]. So, there is a tradeoff between reliability and data/processing loads [35].

Single point of failure [3], [7], [14] is a common obstacle of centralized and hierarchical techniques such that a large part of the network may depend on central servers function [19], [39]. So, in the case of the central servers leave or fail, some resource information could be lost [8]. The hierarchical organization [40], [41] can reduce the bottleneck probability, but still they deal with single point of failures because failing of an index server in the hierarchy causes the resource information of descendant nodes to be lost and many resources will be inaccessible [8]. Another weakness of hierarchical structure is lack of fresh resource information in higher level index servers in the hierarchy since resources are located in lower layers and their information will be updated in upper layers servers with latency [43]. Many researchers choose peer-to-peer approaches to tackle this problem [7], however considering the accuracy and speed of centralized indexing; it is a reasonable tendency to use hybrid (e.g. super-peer) structures [43], [13], [14].

The third factor related to resource discovery reliability is load balancing. An unbalanced grid network and/or resource discovery system can easily be infected by inefficiency, because some parts/participants of the network may deal with a large amount of data and processing loads, while some others are idle and their capabilities are wasting. Resource indexing technique and search strategy determines that a grid resource discovery system is balanced or not. Centralized and hierarchical resource discovery methods deal with unbalanced load significantly because central or higher level index servers suffer from high computation and traffic overheads [43]. Unbalanced loads imply unbalanced importance of participants/nodes in the grid so that some participants/nodes may act a vital and some others a negligible role in the system. Some techniques to increase the load balancing are exploiting polling, agents, global random choices, randomized algorithms, and local diffusion methods [44], [45], [46], [47], [48], [49].

Peer-to-peer based approaches attempt to distribute the loads between all the grid participants, however structured peer-to-peer methods can potentially cause unbalanced loads due to using specific index nodes (e.g. hash tables) [42], [20]. Unstructured peer-to-peer systems are not dependent to a limited number of index servers, which results more autonomy to the participants [11], so they are inherently load balanced. **Error! Reference source not found.** represents a summarized comparison of different grid resource discovery approaches behavior in terms of reliability factors.

Table 1. Reliability in different grid resource discovery approaches

	Scalable	Bottleneck	Fault tolerant	Load balanced
Centralized	Limited	Yes	No	No
Hierarchical	No	Yes	Limited	Limited
Structured peer-to-peer	Yes	No	Limited	Yes
Unstructured peer-to-peer	No	No	Yes	Yes
Super-peer /cluster	Yes	Limited	Limited	Yes

4 CONCLUSION

Different approaches of resource discovery deal with some challenges of reliability. This paper investigated the behavior of current main approaches from reliability's point of view. Centralized methods have fundamental problems to achieve reliability since high data and processing loads in central servers makes them a point of bottleneck and also central servers are potentially single point of failure. Hierarchical methods tried to distribute the loads on multiple points, but still higher level index servers suffer from high data and processing loads. Also, failure to the higher level index servers causes all their descendants to be inaccessible.

Because of distributed nature of unstructured peer-to-peer methods, the loads are balanced and there is no bottleneck or single point of failure, but they are reliable only in small-scale grids since their typical flooding approach for resource dissemination and discovery limits their scalability. Super-peer/cluster methods combined the centralized and peer-to-peer approaches in their designation, so they do not suffer from single point of failure and bottleneck critically and also the load is almost balanced. Moreover, they are scalable since they divide a large-scale network to multiple small-scale networks by using super peers. Because of distributed nature of structured peer-to-peer methods, the resource information gathering and indexing are handled distributedly, for example by using distributed hash tables (DHT). Also, the load is almost balanced and they are not dealt with single point of failure or bottleneck problems seriously. Therefore with regard to the behavior of investigated approaches, it can be concluded that structured peer-to-peer and super-peer approaches are more reliable for resource discovery in global grids.

REFERENCES

[1] R. Buyya, *et al.*, "Nimrod/G: an architecture for a resource management and scheduling system in a global computational grid," in *High Performance Computing in the Asia-Pacific Region, 2000. Proceedings. The Fourth International Conference/Exhibition on*, 2000, pp. 283-289 vol.1.

[2] "A taxonomy and survey of grid resource management systems for distributed computing," *Softw. Pract. Exper.*, vol. 32, pp. 135-164, 2002.

[3] P. Trunfio, et al., "Peer-to-Peer resource discovery in Grids: Models and systems," Future Gener. Comput. Syst., vol. 23, pp. 864-878, 2007.

[4] J. C. Kesler. (2003, Overview of Grid Computing. MCNC, NC. [Online].Available: www.ncbiogrid.org/resources/slides/grid-overview.ppt

[5] G. Fagg. (2003, Grid/MetaComputing Lecture. University of Tennessee. [Online]. Available: www.cs.utk.edu/~dongarra /WEBPAGES / SPRING-2003/lect074.pdf

[6] I. Foster, "What is the grid? A three point checklist," Grid Today, vol. 1, 2002.

[7] A. Hameurlain, et al., "Resource discovery in grid systems: a survey," Int. J. Metadata Semant. Ontologies, vol. 5, pp. 251-263, 2010.

[8] A. H. D. COKUSLU , K. ERCIYES, "Grid Resource Discovery Based on Centralized and Hierarchical Architectures," International Journal for Infonomics, vol. 3, pp. 283-292, 2010.

[9] R. Ranjan, et al., "Peer-to-peer-based resource discovery in global grids: a tutorial," Communications Surveys & Tutorials, IEEE, vol. 10, pp. 6-33, 2008.

[10] H. C. A. X. C. YULAN YIN, "The Grid Resource Discovery Method Based on Hierarchical Model," Information Technology Journal, vol. 6, 2007.

[11] D. A. R. TSOUMAKOS, NICK, "A Comparison of Peer-to-Peer Search Methods," presented at the WebDB, 2003.

[12] A. Clematis, et al., "A Distributed Approach for Structured Resource Discovery on Grid," in Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on, 2008, pp. 117-125.

[13] C.Mastroianni, et al., "A super-peer model for resource discovery services in large-scale Grids," Future Gener. Comput. Syst., vol. 21, pp. 1235-1248, 2005.

[14] B. Beverly Yang and H. Garcia-Molina, "Designing a super-peer network," in Data Engineering, 2003. Proceedings. 19th International Conference on, 2003, pp. 49-60.

[15] D. A. Menasc and L. Kanchanapalli, "Probabilistic scalable P2P resource location services," SIGMETRICS Perform. Eval. Rev., vol. 30, pp. 48-58, 2002.

[16] I. Foster, "The anatomy of the grid: enabling scalable virtual organizations," in Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on, 2001, pp. 6-7.

[17] K. Czajkowski, et al., "Grid information services for distributed resource sharing," in High Performance Distributed Computing, 2001. Proceedings. 10th IEEE International Symposium on, 2001, pp. 181-194.

[18] I. a. C. K. Foster, "Globus: A metacomputing infrastructure toolkit," International Journal of Supercomputer Applications,

pp. 115-128, 1997.

[19] D. a. J. S. Kaur, "Resource Discovery in Web-Services Based Grids," in Proceedings of World Academy of Science, Engineering and Technology, 2007, pp. 284-288.

[20] M. Cai and K. Hwang, "Distributed Aggregation Algorithms with Load-Balancing for Scalable Grid Resource Monitoring," in Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International, 2007, pp. 1-10.

[21] R. v. Renesse, "The importance of aggregation," in Future directions in distributed computing, Andr, et al., Eds., ed: Springer-Verlag, 2003, pp. 87-92.

[22] Gnutella website. Available: http://gnutella.wego.com/

[23] B. Y. Zhao, et al., "Tapestry: a fault-tolerant wide-area application infrastructure," SIGCOMM Comput. Commun. Rev., vol. 32, pp. 81-81, 2002.

[24] A. I. T. Rowstron and P. Druschel, "Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems," presented at the Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg, 2001.

[25] I. Stoica, et al., "Chord: A scalable peer-to-peer lookup service for internet applications," SIGCOMM Comput. Commun. Rev., vol. 31, pp. 149-160, 2001.

[26] S. Ratnasamy, et al., "A scalable content-addressable network," SIGCOMM Comput. Commun. Rev., vol. 31, pp. 161-172, 2001.

[27] A. S. Cheema, et al., "Peer-to-peer discovery of computational resources for Grid applications," in Grid Computing, 2005. The 6th IEEE/ACM International Workshop on, 2005, p. 7 pp.

[28] V. M. Y. M. Teo, and X. Wang, "A DHT-Based Grid Resource Indexing and Discovery Scheme," presented at the Singapore-MIT Alliance Annual Symp., 2005.

[29] D. Oppenheimer, et al., "Design and implementation tradeoffs for wide-area resource discovery," in High Performance Distributed Computing, 2005. HPDC-14. Proceedings. 14th IEEE International Symposium on, 2005, pp. 113-124.

[30] J. Li and V. Son, "A scalable semantic routing architecture for grid resource discovery," in Parallel and Distributed Systems, 2005. Proceedings. 11th International Conference on, 2005, pp. 29-35 Vol. 1.

[31] W. Xuan and K. Ling-Fu, "Resource Clustering Based Decentralized Resource Discovery Scheme in Computing Grid," in Machine Learning and Cybernetics, 2007 International Conference on, 2007, pp. 3859-3863.

[32] H. Shen, "SEMM: Scalable and Efficient Multi-Resource Management in Grids," in Proc. of the 2008 International Conference on Grid Computing and Applications (GCA), 2008.

[33] S. Haiying, et al., "LORM: Supporting low-overhead P2P-based range-query and multi-attribute resource management in grids," in Parallel and Distributed Systems, 2007 International Conference on, 2007, pp. 1-8.

[34] M. L. V. Sahota, M.A. Baker, and N. Antonopoulos, "A grouped P2P network for scalable grid information services," presented at the Peer-to-Peer Networking and Applications, 2009.

[35] D. Puppin, Moncelli, S., Baraglia, R., Tonelotto, N. and Silvestri, F., "A grid information service based on peer-to-peer," presented at the EuroPar, Springer LNCS, 2005.

[36] W. Li, et al., "Grid Resource Discovery Based on a Routing-Transferring Model," presented at the Proceedings of the Third International Workshop on Grid Computing, 2002.

[37] M. Marzolla, et al., "Resource Discovery in a Dynamic Grid Environment," in Database and Expert Systems Applications, 2005. Proceedings. Sixteenth International Workshop on, 2005, pp. 356-360.

[38] M. MollaMotalebi, et al., "A New Model for Resource Discovery in Grid Environment, Informatics Engineering and Information Science." vol. 253, A. Abd Manaf, et al., Eds., ed: Springer Berlin Heidelberg, 2011, pp. 72-81.

[39] V. H. G. Molt, J. M. Alonso, "A service-oriented WSRF-based architecture for metascheduling on computational Grids," Future Gener. Comput. Syst., vol. 24, pp. 317-328, 2008.

[40] T. Gomes Ramos and A. C. Magalhaes Alves de Melo, "An Extensible Resource Discovery Mechanism for Grid Computing Environments," in Cluster Computing and the Grid, 2006. CCGRID 06. Sixth IEEE International Symposium on, 2006, pp. 115-122.

[41] S. M. R. Beheshti and M. S. Moshkenani, "Development of Grid resource discovery service based on semantic information," presented at the Proceedings of the 2007 spring simulaiton multiconference - Volume 1, Norfolk, Virginia, 2007.

[42] R. Ranjan, et al., "Decentralised Resource Discovery Service for Large Scale Federated Grids," presented at the Proceedings of the Third IEEE International Conference on e-Science and Grid Computing, 2007.

[43] C. Mastroianni, et al., "Designing an information system for Grids: Comparing hierarchical, decentralized P2P and super-peer models," Parallel Comput., vol. 34, pp. 593-611, 2008.

[44] M. Hauswirth and R. Schmidt, "An Overlay Network for Resource Discovery in Grids," in Database and Expert Systems Applications, 2005. Proceedings. Sixteenth International Workshop on, 2005, pp. 343-348.

[45] Y. Murata, et al., "A distributed and cooperative load balancing mechanism for large-scale P2P systems," in Applications and the Internet Workshops, 2006. SAINT Workshops 2006. International Symposium on, 2006, pp. 4 pp.-129.

[46] M. Mitzenmacher, "The Power of Two Choices in Randomized Load Balancing," IEEE Trans. Parallel Distrib. Syst., vol. 12, pp. 1094-1104, 2001.

[47] Y. Drougas, et al., "Load balancing techniques for distributed stream processing applications in overlay environments," in Object and Component-Oriented Real-Time Distributed Computing, 2006. ISORC 2006. Ninth IEEE International Symposium on, 2006, p. 8 pp.

[48] J. Bustos-Jiménez, et al., "Load Balancing: Toward the Infinite Network and Beyond Job Scheduling Strategies for Parallel Processing." vol. 4376, E. Frachtenberg and U. Schwiegelshohn, Eds., ed: Springer Berlin / Heidelberg, 2007, pp. 176-191.

[49] A. Y. S. B. Yagoubi, "Task Load Balancing Strategy for Grid Computing," Journal of Computer Science, vol. 3, pp. 186-194, 2007.