TOWARDS A SCALABLE SCIENTIFIC DATA GRID MODEL AND SERVICES

AZIZOL ABDULLAH[1], MOHAMED OTHMAN[1], MD NASIR SULAIMAN[1], HAMIDAH IBRAHIM[1], ABU TALIB OTHMAN[2]

[1]Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

[2]Universiti Kuala Lumpur
50250 Kuala Lumpur, Malaysia

E-mails: {azizol, mothman, nasir, hamidah}@fsktm.upm.edu.my
abutalib@unikl.edu.my

ABSTRACT: Scientific Data Grid mostly deals with large computational problems. It provides geographically distributed resources for large-scale data-intensive applications that generate large scientific data sets. This required the scientist in modern scientific computing communities involved in managing massive amounts of a very large data collections that are geographically distributed. Research in the area of grid has given various ideas and solutions to address these requirements. However, nowadays the number of participants (scientists and institutions) that are involved in this kind of environment is increasing tremendously. This situation has lead to a problem of scalability. In order to overcome this problem we need a data grid model that can scale well with the increasing number of users. Peer-to-peer (P2P) is one of the architectures that is a promising scale and dynamism environment. In this paper, we present a P2P model for Scientific Data Grid that utilizes the P2P services to address the scalability problem. By using this model, we study and propose various decentralized discovery strategies that intend to address the problem of scalability. We also investigate the impact of data replication that addresses the data distribution and reliability problem for our Scientific Data Grid model on the propose discovery strategies. For the purpose of this study, we have developed and used our own data grid simulation written using PARSEC. We illustrate our P2P Scientific Data Grid model and our data grid simulation used in this study. We then analyze the performance of the discovery strategies with and without the existence of replication strategies relative to their success rates, bandwidth consumption and average number of hop.

KEYWORDS: *Scientific Data Grid, Peer-to-peer, Data Grid Model, Data Discovery*

1. INTRODUCTION

In modern scientific communities, the number of researchers involved in managing massive amounts of very large data collections through a geographically distributed environment is increasing. They need an infrastructure that can provide services to support their requirements to create a high-performance computing environment. Recently, research in the area of grid has given various ideas and solutions to address these requirements. However, as the numbers of participants (scientists and institutions) that involve in this collaboration environment keep increasing with the awareness of the

benefit from grid implementation, we are now facing with a new problem on how to provide a scalable infrastructure with scalable services for this environment. It is very difficult task to provide a scalable infrastructure with scalable services that can support high-performance computing environment. As discussed by Venugopal *et al.* [27], currently there are four different data grid model that have being manifested. However, these models are not scale when the number of user increased since they are based on centralized architecture. Over the last few years, the peer-to-peer (P2P) network has become a major research topic. The success of P2P networks was originally boosted by some very popular file sharing applications (e.g., Napster). It seems that The P2P network has provides a scalable file sharing environment. We argued that, the issues of scalability in grid might also be solved by using P2P model and P2P services since these two: the grid and P2P computing represents the same notion of sharing resources available at the edge of the Internet but with different purpose and target users.

A data grid system connects a collection of geographically distributed computer and storage resources that may be located in different parts of a country or even in different countries, and enables user to share data and other resources [1]. Recently, there are a few research projects directed towards the development of data grid such as Particle Physics Data Grid (PPDG) [2], Grid Physics Network (GriPhyN) [3], The China Clipper Project [4], and Storage Request Broker (SRB) [5] which are aim to build scientific data grid that enable scientists sitting at various universities and research labs to collaborate with one another and share data sets and computational power. The size of the data that needs to be accessed on those data grid is on the order of Terabytes today and is soon expected to reach Petabytes. As an example, The Large Hadron Collider (LHC) experiment [12] at CERN is producing Terabytes of raw data a year, and by 2008, this experiment will produce Petabytes of raw data a year that needs to be pre-processed, stored, and analyzed by teams comprising 1000s of physicists around the world. Through this process, more derived data will be produced and 100s of millions of files will need to be managed and stored at more than 100s of participating institutions. Off course, the number of participations will increase and there is a need to provide a more scalable environment that could support this requirement.

In scientific environments, even the data sets are created once and then remain reads only, but the data sets usage often leads to creation of new files, inserting a new dimension of dynamism into a system [13, 14]. Ensuring efficient access to such huge and widely distributed data is a serious challenge and critical to network and data grid designer. The major barrier to supporting fast data access in a data grid system are high latencies of Wide Area Network (WANs) and the Internet, which impact the scalability and fault tolerance of total data grid systems. There are a number of research groups investigating a number of data replication approaches on the data distribution for data grid, in order to improve its ability to access data efficiently [17-20]. In this study, we used a different approach in addressing the scalability problem by proposing a Scientific Data Grid model based on Peer-to-Peer (P2P) and this model used P2P services. Using this model we then study various decentralized discovery strategies which the heart of this model. In this paper, we also investigate the data replication approaches for data distribution to address the data distribution and reliability problem for our Scientific Data Grid Model and study the impact of these approaches on the propose discovery strategies. To evaluate the model and selected discovery strategies, we have developed our own data grid simulation using

PARSEC. The simulation used to generate different network topologies to study the impact of discovery strategies on the cost of locating data and the data access on the overall Scientific Data Grid.

The paper is organized as follows. Section 2 gives an overview of previous work on discovery strategies. In Section 3, we describe our Scientific Data Grid model. Section 4 describes the simulation study and our simulation results. Finally, we present a brief conclusion and future directions in Section 5.

2. RELATED WORKS

In this section, we discussed more on discovery since discovery is a core grid functionality that enhances the accessibility of data in our proposed model. Support for handling and discovering resource capabilities already exist in some metacomputing system such as Globus [21] and Legion [22]. In Globus, the Globus Resource Allocation Manager (GRAM) acts as a Resource Broker is responsible for resource discovery within each administration domain, which works with an Information Service, and a Co-allocator for monitoring current state of resources, and managing an ensemble of resources respectively. In Legion system, resources and tasks are describe as a collection of interacting objects. A set of core object, enable arbitrary naming of resources based on Legion Object Identitifers (LOIDs) and Legion Object Addresses (LOA). Specialized services such as Binding agents and Context objects are provided to translate between arbitrary resource name and its physical location – enabling the resource discovery to be abstracted as a translation between LOIDs and physical resource location. In CONDOR [23], resource describes their capabilities as an advertisement, which is subsequently matched with an advertisement describing the needs of an application. Most of these strategies based on a centralized approach, which may not be as scale. In a large and dynamic environment such as grid, the discovery service should be decentralized in order to avoid potential computation bottlenecks at a point and be more scale.

In the last couple of years, Peer-to-Peer (P2P) system became fashionable [6-9, 11]. They emphasize two specific attributes of resources sharing communities: scale and dynamism. Most of existing P2P systems such as Gnutella, Freenet, CAN [15] and Chord [16] provide their own discovery mechanisms and focus on specific data sharing environment and, therefore on specific requirements. As example, Gnutella emphasis on easy sharing and fast file retrieval, but with no guarantees that files will always be located. In Freenet, the emphasis is on ensuring anonymity. CAN and Chord guarantee that files are always located but accepting increased overhead for file insertion and removal. In scientific collaboration, the typical uses of shared data in scientific collaboration have particular characteristics: *Group locality* – users tend to work in a group and use the same set of resources (data sets). The newly produced data sets will be of interest to all users in the group. *Time locality* – the same users may request the same data multiple times within short time intervals [10]. However, the discovery mechanisms that are proposed in existing P2P systems such as Gnutella, CAN, Chord, etc., do not attempt to exploit this behavior.

A large amount of work on P2P discovery mechanisms has been done either in structured and unstructured P2P system [15, 16, 24-26]. In this paper, we focus our work on unstructured P2P Scientific Data Grid system that will answer one of the questions that have been raised up by the authors in [10]. The question that the most interest to us is how do we translate the dynamics of scientific collaborations in self-configuring network protocols such as joining the network, finding the right group of interests, adapting to changes in user's interest, etc.? This is relevant and challenging in context of self-configuring P2P networks. The services that are provided by unstructured P2P system could be the answer for this question, where it does not rely on a specific network topology and supports the discovery in arbitrary forms. The unstructured P2P provides more scale and dynamism for Scientific Data Grid environment. We have proposed and developed an unstructured P2P Scientific Data Grid model which is discussed in the next section.

3. THE P2P DATA GRID MODEL

A variety of models are in place for the operation of a data grid. These data grid models are manifest in different ways in different systems. The model is the manner in which data sources are organized in a system. These are dependent on the source of data, whether single or distributed, the size of data and the mode of sharing [27]. Four of the common models are *monadic*, *hierarchical*, *federation* and *hybrid*. In this section, we only illustrate our proposed model which provides a generic infrastructure to deploy P2P services and applications.

Figure 1 shows our proposed model for a Scientific Data Grid that will support scientific collaboration. This model is specified as unstructured P2P model where, peers could be any network devices and in our implementation, the peers can include PCs, servers and even supercomputers. Each peer operates independently and asynchronously from all other peers and it can be self-organized into a peer group. Peer group contains peers that have agreed upon a common set of services, and through this peer group, peer can discover each other on the network. Once a peer joins a group it uses all the services provided by the group. Peers can join or leave the group at anytime that they want. In this model, once a peer joins the group, all the data sets that are shared by other peers in the group will be available to him/her. Peers can also share any of their own data sets with other peers within the group. A peer may belong to more than one group simultaneously.

In this scientific communities data-sharing model, we try to provide the same concept as the analogy of electrical power grid, where the users or scientists (in our case is peers) can access their required data sets without knowing which peers deliver the data sets. In other words, they can execute their applications, obtain the remote data sets (did not concerned with the provider) and then wait for the results. This will be done by the discovery service. Our focus in this research is to propose a decentralized discovery strategy for Scientific Data Grid that addresses the scalability problem and also reliability problem.
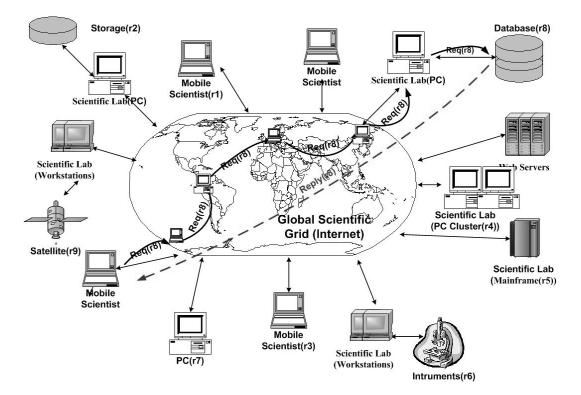
Fig. 1: P2P Model for scientific data grid.

4. THE SIMULATION STUDY

The presented results were obtained with simulation of only read requests without any background traffic in network and the stream of requests for the Scientific Data Grid. Our simulation is written in PARSEC simulation language. PARSEC is a C-based discrete-event simulation. The simulator consists of three parts. The first part is the entity that is responsible for creating the rest of entities: the entities that simulating the network nodes and network layer. It also reads all the inputs needed for the simulation. The second part is the network layer, which comprises of two entities: The entity that simulates a network forwarding protocol and the entity that simulates the Distance Vector Multicast Routing Protocol. The third part is the entity that simulates the various nodes in the Scientific Data Grid.

The Scientific Data Grid Simulation is developed to identify suitable resource discovery strategies and mechanism for P2P Scientific Data Grid model. Starting a simulation first involves specifying the topology of the grid, including the number of network nodes, how they are connected to each other and the location of the files across various nodes. The bandwidth of each link is not specified in this simulation model in order to simplify the model. We assumed that all links have the same bandwidth. In this simulation model, various file requests will be triggered originating according to file access patterns to be simulated. In this study, a file access pattern is generated randomly. The request will be forwarded to the selected neighbour nodes using the information in the routing table in each nodes. Before it forwards the request to several neighbor nodes, it

will check the file required and submit the request if the node has the requested file. In this case, the simulator will record the hop count for the successful request. If not, it will forward the request to other nodes until its time-to-live (TTL) for the request expired.

There are various proposed strategies and mechanism to discover or locate the resources in distributed system. We have studied and compared different discovery strategies using this simulation. Through this simulation, various parameters such as hop count, success rates, response time and bandwidth consumption will be measured. However, in this paper, we just show only the hop count measurement. The simulation was first run on the access patterns that were generated randomly. Table 1 is a sample access pattern file randomly generated during simulation. This being the worse case scenario, more realistic access pattern that contained varying amounts of temporal and geographical locality will be generated in future study.

Table 1: A sample access pattern file.

Time	File Name	Requestor Node
20	File1	1
39	File2	5
70	File2	6
75	File3	3
90	File1	2
92	File2	3
98	File3	5

5. EXPERIMENTS AND RESULTS

In a Scientific Data Grid system, several factors influence the performance of discovery mechanism; the network topology, resource frequency and resource location. However, in these studies, we will only focus on the last two factors. In this section, we show the simulation results of the discovery strategies performance with and without the existence of replication strategies. We then show the results on the impact of data replication strategies that we proposed to address the data distribution and reliability problem for our Scientific Data Grid model, on the propose discovery strategies.

Currently, in our simulation, we used generated random flat networks topology consisting of 20 - 220 nodes. We placed the data file ourselves randomly, and measured data access cost in average number of hop, success number of request and bandwidth consumption. However, in this paper we only show a few of the experiment results. In this simulation study, we assumed that the overlay network graph topology does not change

during the simulation as to maintain simplicity. In this study, four discovery schemes have been model and studied. The study was focused on the discovery and also replication aspects that would improve the performance of data discovery strategy.

In a flooding-based protocol, nodes flood an overlay network with queries to discover a data file. In a request forwarding technique, nodes will forward the queries to the selected neighbor by utilizing information range from simple forwarding hints to exact object locations. The discovery strategies that we have modeled in this simulation included Flooding(F), Random Neighbour(RN), Learning Flooding(LearningF) and Learning Random Neighbour (LearningRN). We have run a number of simulations by creating two different scenarios. In the first scenario, we assumed that no replicas are created at the beginning and during the simulation. In the second scenario, we created replicas at the beginning of simulation. This scenario is to study the impact of replication on the discovery strategies. Figure 2 shows a graph representing the data access cost in average number of hop for topologies with different number of nodes without replication for four different discovery strategies.

As shown by Fig. 2, the performance of all four strategy is worst when the number of nodes increases in the collaboration network. We expect that in real Scientific Data Grid Environment, it will become worst as this strategy does not scale well when the number of nodes increases. The strategy based on flooding technique might give optimal results in a network with a small to average number of nodes. Fig. 2 also shows that the performance of random neighbour discovery strategy is worst compared to other discovery strategies where the number of hops reaching nearly 200 hops to reach to an object in a network topology that consist 160 nodes.
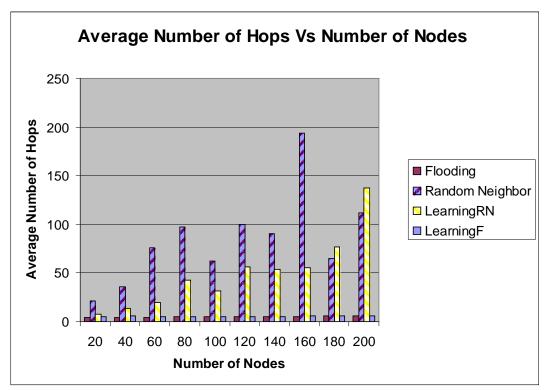

Fig. 2: Average number of hop in different number of nodes.

Figure 3 shows a graph representing the data access cost in term of the number of success queries for topologies with a different number of nodes with and without replication using flooding discovery strategy. The results show that the number of success queries increase when the data is replicated in the collaboration network. The number of replica that created in the network can gives a better performance to discovery strategies as shown by Fig. 3 and Fig. 4.
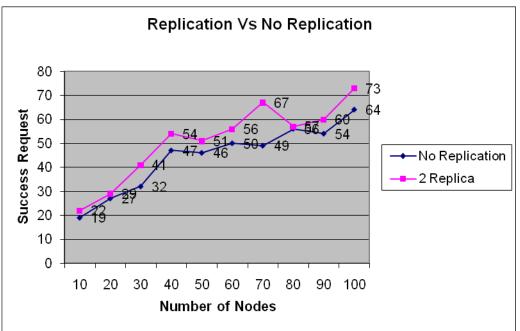

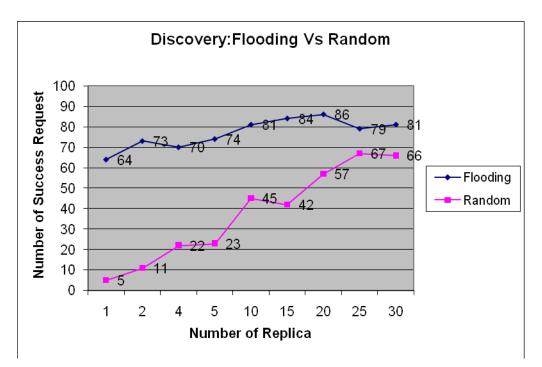Fig. 3: Number of success queries for Flooding Mechanism.

Fig. 4: The 100 nodes network topology with various numbers of replicas.

The results from Fig. 4 show that when the number of replica increase, the number of success queries also increase. For discovery strategy that is based on random, it shows that this strategy can perform better when more data is replicated in the collaboration network. However, we need to monitor and maintain the optimum number of replica that should be replicated in the network in order to prevent other problems such as the storage usage problem etc.
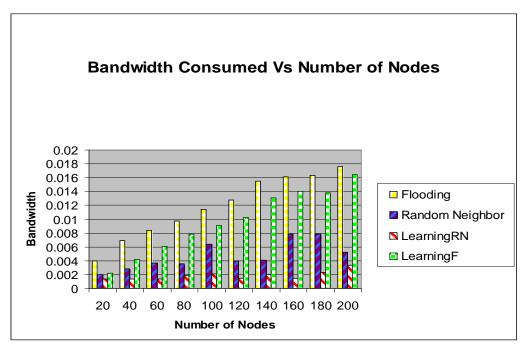

Fig. 5: Bandwidth consumption for four different Discovery Mechanisms.

Figure 5 shows a graph representing the bandwidth consumption for different topologies with different number of nodes without replication using four different discovery strategies. The results show that both flooding-based discovery strategies used more bandwidth to flood the queries. The bandwidth consumption increase accordingly when the number of node increase. However, by having a little bit of checking process for Learning Flooding strategy, it shows that we have slightly reduced the bandwidth consumed by the strategy. It is also the same with the Learning Random strategy.

6. CONCLUSION AND FUTURE WORK

We have addressed the problem of scalability in discovery process for Scientific Data Grid by proposing decentralized discovery strategies (Learning Flooding and Learning Random Neighbour). We also have addressed the issue of having a scalable Scientific Data Grid model by proposing P2P Scientific Data Grid model. We have described our P2P model for Scientific Data Grid Environment and our simulation study. Although the model is conceptually simple, it has provided a scalable service that could be remarkably useful for a Scientific Data Grid Environment that is based on P2P. We believe that with a very simple simulation experiment that we have done on our model using our Scientific Data Grid simulation, it shows that this approach can be used to support good services for

data discovery in a large-scale unstructured scientific collaboration environment. It also represents an interesting approach to data management in Scientific Data Grid through our study on replication impact to discovery strategy.

Our results show that, for a Scientific Data Grid Environment with large population of users, we need a discovery mechanism that could give the shortest time to discover and locate data resources, it also provide the guarantees required by collaborators and, with low bandwidth consumption, and also scalable. In our future work, we will study other strategies to discover or locate the resources in decentralized environment. Various replication strategies will be incorporated as to support more dynamic Scientific Data Grid scenario and also to address the reliability problem.

REFERENCES

[1] Chervenak, A., Foster, I., Kesselman, C., Salisbury, C., Tuecke, S. The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Data Sets. Journal of Network and Computer Applications. (2000).

[2] PPDG: Particle Physics Data Grid, http://ppdg.net

[3] GriPhyn : Grid Physics Network, http://www.griphyn.org

[4] Johnston, W., Lee, J., Tierney, B., Tull, C.,Millsom, D. The China Clipper Project: A Data Intensive Grid Support for Dynamically Configured, Adaptive, Distributed, High-Performance Data and Computing Environments. In Proceeding of Computing in High Energy Physics 1998, Chicago (1998).

[5] SRB: Storage Request Broker, http://www.npaci.edu/DICE/SRB

[6] Gnutella website, http://www.gnutella.wego.com

[7] The Free Network Project website, http://freenetproject.org

[8] Napster website, http://www.napster.com

[9] Konspire website, http://konspire.sourceforge.net

[10] Iamnitchi, A., Ripeanu, M., Foster, I. Locating Data in (Small-World?) Peer-to-Peer Scientific Collaborations. 1st International Workshop on Peer-to-Peer Systems, Cambridge, Massachusetts, March (2002).

[11] Gong, L. Project JXTA: A Technology Overview. Sun Microsystems, Inc. April 25, (2001).

[12] LHC- The Large Hadron Collider website,

http://lhc-new-homepage.web.cern.ch/lhc-new-homepage/

[13] Loebel-Carpenter, L., Lueking, L., Moore, C., Pordes, R., Trumbo, J., Veseli, S., Terekhov, I., Vranicar, M., White, S. and White, V. SAM and the Particle Physics Data Grid. In Proceedings of Computing in High-Energy and Nuclear Physics. Beijing, China. (2001)

[14] Lueking, L., Loebel-Carpenter, L., Merritt, W., Moore, C. The DO Experiment Data Grid – SAM. C.A. Lee(Ed.): GRID 2001, Lecture Notes in Computer Science, Vol 2242. Springer-Verlag Berlin Heidelberg (2001). 177-184

[15] Ratnasamy, S., Francis, P., Handley, M., Karp, R., Shenker, S. A Scalable Content-Addressable Network. In SIGCOMM 2001, San Diego, USA.(2001).

[16] Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., Balakrishnan, H. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. In SIGCOMM 2001, San Diego, USA. (2001).

[17] Ranganathan, K. and Foster, I. Identifying Dynamic Replication Strategies For a High Performance Data Grid. In Proceedings of the International Grid Computing Workshop. Denver. (2001)

[18] Ranganathan, K. and Foster, I. Design and Evaluation of Replication Strategies For a High Performance Data Grid. . In Proceedings of Computing in High-Energy and Nuclear Physics. Beijing, China. (2001)

[19] Stockinger, H., Samar, A., Allcock, B., Foster, I., Holtman, K., Tierney, B. File and Object Replication in Data Grids. In Proceedings of the Tenth International Symposium on High Performance Distributed Computing. IEEE Press. (2001).

[20] Vazhkudai, S., Tuecke, S., Foster, I. Replication Selection in the Globus Data Grid. In Proceedings of the First IEEE/ACM International Conference on Cluster Computing and the Grid. Pp. 106-113, IEEE Computer Society Press. (2001).

[21] Foster, I., Kesselman, C. Globus: A Metacomputing Infrastructure Toolkit. *Intl J. Supercomputer Applications*, 11(2) (1997)115-128

[22] Grimshaw, A., Lewis, M., Ferrari, A., Karpovich, J. Architectural Support for Extensibility and Autonomy in Wide-Area Distributed ObjectSystems. In Proceedings of the 2000 Network and Distributed System Security Symposium (NDSS2000). (2000)

[23] Frey, J., Tannenbaum, T., Foster, I., Livny, M., Tuecke, S. Condor-G: A Computation Management Agent for Multi-Institutional Grids. In Proceedings of the Tenth IEEE Symposium on High Performance Distributed Computing (HPDC10) San Francisco, California, August 7-9. (2001).

[24] Gkantsidis, C., Mihail, M., Saberi, A. Hybrid Search Schemes for Unstructured Peer-to-peer Networks. Proc. Of IEEE INFOCOM'05, Miami, USA (2005).

[25] Lv, Q., Cao, P., Cohen, E., Li, K., Shenker, S.: Search and Replicating in Unstructured Peer-to-peer Networks. Proc. Of 16th Annual ACM Int. Conf on Supercomputing (ISC'02), New York, USA (2002).

[26] Crespo, A., Garcia-Molina, H.: Routing Indices for Peer-to-peer Systems. Proc. Of Int. Conf. On Distributed Computing Systems(ICDCS'02), Vienna, Austria (2002).

[27] Venugopal, S., Buyya, R., Ramamohanarao, K., A Taxonomy of Data Grids for Distributed Data Sharing, Management and Processing. in: ACM Computing Surveys, vol. 38(1), pp. 1–53. ACM Press, New York, USA, 2006.