Dynamic Updating and Management of Virtual Resource Database in Grids Using Mobile Agents

Kameshwari S, Valliyammai C, ThamaraiSelvi S, Divya S, Sharmila R

Anna University, India

Summary
Grid Computing Element (CE) is a distributed heterogeneous system of computational data and informational resources. The goal of any Grid CE is to create an integrated virtual system of resources available to the users. The nature of resource utilization is unpredictable in such computational resource rich Grid CE. So continuous and dynamic updating of the utilization of the resources in the grid clusters becomes mandatory. Virtual database decreases the information/query processing, which is a bottleneck in Centralized Monitoring mechanism. In this paper, the authors propose a Mobile-Agent based mechanism for the dynamic resource monitoring in the Grid CE. The process involves the creation and maintenance of a Virtual Resource Database for every Resource Consumer group.

Key words:
Grid, Dynamic resource monitoring, Virtual Database, Virtual Organization, Monitoring Mobile Agents

1. Introduction

Grid environments have recently emerged as integrating infrastructure for distributed high performance scientific applications. This involves the combined effective utilization of the Grid resources to achieve high performance. But locating the resourceful sites in the grid needs continuous monitoring. The more famous Grid Middleware, Globus Toolkit 4 also does this major task or resource monitoring in grids. It uses LDAP (Lightweight Directory Access Protocol) service to access the network metrics according to the defined LDAP schema. The LDAP service gathers and maintains the metric data via scripts that fetches the current measurement from the local resource monitoring data store. Various tools can be used to collect the grid resource monitoring data from the LDAP service. But the LDAP service enables only simple query processing. The relational implementation would enhance the complex query processing, like in R-GMA (Relational Grid Monitoring Architecture). This concept is the inception for our idea of maintaining a virtual relational database to enable complex querying and lesser information/query processing.

In this scenario, the authors think that mobile agent technology can play a central role since their capability to cope with systems' heterogeneity and to deploy user customized procedures on remote sites seems to be adequate to Grid Environment. By interacting with the remote host, after migrating on it, an agent is able to make complex operations on remote data without transferring them. The basic idea under this paradigm is that of moving the application logic near the data it need. This may produce a significant saving of bandwidth on one hand, and the possibility to analyze remote performance data with user customized algorithms encapsulated in an agent, on the other hand.

2. Related Work

Resource Monitoring in Grid has seen vast research and many architectures have been proposed as in [1, 2, 3 and 4]. Mobile agents play a major role in resource monitoring [11]. Moving a step ahead, mobile agents have also been used in grids for the purpose of resource monitoring [7 and 12]. The concept of Virtual Database has been taken from [16, 17 and 18]. The paper discusses a novel two phase mechanism, for Resource Monitoring in grids, the first phase being taken care of by a Mobile Agent Tool and the second phase being taken care of by a Grid Middleware like GT.

3. Proposed Agent Based Model

The proposed model, with two clusters is shown in Figure 1. Every cluster contains a head node, also called the Resource Provider, does the monitoring activities. That is, the head node runs the Grasshopper tool and migrates monitoring mobile agents to other nodes in the Grid CE, within the cluster. The agent code is written is such a way that it accesses the dynamic processes' files in the nodes and gets them back to the head node. The head node thus aggregates all the resource information of the Grid CE. The head node does this periodically, such that the database contains, almost the current resource values of all the nodes in the Grid CE. Note that the cluster

Manuscript received May 5, 2007
Manuscript revised May 20, 2007

nodes do not know the identity or the location of the Registry Database. The head node alone knows the identity and the location of the Registry Database. The significance of this fact will become evident in the Virtual Database section of this paper.

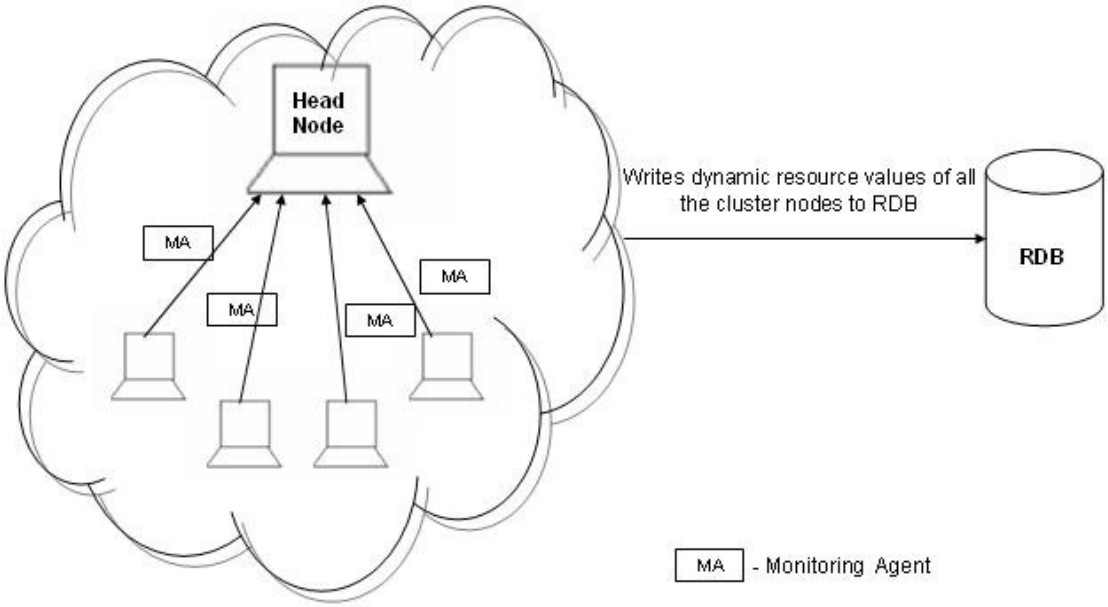

Fig 2. Proposed Agent based model

As head nodes are the actual Resource Providers, job submission takes place only at the RPs. Although the head nodes are called the Resource Providers, they do not actually do the submitted job. They instead, distribute the job among the member nodes of the cluster. The resource values of the member nodes are also monitored, because the head node needs to know the resource values of the member nodes, in order to distribute the submitted job in the cluster.

4. Virtual and Registry Database Design

The location of the Virtual and the Registry Database are shown in Fig 2. Certain kind of requirements can be satisfied only by certain Resource Providers (RP). VDB is one which contains all the Resource Providers' URL, who can satisfy the resource requirements of a particular Resource Consumer group. This Virtual Database is created and maintained at every cluster. Registry Database, RDB is the actual database which contains the resource metrics of all the Resource Providers, aggregated by the Head node of each cluster.

The head node updates the aggregated information, onto the Registry Database (RDB). The Virtual Database is the key in reducing processing time involved in querying. Registry and Virtual Database updating takes place in two phases.

1) Dynamic updating of the Registry Database with the resource metrics of the cluster nodes, at regular intervals.
2) Creating and managing a Virtual Database.

The data structure of the Virtual Database is a relational table consisting of the identity and the location (URL) of the Resource Providers, who can satisfy the resource needs of the Resource Consumers. For example, CpuVDB contains all the resource providers' URL, rich in cpu speed.

.

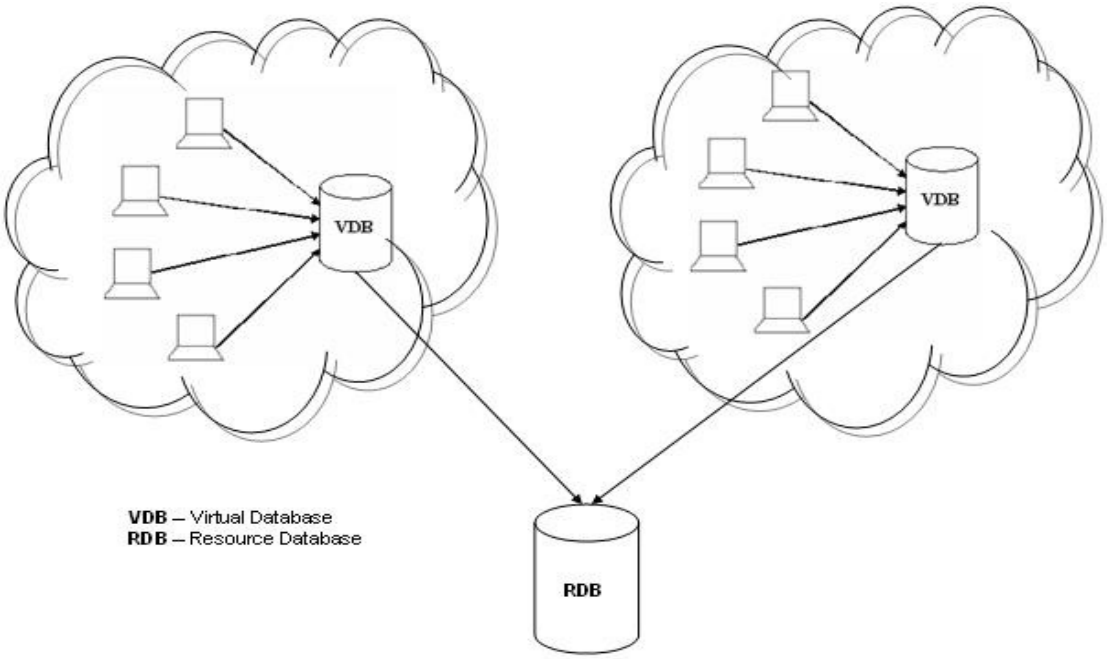

Fig.2 Dynamic updating of virtual and Resource Database

Whenever a consumer asks for resource requirements, the Resource Providers satisfy the requirements of the consumer, from the VDB. The Registry Database (RDB) is then queried with only a subset of Resource Providers that is collected from the VDB, instead of traversing all the records in the RDB. Thus the processing complexity in querying the database gets reduced, which is proved mathematically in the next section of the paper.

5. Mathematical Evaluation of Processing Complexity

Let
N = total number of Resource Providers in the VO, available in RDB.
Let
i = be the number of Resource Consumer groups, who can avail the resources of the Grid CE to satisfy their resource requirements.
Let
N_i = be the number of Resource providers in VDB of each cluster,

such that

$N >= N_1 + N_2 + N_3 + ... + N_i$ …Eq 1

and

$\boldsymbol{N_i << N}$ …Eq 2

In normal case, to get the resourceful node to submit the job,

Number of comparisons required = $\boldsymbol{N}$

That is, it requires that all the records of the RDB to be traversed once, before selecting the Resource Provider. While in the case of having a VDB in between the consumer and the RDB,

Number of comparisons required = $\boldsymbol{Ni}$ ....Eq 2

Active Operation Complexity (AOC)

Without VDB = $\boldsymbol{N}$ …Eq 3

With VDB = $\boldsymbol{N_i / N}$ …Eq 4

Reduction Factor}
In AOC} = $\boldsymbol{1 - (N_i / N)}$ …Eq 5

= $\boldsymbol{(N-N_i) / N}$ ....Eq 6

From the Eq 6, it can be inferred that, as the margin between Ni and N increases, the processing (comparison) complexity is greatly improved, with the presence of VDB.

6. Implementation

The model was implemented at the Grid Computing Labs of MIT, Anna University, with a total of three clusters. The implementation was carried on in two phases. The first phase is taken care by a Mobile Agent Tool, Grasshopper and the second phase is taken care of by a Grid

Middleware like GT.

In the first phase, the Grasshopper tool is used to code and migrate the mobile agents. The Monitoring Mobile Agent is used to monitor the resources in the nodes of the cluster and Grid CE. The Monitoring Mobile Agent (MMA) was invoked at head node of every cluster. The MMA migrates to every member nodes of the Cluster, collects resource metric values of the member nodes and returns the result to the Registry node. The Registry node is the one which maintains the resource metrics of the RPs. The monitoring is done in periodic intervals, so as to dynamically update RDB with the new resource metric values. Migration is done by all the RPs that are registered within the Grid CE.

In the second phase of the implementation, based on the domain, the head node chooses certain RPs and stores their location (URL) along with certain keywords for the RP. Any consumer needing the resource, contacts the corresponding Virtual Database. The VDB contains only the URLs of the RPs, who can cater to the resource requirements of a particular domain. By using the First Fit Algorithm, the RP was selected and the URL of the RP is obtained by the consumer. Then the job was submitted using Globus Toolkit 4, at the selected URL of the RP by the consumer.

7. Results and Discussion

The model was also tested with and without VDB, with varied number of simultaneous requests. This is done to practically check the variance in the processing time of the requests, with direct and indirect access to RDB. The scalability up to 1000 clusters were simulated using GridSim and the results are shown graphically in Fig. 3

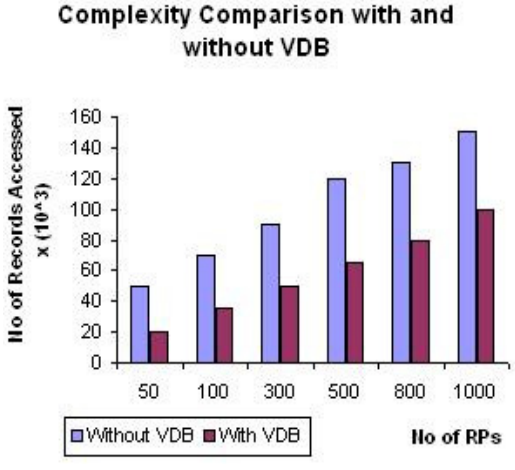

Fig no. 3 Complexity Comparison Graph

From Figure no.3 it is evident that, as the number of requests increases at a particular point of time, the processing time with the use of VDB decreases to a great extent, compared to the one without VDB.

8. Conclusion

Resource Monitoring in Grid Using Mobile Agent technology has already been proved efficient. Further, with the introduction of VDB in a busier grid will have a much more significant improvement in the processing time involved.

Acknowledgments

The authors sincerely thank the Grid Computing Labs of Anna University for their Infrastructure assistance, by allowing us to configure the clusters and test the system for its practicality. We also thank Mr. G Kannan of Grid Computing Labs for offering timely technical assistance.

References

[1] Mark Baker and Garry Smith, "*GridRM: An Extensible Resource Monitoring System"*, Proceedings of the IEEE International Conference on Cluster Computing, 2003

[2] Fawad Nazir, Arshad Ali, Hafiz Farooq Ahmad, Hiroki Suguri, Hamid Abbas burki and Tallat Hussain Tarar , "*MoGiNMA: A Negotiation Framework for Autonomous Grid Node Monitoring Agent*", Multitopic Conference, 2004. Proceedings of INMIC 2004, 8th International

[3] Ghazala shaheen, Muhammad Usman Malik, Zohair Ihsan, Nadeem Daudpota, *"Grid Visualizer: A Monitoring Tool for Grid Environment*", Proceedings of the 16th International Workshop on Database and Expert Systems Applications,2005

[4] David Abramson, Jon Giddy, LewKotler, *"High Performance Parametric Modeling with Nimrod/G: Killer Application for the Global Grid*" ,14th International Parallel and Distributed Processing Symposium,2000

[5] Karl Czajkowski, Carl Kesselman, Steven Fitzgerald, Ian Foster, *"Grid Information Services for Distributed Resource Sharing",* 10th IEEE International Symposium on High Performance Distributed Computing,2001

[6] Rajkumar Buyya, *"Economic-based Distributed Resource Management and Scheduling for Grid Computing",* Journal-ref: Monash University, Australia , 2002.

[7] Junwei Cao Spooner, D.P. Jarvis, S.A. Saini, S. Nudd, G.R, *"Agent-based grid load balancing using performance-driven task scheduling,*Parallel and Distributed Processing Symposium", 2003.

[8] Fukuda, M. Tanaka, Y. Suzuki, N. Bic, L.F. Kobayashi, S, "*A mobile-agent-based PC grid"*, AutonomicComputingWorkshop,2003

[9] Rajkumar Buyya,David Abramson, Jonathan Giddy,*"A Case for Economy Grid Architecture ",*for Service Oriented Grid Computing ,10th Heterogeneous Computing Workshop, 2001

[10] Rajkumar Buyya, Steve Chapin, David DiNucci, *"Architectural Models for Resource Management in the Grid*", First IEEE/ACM International Workshop, 2000.

[11] Ying-Hong Wang, Huan-Chao Keh, Tsang-Ching Hu, and Cheng-Hong Liao, *"A Hierarchial Dynamic Monitoring Mechanism for Mobile Agent Location",* Proceedings of the 19th International Conference on Advanced Information Networking and Applications, 2005

[12] Orazio Tomarchio, Lorenzo Vita, Antonio Puliafito, *"Active monitoring in Grid environments using mobile agent technology",* DATAGRID workshop, 2000

[13] Lee,J. Gunter,D.Stoufer,M.Tierney , *"Monitoring DataArchives for Grid Environments*", Supercomputing, ACM/IEEE Conference, 2002

[14] Xuechai Zhang Freschl, J.L. Schopf, J.M, *"A performance study of monitoring and information services for distributed systems", High Performance Distributed Computing,* Proceedings, 12th IEEE International Symposium, 2003.

[15] *Resource and Job Monitoring in the Grid* Euro-Par 2003 Parallel Processing, Grid Computing and Middleware Systems.

[16] A. W. Cooke , A. J. G. Gray , W. Nutt, J. Magowan, M. Oevers P. Taylor, Cordenonsi, R. Byrom, L. Cornwall, A. Djaoui, L. Field, S. M. Fisher , S. Hicks, J. Leake R. Middleton, A. Wilson, X. Zhu, N. Podhorszki, B. Coghlan, S. Kenny, D. O ' Callaghan and J. Ryan, "*The Relational Grid Monitoring Architecture: Mediating Information about the Grid"*, Journal of Grid Computing ,2005

[17] A Gupta, V Harinarayan, A Rajaraman Virtual database technology ACM SIGMOD Record Volume 26 , Issue 4 (December 1997) Pages: 57 – 61 ISSN:0163-5808

[18] Anand Rajaraman, Peter Norvig, "*Virtual Database Technology: Transforming the Internet into a Database,*" IEEE Internet Computing, vol. 02, no. 4, pp. 55-58, Jul/Aug, 1998.

S Kameshwari is pursuing the final year Bachelor of Technology degree in Information Technology, at Madras Institute of Technology, Anna University, Chennai, India. Her research interests include Grid/Cluster Computing, Network Security and Mobile Agents

C Valliyammai received her M.Tech. degree in computer science and engineering from Visveswaraiah Tech University. She is currently doing her PhD in computer science and engineering at Anna University. She has six years of teaching experience. Her area of interest includes Network Management, Grid computing and Mobile Agents.

S Thamarai Selvi received her Ph.D Computer Science and Engineering in Manonmaniam Sundaranar University, Tirunelveli. Her area of interest includes Artificial Neural Networks, and Grid computing. She is handling five research projects funded by Central government Agencies. She is currently working as a Professor and Head of the Dept of IT at MIT, Anna University, India.

S Divya is now a final year student aspiring for Bachelor of technology degree in Information technology, at Madras Institute of Technology, Anna University, Chennai, India. Her research interests include Grid Technology and Mobile Agents.

R Sharmila is pursuing the final year Bachelor of Technology degree in Information Technology, at Madras Institute of Technology, Anna University, Chennai, India .Her research interests include Grid Technology and DataBases.