International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-2 Issue-4, March 2013

Middleware and Toolkits in Grid Computing

Pooja Rani

Abstract:- The increasing demand for more computing power and data storage capacity in many fields of business, research, engineering, medical and science has raised the emergence of Grid Computing. Grid computing facilitates the environment where computers are interconnected with each other in such a manner that for making the execution faster of their tasks they can utilize the unused processing power of other idle systems. The applications of grid demand the secure access to the computational resources in the distributed environment. The secure and uniform access to resources is provided by Grid Middleware. In this paper, I have presented some of the middleware and Toolkits having some advantages and disadvantages.

Keywords: Grid Computing, Middleware, Toolkits, Grid Applications.

I. INTRODUCTION

A new field named "Grid Computing" is emerged from the conventional distributed computing. "Grid" name is analogous to electrical power grid. The basic motivation behind this is to access the power of computer from computer grid as like to access the power of electricity from electrical grid. Grid is an infrastructure in which many grids are joined together for the creation of a global network of computers, where discovery, selection, management, processing and sharing of large data sets become possible. Grid Computing is a form of distributed computing in which geographical distributed heterogeneous resources are integrated in to a single Virtual Organization, Where these resources can be shared, managed and exchanged to solve the large scale problems in research, science and commerce. Grid uses some open standards and protocols to provide the interoperability, to access and to discover the resources. To access the resources in uniform manner, Grid portal is an entrance for the end user. Grid portal provides the facilities to user like submission, monitoring and management of jobs. Grid middleware is designed to run the various applications over the entire network of computers.

II. REVIEW OF LITERATURE

A. Grid Middleware

The backbone of the grid computing systems is considered to the Grid Middleware. Because the communication across the entire network is not possible without the use of middleware. The purpose of grid middleware is to integrate the heterogeneous resources, efficiently assigning the resources to jobs, monitoring, managing and to provide the secure data access [14].In the past, while facing the grid infrastructure, many of the existing middleware shows the various limitations. To overcome these limitations, new middleware are proposed. We introduced here some of the most widely used middleware to meet the needs of complex applications.

Manuscript published on 30 March 2013.
***Correspondence Author(s)**
Pooja Rani Student, Department of CSE,Lovely Professional University

I. MPI Based Middleware

For the development of parallel applications, MPI (Message Passing Interface) is introduced as the main communication library. For grid, new approach MPICH-G2 is introduced that is the implementation of MPI. With the use of MPICH-G2, across the network of computers users can run Message Passing Interface programs with the help of the same command. It also uses the Globus toolkit services. As a result, it provides the better results than its predecessor. Its applications can be seen in world-wide like, to run and to distribute applications and conventional Message Passing Interface programs across computers located at different sites [7].

II. Java RMI Based middleware

To develop the large scale distributed applications, Java RMI (Remote Method Invocation) based middleware is introduced. For object replication, a new approach is introduced in java that is based on complier. The compiler is helpful for code generation for RMI and to check the consistency. This approach results in better java objects replication and also performance is improved for parallel object based programs [2].

III. NetSolve

In distributed environment, to solve the computational scientific problems, this client/server application is designed. This middleware provides the facility for searching the available resources, selecting the best one resource, and after solving a problem an answer is returned to the user [4] With the Use of TCP/IP sockets, the communication between NetSolve Agents, Clients and Servers is done. For searching and selecting the best one resource NetSolve agents are responsible [5].

IV. Globus based middleware

One of the most widely used middleware is Globus. The most popular Globus toolkit is also provided by Globus. For computational grid, it provides high level services. Various services provides by Globus includes GRAM (Globus Resource Allocation Manager) that is responsible for resource allocation, monitoring and management services. GSI (Grid Security Infrastructure) that provides single-sign-on authentication and authorization facilities [3].

V. gLite

For grid computing, gLite is another middleware. One of the unique features provided by gLite is that according to the requirements of users, they can implement the services without using the system as whole. As we know in medical sector, huge amount of data is produced, processed and manipulated. To run the medical image processing applications over the grid infrastructure, Medical Data Management approach is introduced.

In this approach, advanced gLite data management services are used by medical data management system to manage and store all the data in secured fashion [13].

VI. Legion

By the University of Virginia, Legion middleware project is introduced. For grid applications, Legion is object based metasystem software. Legion provides "vault "mechanism for persistent storage. The limitation of Legion is that it does not provide any mechanism to solve the issues like: Data load and Replica management [1].

VII. Condor and Condor-G

High throughput computing environment is provides by Condor. To perform the computational tasks, Condor harnesses the capacity of idle workstations. Without modifying the applications, it provides the facility to schedule and monitor the applications. The limitation of Condor is that it does not support the parallel applications [8]. The combination of Condor and Globus results in Condor-G software system. For grid applications, the purpose of Condor-G is to provide the proper job management services.

VIII. UNICORE

UNICORE is the Uniform Interface to Computing Resources. It provides the uniform Graphical User Interface (GUI) and security architecture where distributed resources can be accessed in secure fashion .It allows that without the user intervention, data movement function can be performed in well manner [9].

IX. NIMROD AND NIMROD-G

An interface is provided by Nimrod, where jobs can be independently submitted to a resource management system. A software system Nimrod-g is a combination of Nimrod and Globus. For the scheduling and management of computational resources that are geographically distributed world -wide Nimrod-G is introduced. The result of its implementation shows that how it is scheduling the tasks with in time and cost constraints in a well manner [11].

B. Toolkits

I. Grid Application Toolkit

The missing chain between the application level and various grid middleware packages is achieved via Grid application toolkit. To fulfil the needs of grid application developers and users, this toolkit introduced uniform programming interface. The main focus of Grid application toolkit is to the management of data, information, events and resources. As a result, we obtained that it is able to provide high level Application Programming Interface (API) that is really needed by application developers [11].

II. Simgrid Toolkit

In distributed environment, for distributed applications, it is necessary to simulate the scheduling algorithms. Simgrid toolkit is based on C language. For parallel applications, Simgrid toolkit is introduced for the simulation of scheduling algorithms. Its results describe that it provides efficient and accurate simulation. But still it is having some limitations like it is not able to model the computational grid environment aspects [12].

III. Gridsim Toolkit

A toolkit that provides simulation and modelling facilities for grid heterogeneous resources and users. To evaluate the performance of scheduler in controlling manner is very difficult and impossible task in grid systems, because in grid environment, users and resources are distributed across multiple organizations and every organization has different polices. So to overcome this limitation, Gridsim toolkit is developed that is based on java language [6].

IV. Gridbus Toolkit

Gridbus toolkit provides services for both for data and computational grid. For eBusiness and eScience applications, grid middleware technologies are developed by Gridbus project. This project has gained success world-wide. As future work, in grid environment it is trying to develop that kind of technologies that will be helpful for stock exchange creation [14].

V. Globus Toolkit

One of the most widely used for grid applications is Globus toolkit. Globus toolkit provides the facilities to access and to share the data, computing power and other heterogeneous resources across geographically distributed areas [10]. Globus toolkit version 4 is introduced that is based on web services. For implementing the web service interfaces GT4 provide web services containers. This version shows the various improvements in terms of functions that it performs like the better performance that it shows usability and standards that it covers over existing versions [12].

VI. ITK and VTK

ITK is Insight Segmentation and Registration Toolkit and VTK is Visualization Toolkit. In grid environment, for the development of biomedical image analysis applications these toolkits are introduced. It results that image analysis applications can be developed and execute in an efficient manner [13].

Table1: Summary of Grid Computing Middleware

All the above explained Grid middleware with their basic idea, advantages and disadvantages are summarized in the following table.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 1. Summary of Grid Computing Middleware with their advantages and disadvantages.

S. No	Middleware	Key Idea	Advantages	Disadvantages
1.	MPI Based Middleware	Message Passing Interface middleware is for developing the parallel applications.	It provides the better support to tightly coupled parallel applications	It does not provide support to loosely coupled parallel applications.
2.	Java RMI Based Middleware	It is introduced to develop large scale distributed applications.	It is platform independent.	Object replication problem is that on given object when method is invoked, then it can also be accessed by many other objects also [2].
3.	NetSolve	It does best effort to use the best computational resources available over the network.	It is client server application [4]. It is used for high performance computing [14].	It does not provide universal interface [14].
4.	Globus Based Middleware	For grid services, it is one of the most widely used middleware.	It provides various features like discovery of resources, heterogeneity and security issues.	It does not provide any common model or architecture to add any new functionality.
5.	gLite	For building the grid applications, this middleware provides the framework [13].	According to the needs, user can implement the services rather than using the whole system.	To add the more features, it requires further work.
6.	Legion	It is object based metasystem software.	It provides common model or architecture that will support the new functionality added by designer.	It does not address the issues like Management of data load and replica [1].
7.	Condor-G	This is derivative software that combines the features of both Condor and Globus.	Condor-G provides the various features in multi-domain environments like to access the resources in secure manner, to harness the capacity of idle workstations.	The limitation of Condor-G is that it does not support the parallel applications [8].
8.	UNICORE	For computational resources uniform interface is provided by this middleware.	It provides the interface and java based environment for preparing the jobs and to accessing the super computer resources in secure manner.	Interactive processes are not allowed here.
9.	Nimrod-G	The combination of Nimrod and Globus results in Nimrod-G middleware.	For accessing the computational resources, it facilitates transparent, simple and easy mechanism [11].	Scheduling policy followed here is fixed application oriented.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 2: Summary of Grid Computing Toolkits

The following table summarizes the all above explained various toolkits with their basic idea, advantages and disadvantages are summarized.

S. No	Toolkits	Key Idea	Advantages	Disadvantages
1.	Grid Application Toolkit	To establish the link between the application level and grid packages this toolkit is introduced.	It provides the various features like management of resources, data and events.	All possible use cases are not covered by this toolkit.
2.	Simgrid Toolkit	In the distributed environment, for parallel applications, this toolkit is introduced for the study of scheduling algorithms	It facilitates the simulation of scheduling algorithms.	Although it provides the simulation feature, but still it is having some limitation like it is unable to model the complex environments aspects.
3.	Gridsim Toolkit	For grid resources and users that are heterogeneous in nature, it facilitates the modeling and simulation facilities.	It overcomes the limitation of evaluating the performance of scheduler in controlled manner.	For the concurrent implementation it still needs lot of work.
4.	Gridbus Toolkit	Gridbus is introduced for ebusiness and escience applications.	It offers various services for computational and data grid.	For the other complicated applications like stock exchange creation there is need for development of other technologies at here.
5.	Globus Toolkit	For the grid applications, this toolkit results as one of the most widely used.	Various features are introduced like sharing and accessing the processing power, data and other resources over geographical distributed environment.	Globus toolkit is having missing functionality, if vendor does not provide the fabric level behaviour.
6	ITK and VTK	ITK is Insight segmentation Registration Toolkit and VTK is Visualization Toolkit.	It is introduced for the biomedical image analysis applications.	To gain the much better results, we can enhance its features also.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

III. CONCLUSION

From the above discussion of Grid Middleware and Toolkits, Finally I conclude that there are many middleware and toolkits available for the co-ordinated resource sharing, secure resource access, resources allocation, job scheduling, monitoring and management of jobs. Each approach has their own pros and cons. It could not said that anyone of the middleware and toolkits overcome the limitations of others, but with the hybrid approaches like Condor-G and Nimrod-G, we can overcome the limitations of single approach and it also results in better features and achievements.

REFERENCES

1. Hoschek, Wolfgang, Javier Jaen-Martinez, Asad Samar, Heinz Stockinger, and Kurt Stockinger. "Data management in an international data grid project." Grid Computing—GRID 2000 (2000): 333-361.
2. Maassen, J., Kielmann, T., & Bal, H. E. (2001). Parallel application experience with replicated method invocation. Concurrency and Computation: Practice and Experience, 13(8-9), 681-712.
3. Foster, Ian, Carl Kesselman, and Steven Tuecke. "The anatomy of the grid: Enabling scalable virtual organizations." International journal of high performance computing applications 15, no. 3 (2001): 200-222.
4. Baker, Mark, Rajkumar Buyya, and Domenico Laforenza. "Grids and Grid technologies for wide-area distributed computing." Software: Practice and Experience 32, no. 15 (2002): 1437-1466.
5. Krauter, Klaus, Rajkumar Buyya, and Muthucumaru Maheswaran. "A taxonomy and survey of grid resource management systems for distributed computing." Software: Practice and Experience 32, no. 2 (2002): 135-164.
6. Buyya, Rajkumar, and Manzur Murshed. "Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing." Concurrency and Computation: Practice and Experience 14, no. 13-15 (2003): 1175-1220.
7. Karonis, Nicholas T., Brian Toonen, and Ian Foster. "MPICH-G2: a Grid-enabled implementation of the Message Passing Interface." Journal of Parallel and Distributed Computing 63, no. 5 (2003): 551-563.
8. Goldchleger, Andrei, Fabio Kon, Alfredo Goldman, Marcelo Finger, and Germano Capistrano Bezerra. "InteGrade: object-oriented Grid middleware leveraging the idle computing power of desktop machines." Concurrency and Computation: Practice and Experience 16, no. 5 (2004): 449-459.
9. Yu, Jia, and Rajkumar Buyya. "A taxonomy of workflow management systems for grid computing." Journal of Grid Computing 3, no. 3 (2005): 171-200.
10. Buyya, Rajkumar, and Srikumar Venugopal. "A gentle introduction to grid computing and technologies." database 2 (2005): R3.
11. Allen, Gabrielle, Kelly Davis, Tom Goodale, Andrei Hutanu, Hartmut Kaiser, Thilo Kielmann, Andre Merzky et al. "The grid application toolkit: toward generic and easy application programming interfaces for the grid." Proceedings of the IEEE 93, no. 3 (2005): 534-550.
12. Foster, Ian. "Globus toolkit version 4: Software for service-oriented systems." Journal of computer science and technology 21, no. 4 (2006): 513-520.
13. Montagnat, Johan, Ákos Frohner, Daniel Jouvenot, Christophe Pera, Peter Kunszt, Birger Koblitz, Nuno Santos et al. "A secure grid medical data manager interfaced to the glite middleware." Journal of Grid Computing 6, no. 1 (2008): 45-59.
14. Xhafa, Fatos, Sabri Pllana, Leonard Barolli, and Evjola Spaho. "Grid and P2P middleware for wide-area parallel processing." Concurrency and Computation: Practice and Experience 23, no. 5 (2011): 458-476.
15. Buyya, R., Abramson, D., & Giddy, J. (2000, May). 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 (Vol. 1, pp. 283-289). IEEE.
16. Casanova, H. (2001). Simgrid: A toolkit for the simulation of application scheduling. In Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on (pp. 430-437). IEEE.
17. Hastings, S., Kurc, T., Langella, S., Catalyurek, U., Pan, T., & Saltz, J. (2003, May). Image processing for the grid: A toolkit for building grid-enabled image processing applications. In Cluster Computing and the Grid, 2003. Proceedings. CCGrid 2003. 3rd IEEE/ACM International Symposium on (pp. 36-43). IEEE.
18. Buyya, R., & Venugopal, S. (2004, April). The gridbus toolkit for service oriented grid and utility computing: An overview and status report. In Grid Economics and Business Models, 2004. GECON 2004. 1st IEEE International Workshop on (pp. 19-66). IEEE.

AUTHOR PROFILE

Pooja Rani is doing M.Tech (CSE) from Lovely Professional University, Phagwara, and Punjab, INDIA. Her research interests include software engineering.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)