European Online Journal of Natural and Social Sciences 2015; www.european-science.com
Vol.3, No.3 Special Issue on New Trends in Architecture, Civil Engineering, and Urban Studies
ISSN 1805-3602

Evaluation & measurement of Factors Affecting on Efficient Implementation of Grid Network: Case of Tarbiat Modares University (TMU)

Baharak Pourrajab[1*], Amir Albadvi[2]
[1]MSc. in Information Technology Engineering, Faculty of Engineering & Technology, Tarbiat Modares University, Tehran, Iran; [2]Prof., Faculty of Engineering & Technology, Tarbiat Modares University, Tehran, Iran
[*]E-mail: Pourrajab@modares.ac.ir

Abstract

Integrating computational resources is one of the aims of grid computing in order to quickly solve large-scale computational intensive problems, which are connected through network. This could make them integrated into virtual but very powerful supercomputers. Since the advent of grid technology, it has undergone major changes and among the important issues discussed in the grid network, optimal implementation of this technology could be mentioned. One of the effective factors in grid optimization is emplo 200 ideal broadband and desired communicational infrastructure, commensurated with the sent workload in grid environment. In this study, the requirements and effective factors in optimized implementation of service-based grid computing is presented, which expresses the way universities are connected to grid computing of the Institute of Basic Sciences and universal grid computing in three layers, along with a description of services in each layer. Then, the impact of optimal implementation of this technology is evaluated by investigating the factors affecting the grid performance.

Keywords: optimization of Grid, national Grid, architecture of service-oriented Grid

Introduction

The purpose of developing a more general concept of Grid technologies in sharing resource is to build a virtual organization of one or more management domains. Grid computing is the next generation of distributed systems and aims to create a powerful virtual machine, large and autonomous that of being together numerous disparate systems was created with the purpose of sharing resources. A Grid Computing can be in term of topologies one of three types of domestic, foreign and inter-organization.

Internal Grid Computing is called to organization without integration and with a cluster. External Grid Computing is called to multiple integrated clusters organizations and intra-organization Grid Computing is called to several organizations with multiple collaborations and a large number of multiple clusters.

In fact, the purpose of Grid is the exchange of information from a network environment to the other network environment in which computers can communicate with each other, without concern about being Heterogeneous facilities in exchange of processing information and saving. Respond to the appeals received. The more power in Grid Computing and total available resources Grid Computing is at higher level. Performing complex calculations on Grid Computing will be more successful and more computing power

In this study, evaluating affecting factors on efficient performance of Grid will be discussed. At first the components of evaluation and simulation software will be introduced, and then expressing used data collection and output of simulation software. Finally, we will discuss about assessing on obtaining results and evaluate affecting factors on efficient operation of Grid

Statement of problem

For that a project has been successfully done correctly and should meet the requirements for it, this case becomes more important for infrastructure projects such as Grid Computing. Due to the nature of Grid Computing infrastructure consisting of a set of resources that are geographically distributed and these resources must work together to do so. And other features such as seamless access to computing resources, requires the identification of infrastructure Grid Computing. Due to given explained, the organization must have a good networking hardware infrastructure to have successful Grid

Given the title, the main aim of study was to evaluate the affecting factors on efficient implementation of Grid computing in universities and scientific centers. According to this goal, the most important secondary objectives of this study was to examine the model and identify affecting factors in optimal implementation of Grid computing in universities and achieve a comprehensive understanding of that factors on the basis of the above objectives.

Research questions

What is proper implementation of Grid computing model in universities and scientific centers?

What are affecting factors on implementation Grid Computing?

What is the effect of factors on the implementation of optimized Grid Computing?

Literature review

With creating the idea of grid Grid Computing first time in 1999, as a new software platform to run applications on a large number of disparate sources and distributed, it is proposed passing through different areas of management and policies. In recent years, the idea of using network-level resources was created in this idea there is no central control over vast resources throughout the network level. But any computer as a member enter into network and was used certain techniques to control resources and it is responsible for controlling their resources

Grid includes systems and applications that manage and unify resources and services that have been distributed in several management areas. At first, Grid computing appeared for high-performance computing, as well as a source for sharing, with developing abilities, along with enhanced performance and cost reduction for organizations which want have ownership of their resources

Grid computing is sharing resource and solving problem coordinately in virtual dynamic organizations at multiple levels of the organization.

Grid provide good service and quality by open and standard multi-purpose relations, protocols for resources does not central control over them this definition make turbine feature bold

Grid by using the multi-purpose protocols and relations made that covers fundamental issues such as authentication, access control, resource discovery and access to resources. Standard and openness of this protocols and the equation is very important, otherwise we will be faced with a single purpose system

Quality services related to response time, throughput, availability and security provided or in other words, it is done the joint allocation of some resources for the implementation of complex operations so that the total resources under control have more value than individual resources.

Grid computing placed in the overall classification of distributed computing among the Standards and Architecture-of Grid, the most important standard which formed until now was open service-oriented architecture of Grid, which was created by the World Association of Grid. Open service-oriented architecture of Grid is an information description that defining a common architecture, standard and open for Grid functional programs. Its aim is to standardize all services

that a Grid application can use it. Such as resource management services, management-labor services communications, and security. The underlying basic technology for patterns of distributed computing, are web services. Many standards exist to build a service-oriented architecture for Grid. These standards form the main base of the building and let applications have ability to execute service requests

Grid divided into two general categories Data distribution and Grid Computing. Grid computing is a kind of Grid that is composed of high-performance computing capabilities and can do heavy computing tasks. In this type of Grid Computing have a network with high bandwidth is advantage. Data Grid has responsibility for the management and control massive amounts of distributed data sharing mechanism. These systems are often combined with computing systems Grid. The major difference is data Grid and Grid Computing, also it provides a special bed for using storage management and data access. But in Grid computing, this is an application that representing storage management. Sample of Grid applications is data Grid and mining data for specific purposes that information from different data sources make related together

In the mid-1990s, distributed processing was very interesting subject and multiple academic research projects was defined in this field.The main objective of this research designs was create a tool that can be managed distributed multiple computing resources just as a single and large system. The basic of distribution network create virtual organizations based on need that can achieve to goals such as cooperation, sharing, understanding, and trust. The purpose of the distribution network is cooperating to share resources and solving problems in virtual dynamic organization. Grid computing technology since the appearance till so far has four phases of major developments in both the scientific and industrial fields. Each of these steps is a major step in the evolution and spreading of this technology

After the initial private solution that was introduced in the early 90s (Cattlet, 1992), in 2002, the world of Grid Computing was witness the arrival of an open service-oriented architecture Grid as a comprehensive standard. Where all against implementations could work together and it was created with evolution of its open service-oriented architecture Grid

Open architecture services of Grid covers items such as how the authentication, create access level to services, dialogue policy and discovery services, how to integrate resources and manage them. FeaGridg of advantage and special services of object-oriented architecture and develop in applying this architecture in practice follow the related security issues with SOA, which in some cases differ from security principles of traditional information systems. Techniques like using of pathway organization service for implementation of security controls in architectural Design part and use of Ws-Security for message-level security are devoted to service-oriented architecture (Taghva, 2013)

Main and Basic services that provide open service-oriented architecture of Grid include:

Job management Service: The tasks of this service is distribution of tasks on the compute nodes of Grid and make coordination between them in such a way that the created complexity of the distribution remain hidden from the user. The main components of this service inquiries management system and other components of that service is chronicles that with saving happened events that happened in the process of executing a request on the Grid Computing provide complete history of them (Aiftimiei,2010)

Data Management Service: The service is hiding the way of doing data storage systems, checking permissions for data access, dissemination and use of information to monitor and explore the available resources, provided related index to data and services which using of them the user can informed of capacity and Grid Computing power in calculating distributed computing (Laure,2006)

Access service: a common framework for the possibilities access to the Grid services would be presented to the user and has permission for life cycle available services management for the user according to the user's access level (Ciaschini, 2006)

Security Service: it is responsible for verifying the identity of users, systems, services and access or cancel permits to resources and services. (Alfieri, 2005)

Located at a higher level set of available resources of Grid Computing, it will increase computing power. It is obvious that without a model to provide how this relation and explain created group requirements of Grid Computing is hard to reach such network. Given the importance of sublet will be described the structure of computing Grid model at universities in below.

Grid computing model structure in universities

According to the done survey of knowledge base research as director of Grid Computing in Iran, national Grid implemented in universities is composed of three layers as follows:

The first layer is services located in universities and scientific centers and are unique for each university or scientific center, including storage computing node, doing node and user interface. Of course, all these services are not required for each university, universities and scientific centers can share and compute nodes, operating and data warehouses in Grid computing. The source must be accessible by informing and recovery service that is implemented in IPM,

The second layer is consists of five service open service-oriented Grid architecture, Services layer is central and deal with concepts such as request and receive a digital certificate, create a proxy and submit job in Grid Computing. In fact, Users of Grid computing by utilizing this mechanism, will benefit of available computation and storage resources, in order to perform their tasks. Other central services is accessible so users can sent request to enter Grid Computing and, if approved, their appeals to be sent on Grid Computing resources to perform calculations. Data management service is another service that store, retrieve, read and write data on the Grid Computing. Universities and scientific centers can join optimally with the model and clustering their resources to Grid Computing. And in order to increase its computing power. Security services in layers, infrastructure and services have been seen. Each university with sharing its clusters and collaboration with other universities and scientific centers which they also share their resources and we will have an inter-organizational Grid at national level. On the other hand, if other services go forward to universities layer than all universities have internal Grid independently and by joining together can improve computing power of each other

Computing node is the interface node between Grid Computing and work nodes and its duty is management and labor division between work nodes

Work node consider as computing resources to do the job and its task do the user request. in fact,the computational power of Grid supply these nodes and do operation by making available a variety of resources

The third layer represents Grid Europe, which means that all authentication centers of Grid are under the Global legislative body and with joining Iran in the collection a way is made to join other countries virtual organizations

Experimental sample of model raised, in 2010 by Hashemi, and his colleague run at tarbiat Modares University. The results of this study state the advantage of using Grid Computing. Now, due to being efficient model, the factors affecting on the performance of layer related to Universities will be evaluated and reviews

Data collection methods and instrumenst of study

This research, in terms of the user goal and Descriptive survey data collection method, and it will be on the basis of quantitative research approach. In this study, gaining the main component to the study, was used the analysis of previous studies and comments of experts. In early reviews of

studies we check and gain the comments of several expert in this area and then the factors affecting on the performance of Grid is extracted and then check in simulator environment. In the meantime the present component of research have selected based on the frequency

Frequency component that they are less than 3 were removed and finally 3 components selected that expert has been confirmed these components. Professors and students (masters and PhD) are Statistical Population of the study in Tarbiat Modares university and the reason of doing that gain a mere feedback of experiences and understanding on influential factors on computing performance of Grid in order to implement this technology in the University

All Statistical population is 30 people. From a total of 30 people were study people with doctoral degrees have the highest frequency with the number 15 (of 50%) and professors with 5 (equivalent to 6.16 percent) had the lowest rate. Group Engineering with 22 (equal to 73.3 percent) have the highest one natural resources group 1 people (3.3 percent) had the lowest frequency. Most observation group had PhD in Drew engineering and adequacy of respondents with respect to education records have a good frequency. In order to achieve the exact information after obtaining the effective components, to evaluate the effect of factors is used simulator GridSim and output are expressed as R charts

Gridsim is discrete event simulation of Grid based on Java that has been used environment Netbeans for its development. To define and implement their several experiments, the simulator code has been added to the source code. This is a comprehensive tool allows you to simulate a different class of heterogeneous resources, users, applications, servers. This tool can be used to simulate the schedules application for distributed computing systems and used as clusters and Grids for single or multiple administrative domains

A simulated environment requires that all entities and all dependent interactions to time in real systems form in abstract. It supports the creation of user-defined functions dependent to time for providing communication needs. Including entities for the user, servers, resources, intelligence service, network and statistics

Findings of the study

In terms of hardware, Grid Computing entry will be provided by existing computers with the same wires connections, optical fibers and satellites. But, in terms of the application form is different. In Grid computing run software will be done through an middleware and determine an appropriate location to run the software depends on user choice or select an intermediate software in allowed systems and after sending the result, the bills utilization of the processor, sensors and databases were provided.

Because Grid computing is a distributed system in a distributed system, is the only method of communication between network resources, so the network infrastructure is the main requirements of Grid computing infrastructure. With regard to this case, all hardware devices and tools are used in the development of network infrastructure Grid Computing. It consists of cables, switches, routers and firewalls and must act correctly to establish a connection for network infrastructure. In other words, all network services related to Grid Computing by network infrastructure will be able to interact with each other.

It is obvious that joining the network is to Grid Computing to do things on a large scale, achieving high-speed and exchange of large amounts of data need high bandwidth. In our country, the current connection for Internet and data exchange, and run program needs to increase bandwidth. Of course, the bandwidth is very important in scientific institutions and major universities

Finally, the existing resources in a computing infrastructure, computing or storage facilities provide for Grid Computing. Computing power, storage space and computing infrastructure Grid

roughly equivalent to the total computational storage facilities resources. So, the more Grid Computing resources it is clear that computing power of Grid from point to point in terms of computational and data will be increased. The case is also correct about the capacity of resources, which means that the more capacity of resources, the more computing infrastructure Grid in computing power and data

Compute nodes consist of some working node in Grid processing environment, connected to each other through network ties. The main feature of these links, is bandwidth evaluated in units of bits per second. Bandwidth with units of bits per second means the number of bits that any link can pass in a second. In addition to these features, delay of sending link is also very important and its effect is considerable.

To evaluate the performance, efficiency can be defined as done work per unit of time. In the experiments, for those sent to run in computational nodes, are defined features such as time, run time on work node and latency expectations. Another feature is defined as network latency which includes the period between sending and receiving it by sending nodes except run time on host processor This measure consists of all network delays such as a delay in the router and wait in the queue, propagation delay and send in communications links and define as following criteria:

$$\propto_i = Jc_i - Js_i - Jx_i \quad (1)$$

Which represents a network delay job number and the time of receiving job results and sending time and it also show running time. It also reflects the time of execution of work. In other words, this number includes all delays sent through the network and independent of the runtime on the target processor. As a result, a good benchmark for network efficiency Grid when assessing bandwidth of communication links. Based on conducted research in several colleges of the study university and using the component values listed in Table 1, in which input parameters of emulator software were specified.

Table 1: Determination of input parameters of emulator software

Civil engineering	Electronic engineering	Mechanic engineering	biology	Parameter
Server 2008	Server 2008	Ubuntu 10.04	Red hat 5.5	operating system
1G/cat6	1G/cat6	1G/cat6	100M/cat5	Communication cables
Flow 3D	Comsol 4.3	OpenFoam	Amber v. 10.0	Application program
500 KB	20 MB	14 KB	1 MB	The size of the input file
70 GB	10 GB	500 MB	10 GB	Output file size

Checking the effects of bandwidth at the time of execution of the Works

In doing simulation through relation 1 to any of the works submitted, the total execution time is calculated. In Figure 1 and 2, the result of this test can be seen. In Figure 1, the amount of bandwidth varies between 1 and 15 Mbps. The result of total running time has been reported between zero and one for all operations. It is observed that by changing the bandwidth of one Mbps to 15 Mbps, the running time is reduced by 80 percent. Apperantly, this result shows that by increasing bandwith, the running time is reduced so much regard to this point i mb is a small number and link in universities support higher than 100MBS per second. In practice, such an outcome does not happen. Figure 2 repeats the same experiment for 100 Mbit bandwidth up to 1 Gbps. In this experiment, it was observed that changes in runtime is minimal and can be neglected.

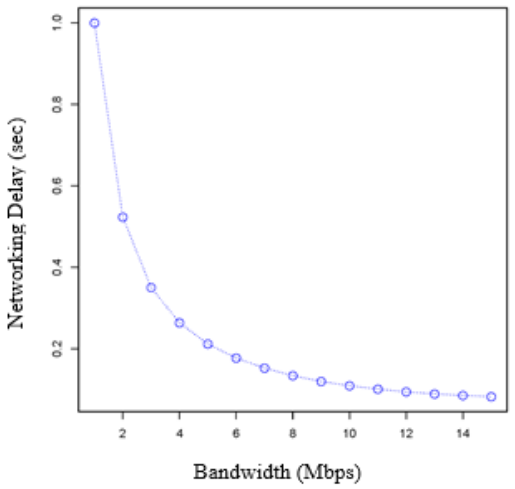

Figure 1: Effects of Bandwidth

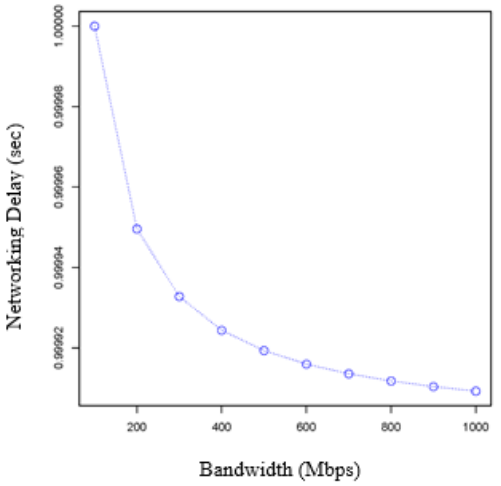

Figure 2: Effect of number of task on run time

Every job has a data input and data output. When running, the input data and are sent to the CN also the result after run the task is returned to the sending node. It is obvious that data exchange takes place via communication link as a result, bandwidth of the links affected on the exchange speed

Figure 3 shows task size effect on the speed of execution while amount of bandwidth is fixed. In this test, a task change of 100 KB to 1 MB and its impact on the running of these reported change. After the workload passes from 400 kilobytes, run time speed increases which can be concluded that the increasing volume of work is needed to exchange data through link go up. As a result, bandwidth of these links becomes a pathway to runtime

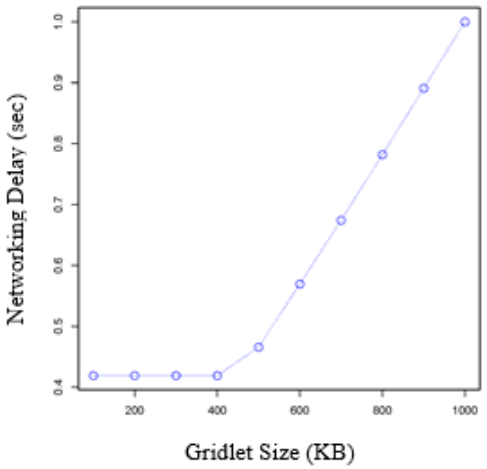

Figure 3: The effect of work size

The effect of the amount of delay in related links on running task

Another criterion in the assessment of effective communication links, is delays communication links. It should be noted that the delay in addition to links depends on the length of link. And, it should be keep in mind that design topology. In this experiment, increasing propagation delay network links Grid, the speed of performing tasks was measured. Figure 4 depicts the results of this test, it is observed that by increasing propagation delay, the length of the task increases almost linearly

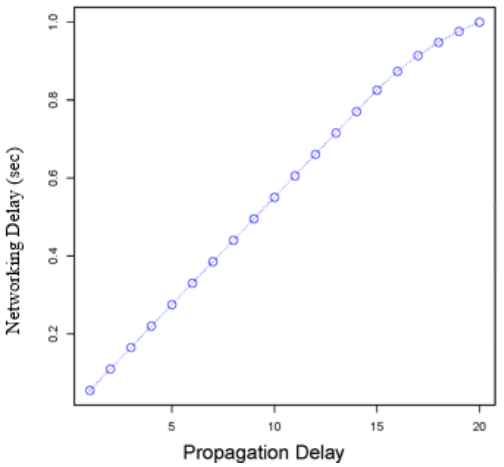

Figure 4: Effects of latency communication links

Checking the effect of Maximum Transmission Unit

Maximum Transmission Unit MTU is another factor that affects on network performance. In other words, the maximum capacity for each packet transmission in the network is sent, and the unit of is byte calculated. In each repeat experiment, the amount of bandwidth increase and observed its effect on efficiency. The result of this experiment is depicted in Figure 5, 512 bytes, 640 bytes and 1 KB shown as values of Maximum Transmission Unit in this graph. by increasing the MTU value network latency is reduced. the reason of this reduction is overall delay due to an increase in MTU smaller size packets with more volume are sent on the network level As a result, network overhead such as delays in router is lower

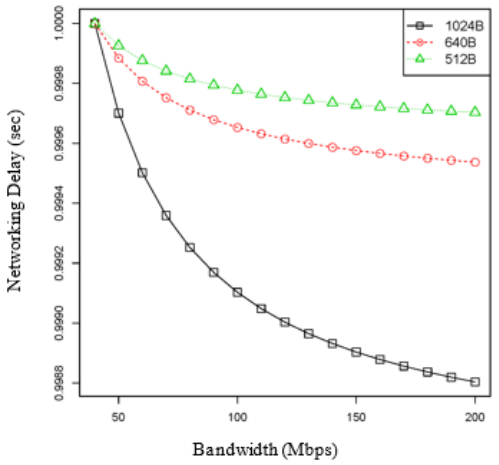

Figure 5: Effect of MTU transmission

Conclusion and Recommendations

For effective implementation of Grid, the different dimensions and aspects should take into consideration in order to increase efficiency. At general glance, this study is different from previous preceding studies .These above studies show a different approach. In other words, the perspective of infrastructure and management has been considered Grid.

For having a system of efficient Grid, we should pay attention to the high level of systems, including the management and distribution of tasks in the system, and considering infrastructure low level of the communication network. Given the nature of the underlying variability in a large Grid system, the management part must have complete information on the current state of infrastructure

of Grid system so that it can decide on the division of work optimally between different nodes. In total, after checking, 3 layers model can be implemented in universities, three components bandwidth, latency in communication links and size of sent task examined as affecting factors on the speed of performing tasks. In this study, 2 final cases related to infrastructure network were identified and the last one is how to distribute a task in Grid environment , in which high-level tasks of Grid are considered. In addition, the MTU value also has a network of low-level settings playing a decisive role in the overall performance of Grid system.

Experiments show a fundamental role of Network communication in the performance of Grid system. As a result, to raise the efficiency of a Grid system, we should pay special attention to the network connections between system node, Not only network infrastructure of Grid system must be optimized according to the requirements of the system, but also the structure and properties of this infrastructure must also be considered in the process of distributing computing tasks.

In terms of management, we must try the hardest to increase bandwidth and reduce latency in network infrastructure. For example, according to test results the length of the link increases in proportion to increased delay and the delay will be reduced if speed in any link increased. As a result, the execution speed linearly changes.

The effect of the MTU is in Grid environment performance and paying attention to low priority to network infrastructure layers. As can be seen, with increasing MTU, the value environmental performance Grid arises. This issue is also important to increase the speed and reduce network latency. The reason of overall reduction delay is due to increasing in MTU will send smaller packages with greater volume in the network. As a result, network overheads such as delay in the router is lower, As a result, by using features such as enabling Jumbo Frame, we can provide take advantage of MTU values higher than 1500 in Ethernet network and a significant improvement in system performance can be created.

After reviewing the two components of the technical infrastructure, the size of work as a component of the applications is examined and tested to show that, in testing conditions after the work, the overload passes from 400 kilobytes and run-time increases quickly that can be concluded by increasing volume of work, we need to exchange data via links improvement. As a result, bandwidth of these links convert to pathway runtime. This issue show the essential improves in the environment of Grid and its distribution among more systems.

Finally, this study shows that having sufficient knowledge about used infrastructure and topology Grid environment and the nature of the jobs posted for system to distribute task in Grid environment is essential. As a result, management software must be manually or automatically acquired these properties. For example, having sufficient knowledge of the processing power of multiple nodes, as well as bandwidth and their communication latency they can distribute task in processing system of Grid. These data related to one user distribute as possible as among nodes that have the lowest latency and highest bandwidth among each other. Also, the task size of each node is determined according to its processing power and communication features with other nodes

This study shows that, to increase the environmental performance of Grid, in addition to increased communication bandwidth, we should reduce the communication delay as much as possible. For this purpose, by optimizing the network settings (such as MTU), the length links is reduced as much as possible. Also, due to the overhead distribution, a task must be adjusted among different systems and the system performance should be reduced in case of work allocated to each node size, and the size of each work should be regulated optimally. To have maximum performance (for example, in our review environment, this size is about 400 kilobytes).

According to the obtained results in the simulation, the use of integration standards in the field of data , design modification applications programing, building a data center in a secure

physical environment with all network standards that all the servers and critical equipment sore are proposed. If possible, even the transfer of Grid processing resources of several organizations to the central location also by reducing communications delays, increases overall performance for the benefit of the organization. Moreover, it is used in dashboard way for managing user requests prior to submission to the Grid environment to estimate the amount of resources, bandwidth, and the required time to perform the posted task can increase the Grid environmental performance.

References

Butt, A., & Sumalatha, A. (2003).Grid computing portals and secure issues. Journal of Parallel and Distributed Computing, 63(10), 1006-1014.

Cattlet, C. (1992). In search of gigabit applications. IEEE Communication Magazine, 42-51.

Christoph, S., & Till, J. (2007). Web2.0 and SOA: Converging Concepts Enabling the Internet of Services.

Chunlin, L., & Layuan, L. (2009).An efficient coordination scheme between layers of grid architecture. Computer Standard & Interfaces, 31,110-119.

Ciaschini, V., Ferraro,A., Ghiselli, A., & Rubini,G. (2006).An Integrated Framework for VO-oriented Authorization, policy-based Management and Accounting. Paper sent at the Computing in High Energy and Nuclear Physics, Mumbai, India.

Cody, E., Sharman, R., Rao, R, & Upadhyaya, S. (2008).Security in grid computing: A review and synthesis. Decision Support Systems, 44,749-764.

Gottschalk, K., Graham, S., & Kreger, H. (2002). Introduction to webservice IBM System journal, 41(2), 170-177.

Grimshaw, A., Humphery, A.S., & Natrajan, A. (Mar 2004).A Philosophical and Technical Comparison of Legion and Globus. IBM Jounnal research and Development, 14(2), 233-254.

Humphery, M., Thompson, M.R., & Jackson, K.R. (March2005) Security for grids. Proc of IEEE, 93(3), 644-652.

Laure,E.,Fisher,S.M.,Frohner,A.,Grandi,C.,Kunszt.,P., & Krenek,A.(2006).Programming the Grid with gLite. Computational Methods In Science and Technology, 72(1), 33-45.

Mim Hashemi, F. (2010). Providing a practical model for the implementation of service-based Grid Computing at Tarbiat Modarres University. Master of Information Technology. Tarbiat Modares University.

Shread, P. (2004). Survey Finds Grid Becoming Strategic ITS Investment. Future generation Computer Systems, 20, 7-18.

Taghva, M. Izadi, M. (2013). The security of information systems with service-oriented architecture. Journal of Information Technology Management, 5 (3), 25-42.

Pourrajab, B. (2013). Assessing the factors affecting the efficient implementation of Grid computing at Tarbiat Modarres University. Master's thesis of Information Technology. Tarbiat Modares University.