A Review On Security In Grid Computing

Abdul Hadi Jehmica Bin Abdullah (Author)
Course Subject : Client-Server Computing (TS6234)
Fakulti Teknologi Dan Sains Maklumat (FTSM)
Universiti Kebangsaan Malaysia (UKM)
Student No. : P50117 ; E-mail : hadiFedEx@gmail.com

Keywords : Grid Computing, Secure Grid Computing, Trusted Grid Computing

Abstract

A Grid computing system is a geographically distributed environment that share resources amongst themselves. One primary goal of such a Grid environment is to encourage domain-to-domain interactions and increase the confidence of domains to use or share resources without losing control over their own resources and ensuring security or confidentiality for others. A comprehensive set of Grid Computing usage scenarios are presented and analysed with regard to security requirements such as authentication, authorization, or Confidentiality, Integrity and Availability (CIA). Grid security is enforced through trust update, propagation and integration across sites. Trusted Grid Computing demands robust resource allocation with security assurance at all resource sites. There is also a paper which analyses the unique security requirements of large-scale Grid Computing and develops a security policy and a corresponding security architecture. These show a broader goal to increase the awareness of security issues in Grid Computing.

1. Introduction

Grid computing is the combination of computer resources from multiple administrative domains applied to a common task, usually to a scientific, technical or business problem that requires a great number of computer processing cycles or the need to process large amounts of data. One of the main strategies of grid computing is using software to divide and apportion pieces of a program among several computers, sometimes up to many thousands. Grid computing is distributed, large-scale cluster computing, as well as a form of network distributed parallel processing [11]. The size of grid computing may be different from being small confined to a network of computer workstations within a corporation, for example, to being large public collaboration across many companies and networks. The idea of a confined grid may also be known as an intra-nodes cooperation whilst the notion of a larger, wider grid may thus refer to an inter-nodes cooperation. This inter or intra nodes cooperation across cyber based collaborative organizations are also known as ”Virtual Organizations” (VOs) [2].

It is a form of distributed computing whereby a ”super and virtual computer” is composed of a cluster of networked loosely coupled computers acting in concert to perform very large tasks [5]. This technology has been applied to computationally intensive scientific, mathematical, and academic problems through volunteer computing, and it is used in commercial enterprises for such diverse applications as drug discovery, economic forecasting, seismic analysis, and back-office data processing in support of e-commerce and Web services. What distinguishes grid computing from conventional cluster computing systems is that grids tend to be more loosely coupled, heterogeneous and geographically dispersed. Furthermore, while a computing grid may be dedicated to a specialised application, it is often constructed with the aid of general purpose grid software libraries and middle ware.

ahja
31 August 2009

2. Background

The goal of Grid Computing is to create a virtual organization (VO) across one or more physical organizations. Securing a Grid environment presents a distinctive set of challenges. The resulting value of the virtual organization to users in each of the physical organizations is that the users can be more productive, either in their own activities or in their collaborations with other people across the virtual organization (VO). This enhanced productivity is achieved by having access to a greater number or variety of resources such as computers, databases, or any others particular equipments.

Computational Grids are motiviated by the desire to share processing resources among many organizations to solve large-scale problems [3, 5]. Normally, a Grid is used for executing a large number of jobs at dispersed resource sites. Each site executes not only local jobs but also

jobs submitted from remote sites. Thus, job outsourcing becomes a major trend in Grid computing. However, job outsourcing faces the problems of inevitable security threats and doubtful trustworthiness of remote resources [6]. Indeed, Grid sites may exhibit unacceptable security conditions and system vulnerabilities [7, 8].

3. Grid Computing Security

Secure operation in a Grid environment requires that applications and services be capable of supporting a variety of security functionality, such as authentication, authorization, credential conversion, auditing, and delegation. Grid applications need to interact with other applications and services that have a range of security mechanisms and requirements. These mechanisms and requirements are likely to evolve over time as new mechanisms are developed or policies change. Grid applications must avoid embedding security mechanisms statically in order to adapt to changing requirements.

While scalability, performance and heterogeneity are desirable goals for any distributed system, the characteristics of computational grids lead to security problems that are not addressed by existing security technologies for distributed systems. For example, parallel computations that acquire multiple computational resources introduce the need to establish security relationships not simply between a client and a server, but among potentially hundreds of processes that collectively span many administrative domains. Furthermore, the dynamic nature of the grid can make it impossible to establish trust relationships between sites prior to application execution. Finally, the interdomain security solutions used for grids must be able to interoperate with, rather than replace, the diverse intradomain access control technologies inevitably encountered in individual domains [9].

4. Issues And Challenges

4.1. Grid Security

We introduce the grid security problem with an example illustrated in Figure 1. We imagine a law enforcer in the Ministry of Home Affairs (MOHA), a member of a multi-institutional fight crime collaboration, who receives e-mail from a colleague regarding a new data set of crime scenes. He starts an analysis program, which dispatches code to the remote location where the data is stored at the Special Branch, Royal Malaysian Police (site C). Once started, the analysis program determines that it needs to run a simulation in order to compare the experimental results with predictions. Hence, it contacts a resource forensics service maintained by the collaboration at the Commercial

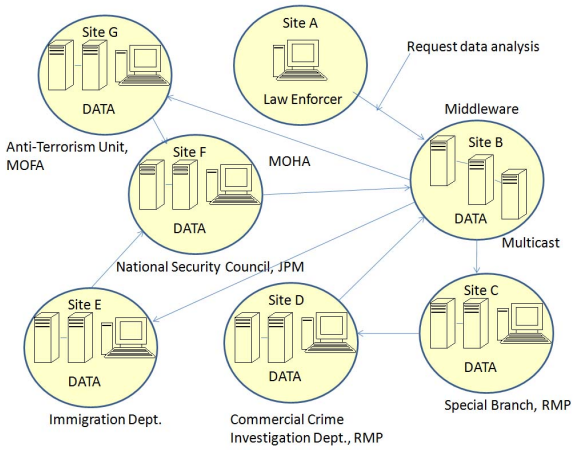

Figure 1. Example of a large-scale distributed computation: user initiates a computation that accesses data and computing resources at multiple locations.

Crime Investigation Department, Royal Malaysian Police (at site D), in order to locate idle resources that can be used for the simulation. The resource forensics in turn initiates computation on computers at Immigration Department and Anti-Terrorism Unit, Ministry of Foreign Affairs (at sites E and G). These computers access parameter values stored on a file system at yet another site at National Security Council, Prime Minister Department (F) and also communicate among themselves (perhaps using specialized protocols, such as multicast) and with the middleware, the original site, and the user.

From a security standpoint, the users, the applications, or the grid middleware or some combination of the three must be trusted [10]. In dynamic, scalable, wide-area computing environments, it is generally impractical to expect that all users can be held accountable for their actions. Accountability comes after the damage has been done, making this a costly solution. Another option is to trust the applications. This is typically accomplished either by constraining the development environment to a point where the generated applications are guaranteed to be safe or by making sure that the applications come from a trusted source. However, limiting the functionality of applications also limits the usefulness of the computing environment. History has shown that it is too easy for applications from trusted sources to contain bugs that compromise the integrity of resources [10].

Figure 1 Example Illustrates Many Of The Distinctive Characteristics Of The Grid Computing Environment :

4.1.1. The user population is large and dynamic. Participants in such virtual organizations as this scientific collaboration

will include members of many institutions and will change frequently;

4.1.2. The resource pool is large and dynamic. Because individual institutions and users decide whether and when to contribute resources, the quantity and location of available resources can change rapidly;

4.1.3. A computation (or processes created by a computation) may acquire, start processes on, and release resources dynamically during its execution. Even in our simple example, the computation acquired (and later released) resources at five sites. In other words, throughout its lifetime, a computation is composed of a dynamic group of processes running on different resources and sites;

4.1.4 The processes constituting a computation may communicate by using a variety of mechanisms, including unicast and multicast. While these processes form a single, fully connected logical entity, low-level communication connections (e.g., TCP/IP sockets) may be created and destroyed dynamically during program execution.

4.1.5. Resources may require different authentication and authorization mechanisms and policies, which we will have limited ability to change. In Figure 1, we indicate this situation by showing the local access control policies that apply at the different sites. These include Kerberos, plaintext passwords, Secure Socket Library (SSL), and secure shell.

4.1.6. An individual user will be associated with different local name spaces, credentials, or accounts, at different sites, for the purposes of accounting and access control. At some sites, a user may have a regular account. At others, the user may use a dynamically assigned guest account or simply an account created for the collaboration.

4.1.7. Resources and users may be located in different countries. To summarize, the problem we face is providing security solutions that can allow computations, such as the one just described, to coordinate diverse access control policies and to operate securely in heterogeneous environments.

4.2. Virtual Organizations (VOs)

Security requirements within the Grid environment are driven by the need to support scalable, dynamic, distributed virtual organizations (VOs) collections of diverse and distributed individuals that seek to share and use diverse resources in a coordinated fashion [6]. From a security

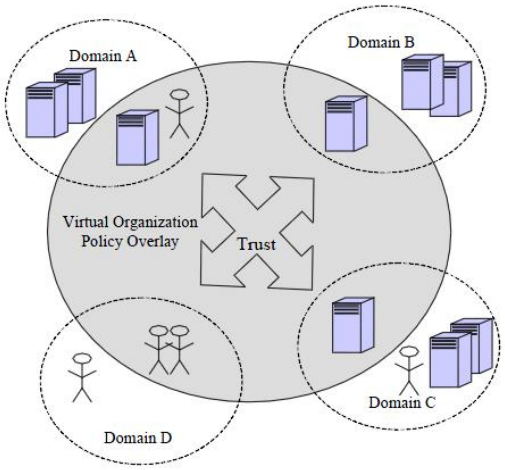

Figure 2. Virtual organization policy domain overlay pulls together participants from disparate domains into a common trust domain [2].

perspective, a key attribute of VOs is that participants and resources are governed by the rules and policies of the classical organizations of which they are members. Furthermore, while some VOs, such as multiyear scientific collaborations, may be large and long-lived (in which case explicit negotiations with resource providers are acceptable), others will be short-livedcreated, perhaps, to support a single task, for example, two individuals sharing documents and data as they write a proposalin which case overheads associated with VO creation and operation have to be small.

A fundamental requirement is thus to enable VO access to resources that exist within classical organizations and that, from the perspective of those classical organizations, have policies in place that speak only about local users. This VO access must be established and coordinated only through binary trust relationships that exist between (a) the local user and their organization and (b) the VO and the user. We cannot, in general, assume trust relationships between the classical organization and the VO or its external members. Grid security mechanisms address these challenges by allowing a VO to be treated as a policy domain overlay as shown in Figure 2 [2]. Multiple resources or organizations outsource certain policy control(s) to a third party, the VO, which coordinates the outsourced policy in a consistent manner to allow for coordinated resource sharing and use.

5. Conclusion

Indeed, to completely specify a job security demand, we need to use complex vectors of attributes to fully specify the

requirements involving all of the aforementioned parameters. This is obviously an unreasonable burden on Grid users. Unfortunately, up to now, there is no effective methodology to assess trust index of resource sites. A weighted sum of security parameter values will not work, because it is difficult to deterministically determine the weights and even the correct set of parameters to be included in real-time [4]. We can see that the matching of the job security demand with the site trustworthiness is an important system issue, which was largely ignored by the cyber security community.

Furthermore, the grid computing systems are being positioned as a computing infrastructure that will enable pools of resources to be shared across institutional boundaries. Unfortunately, the idea of sharing poses some concerns such as privacy and confidentiality. Hence, 'trust' should be addressed in such a distributed environment. Apart from that, the most significant challenges for Grid computing is to develop a comprehensive set of mechanisms and policies for securing the Grid.

Acknowledgment

I am grateful and would like to thank Mr. Mohd Zamri Murah (The Lecturer and Instructer) for his fruitful advice, thought and knowledge in order to complete this paper. This review paper is submitted as required for each students to do so for the TS6234 course under the Masters of Information Technology (Management Information System) at the National University of Malaysia.

References

[1] Humphrey, M. and Thompson, M.R., *Security Implications of Typical Grid computing Usage Scenarios*, Cluster Computing. Springer, 1999.

[2] Welch, V. and Siebenlist, F. and Foster, I. and Bresnahan, J. and Czajkowski, K. and Gawor, J. and Kesselman, C. and Meder, S. and Pearlman, L. and Tuecke, S., *Security For Grid Services*, Twelfth International Symposium on High Performance Distributed Computing (HPDC-12) Springer, 2003.

[3] Berman, F. and Fox, G. and Hey, A.J.G., *Grid Computing: Making The Global Infrastructure a Reality*, Journal of Grid Computing. Wiley, 2003.

[4] Song, S. and Hwang, K. and Kwok, Y.K., *Trusted Grid Computing With Security Binding and Trust Integration*, Journal of Grid Computing. Springer, 2005.

[5] Cosnard, M. and Merzky, A., *Meta-and Grid-Computing*, Journal of Grid Computing. Springer-Verlag London, UK, 2002.

[6] Nagaratnam, N. and Janson, P. and Dayka, J. and Nadalin, A. and Siebenlist, F. and Welch, V. and Foster, I. and Tuecke, S., *The Security Architecture For Open Grid Services*, Journal of Grid Computing. Version, 2002.

[7] Humphrey, M. and Thompson, M.R., *Security Implications of Typical Grid Computing Usage Scenarios*, Cluster Computing. Springer, 2002.

[8] Hwang, S. and Kesselman, C., *A Flexible Framework For Fault Tolerance In The Grid*, Journal of Grid Computing. Springer, 2003.

[9] Hwang, S. and Kesselman, C., *A security Architecture For Computational Grids*, Proceedings of the 5th ACM Conference on Computer and Communications Security. ACM New York, NY, USA, 1998.

[10] Butt, A.R. and Adabala, S. and Kapadia, N.H. and Figueiredo, R.J. and Fortes, J.A.B., *Grid-Computing Portals And Security Issues*,Journal of Parallel and Distributed Computing. New York, NY: Academic Press, 2003.

[11] Humphrey, M. and Thompson, M.R. and Jackson, K.R., *Security For Grids*,Proceedings of the IEEE. New York, NY: Citeseer, 2005.