A SECURE GRID ENABLED SIGNATURE VERIFICATION SYSTEM

M. F. Tolba M. S. Abdel-Wahab I. A. Taha A. M. Al Shishtawy

alshishtawy@yahoo.com, Tel. +20105690130
Scientific Computing Department
Faculty of Computer and Information Sciences
Ain Shams University
Cairo, Egypt

Abstract:

The emerging Grid technologies has changed the way people think about computation by presenting new paradigms and application models. This paper is a part of a joint research project between Ain Shams University in Egypt and George Washington University in USA to build a system for signature verification. The paper present a new approach to solve such problems using Grid approaches to increase performance and resource utilization while reducing the maintenance costs and security risks of the system. The limitations of other possible solutions and the advantages of Grid solutions have make it a good paradigms for future applications. A testbed was created linking the two universities was used to test the proposed architecture and prove its applicability in real applications.

Keywords:

Computational Grids,Grid enabled applications, Hand written signature verification.

1. Introduction

The Grid [4][8] or computational grids are a new approach to computing and considered to be one of the emerging technologies that will change the world [13]. It is defined in [5] as "coordinated resource sharing and problem solving in a dynamic, multi-institutional virtual organizations".The Grid also enables new application models that are more suitable to its nature such as distributed supercomputing, high-throughput computing, on-demand computing, data-intensive computing, and collaborative computing [1]. The work presented in this paper is a part of a research project between Ain Shams University (ASU) in Egypt and George Washington University (GWU) in USA. The goal of the project is to create a hand written signature verification system. This means that the user of the system will be able to check if a hand written signature he/she have (on a contract for example) is a genuine or forged signature. This paper presents the general architecture of the system, focusing on the advantages of using the Grid technologies over other techniques. More details about the implementation of the system architecture can be found in [10]. Details about the signature verification algorithm is presented in [9].

Section 2 briefly describes the problem of hand written signature verification. a discussion of different approaches to solve this problem is presented in section 3 showing the advantages and drawbacks of each and why the grid approach is better. Section 4 discusses the testbed used for testing and the results of different experiments. Finally the general conclusions are presented in section 5.

2. The Signature Verification Problem

The problem of handwritten signature verification has four main components that can be described in the following four points:

1. **The Database:** There are databases storing images of genuine signature. The system can only verify signatures for persons having their genuine signatures stored in these databases. For each person there is one or more signatures

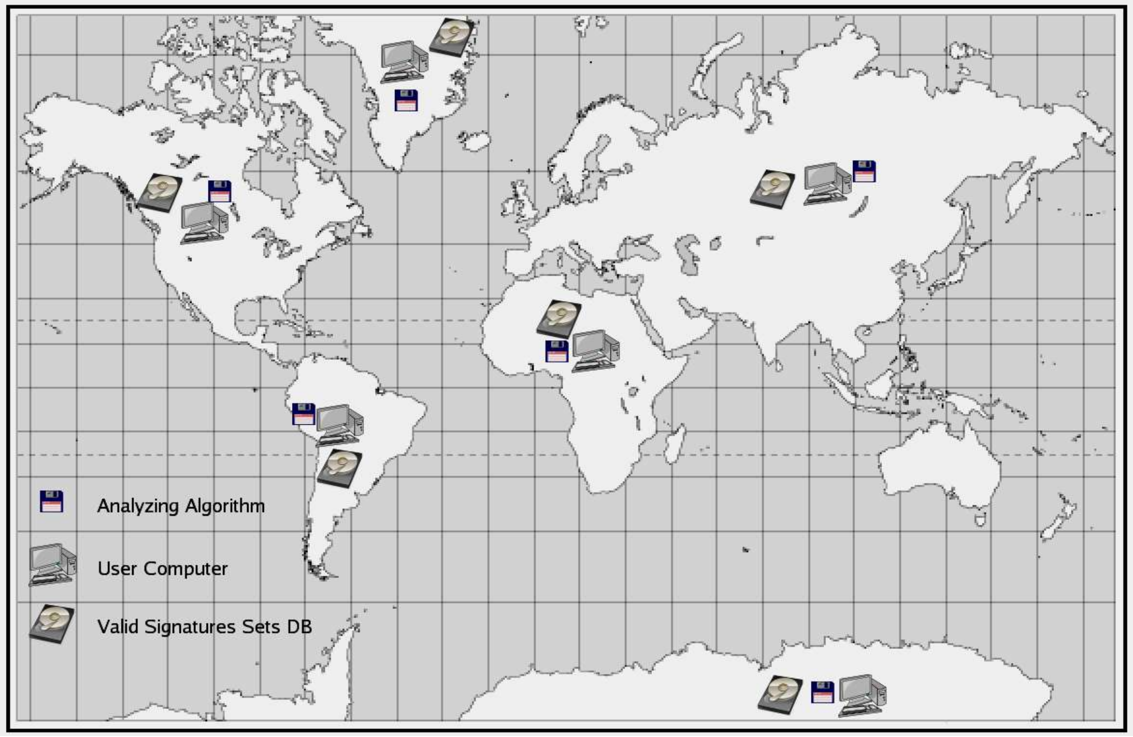

Figure 1. The old scenario.

set. Each set consist of a number of genuine signature images of that person. Each set also has some properties - meta data – describing its contents such as the number of signature images, the date these signatures where signed by the person, the conditions at which the person signed these signatures, the signature image resolution, and so on.

2. **Suspected Signature:** The user of the system should have a person's signature image that is needed to be verified by comparing it with the genuine signature images in one of that person's signatures sets.
3. **The Analyzing Algorithm:** The system can accept one or more algorithm capable of analyzing suspected signature images and decide whether they are forged or not. These algorithms may have different characteristics and requirements such as complexity, execution time, accuracy, signature image format, number of valid signature images needed, and so on.
4. **The User's QoS:** According to the user's desired Quality of Service (QoS). The suspected person signature will be compared to one or more of his signature sets in the database – with different attributes and quality – by one or more analyzing algorithm – with different requirements and accuracies – to achieve the desired user's goals.

3. The signature verification system architecture

Although the problem sounds simple, it contain many complex hidden issues and trade-offs and many possible solutions. But generally these solutions can be classified into one of the possible three scenarios.

3.1. The Old Scenario

This is the simplest – but not the best – solution to this problem. It is simply to have the whole signature verification system as one entity at each user that wants to use it as shown in *Figure 1*. Going back in time a couple of decades at the beginning of the computer revolution maybe it would the only possible solution. In this scenario any user of the system must have a database management system with a database filled with genuine signature sets for all possible persons that the user of the system may need to verify. The user must also have all the signature analyzing algorithms needed to analyze the signatures according to his QoS. A computer system capable of providing the needed computational power whenever it is needed by the algorithms to analyze the signatures must be available and dedicated for this purpose. This system must also be regularly administrated and maintained by updating the database with new signature sets, checking for new algorithms and updated versions, and maintaining and upgrading the computer system

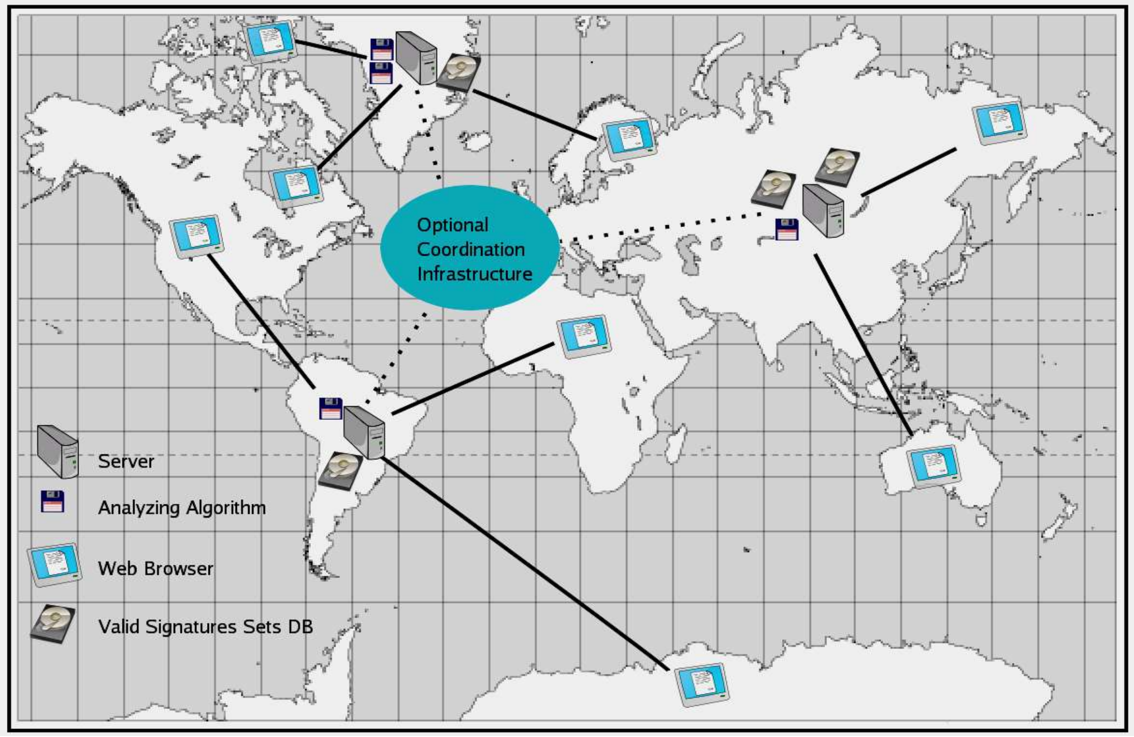

Figure 2. The modern scenario.

hardware.

This scenario with its architecture and requirements introduced many problems. The database size may grow to be very large and hard to maintain by the user of the system when it is filled with millions of person signature sets. Such system – with replicating its components at each user – will have a very high maintenance cost and poor cost/performance ratio. This scenario contains security problems. The user of the system must implement security mechanisms to protect this database and to verify that its content was not compromised by any undetected attack on the system. The user must also verify the origin of the signature sets and the algorithms to assure that they are correct. Most importantly it is unsafe and very risky and to give a database with all genuine persons signatures to all users of the system, because there is no way to guarantee that none of these users will miss use this valuable information to create professional forged signatures that are hard to detect.

3.2. The Modern Scenario

With the widespread of the Internet and web enabled applications this problem may be solved in much elegant ways. modern technologies may be used such as web services and client/server application. In one of the possible scenarios, as shown in *Figure 2*, the user acquires appropriate credentials(private keys, password, ...) to verify his identity. After that, using a web browser, the user browse to one of the system's home page then authenticate and login using his credentials. The user uploads the signature image that needs to be verified. The server analyzes the signature and send back the result to the user.

In this scenario all what the user should have is a valid credential to prove his identity, a web browser, and of course the signature image he want to verify. This scenario sounds quite simple but only from the user (client) side. But from the server side there are several solutions with their problems. It could be just having multiple servers (mirrors) distributed around the world to divide the workload among them. Each server will maintain a complete system with a database of all valid persons signature sets, all analyzing algorithms, and the necessary computational power to handle the users requests. This is simple but have almost the same drawbacks of the old scenario.

An alternative solution for the server side is to have a real distributed system with a distributed database of valid signature sets, for example a database for each country, and having multiple processing nodes each specialized in one or more analyzing algorithms. This system will solve almost all the problems. Each valid signature set is maintained at a single database which ease the maintenance, reduces security risks, and reduces the database size. The cost is distributed among the database and computation servers. Each site will be responsible only for its part of the system not

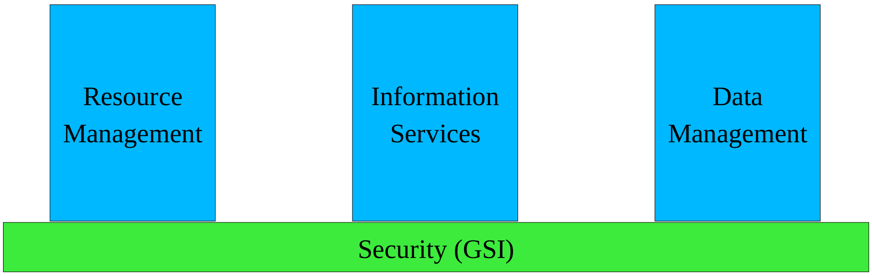

Figure 3: Basic Grid Services

the whole system.

But on the other hand this architecture will require the implementation of a specialized infrastructure (shown by the circle in the middle of *Figure 2*) that is among others capable of:

1. Locating the available signature sets databases.
2. Searching these databases for a person's valid signatures sets.
3. Locating free analyzing nodes with appropriate algorithms.
4. Handling possible failure of any of this system components.
5. Implement appropriate security mechanisms to control access to databases and analyzing nodes.
6. Provide a mechanism to allocate and start computation on the analyzing nodes.

Such architecture is very complex and hard to be specially implemented for a specific application of the hand written signature verification system.

3.3. The Grid and Security

The Grid – from it's definition – is used to coordinate and couple resources that are geographically distributed and administrated independently to create virtual organization that enables collaboration and seamless access to computational resources. Put in another way, the Grid simply implements all the required infrastructure discussed in the previous scenario (shown by the circle in the middle of *Figure 2*). The Globus ToolKit 3 [2][11] is a software toolkit that implements the grid infrastructure. It provides basic services needed by any Grid enabled application. These services are resource management [6], information services [7], data management [12], and security[3] as shown in *Figure 3*. Security is important for the success of a Grid environment. The Globus ToolKit addressed the security issue through the Grid Security Infrastructure (GSI) that is a component of the toolkit and supports all other services as in *Figure 3*, It is currently based on public key infrastructure (PKI) and it requires all users and resources to have appropriate certificates to join the Grid environment. GSI also provide basic services such as encryption, single sign on, mutual authentication, among others.

3.4. The Grid Scenario

Using the Globus ToolKit 3 the scenario – as shown in *Figure 4* – will start by the user login using appropriate credential and creates the user proxy. The user will start the Globus enabled signature verification application and provide it with the signature image needed to be verified. The program will contact the monitoring and discovery service (MDS) – a part of the Grid information services – to find available valid signature sets and processing nodes with appropriate analyzing algorithms. According to the user QoS the application will pick one or more available processing node(s) and then send to them suitable valid signatures set and the suspected signature image. After processing ends the results are sent back to the user application and displayed.

In this scenario the user will have a valid credentials to prove his identity, the signature image, the Globus infrastructure installed, and the Signature verification application. In this scenario we can notice the following:

1. It is a dynamic environment where valid signatures sets can be added or removed and processing nodes can join or leave as they want. This environment is shared and coordinated through the Globus infrastructure.
2. The valid signature sets are protected by being kept at secured databases and only sent to trusted and authenticated

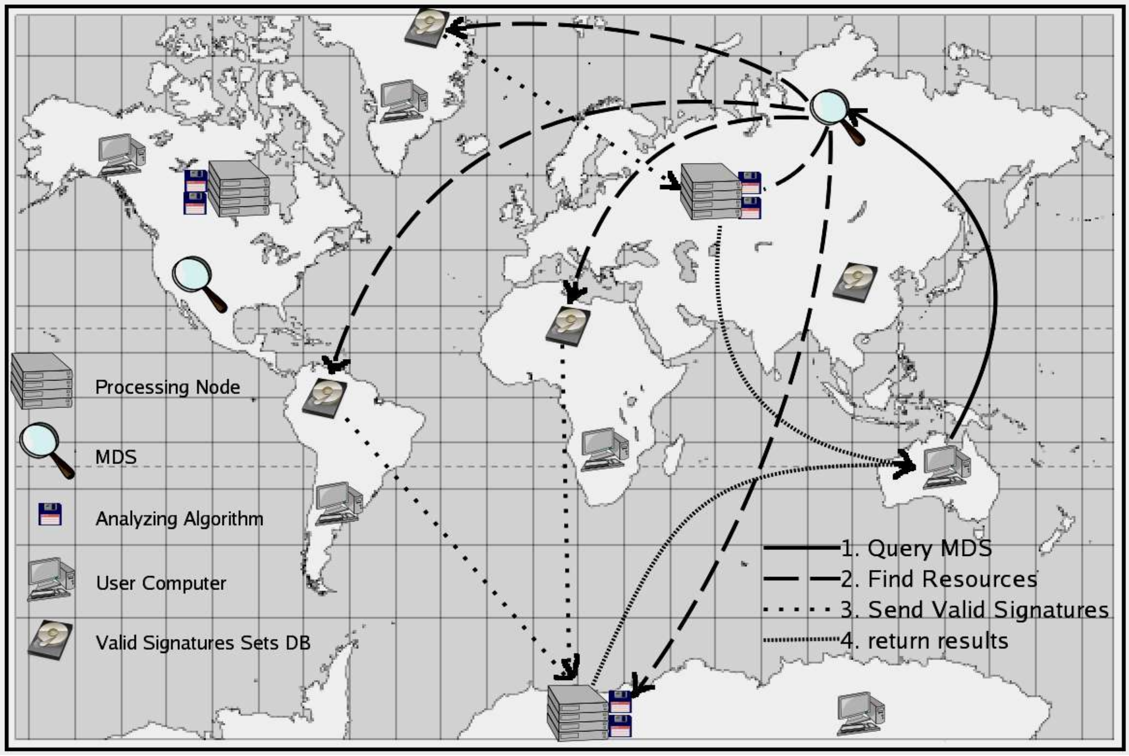

Figure 4. The Grid scenario.

analyzing nodes based on the choice of the local system administrators of the signature database.

3. The coordination process is seamless and all complexities are hidden by the Globus infrastructure. These complexities are in security, resource management, data management, and information services.
4. Signature algorithms are updated and improved locally and then discovered globally by MDS without requiring changing of the user application.
5. Within this large pool of resources the user can deliver the desired QoS by combining the appropriate pair of signature set and analyzing algorithm.
6. User authentication and secure signatures transfer over public insecure networks are done using the Grid Security Infrastructure GSI [3]
7. There can be different trust relationships among users, databases and processing nodes owners. The Grid Security Infrastructure supports this trust relationships. Although it is a one global system, each component (user, database, or processing node) works only with trusted components. This trust relation is defined locally by the decisions of the component's system administrators. For example a database may deny the request to send signatures sets to a processing node if it is not trusted. In the Globus ToolKit 3 this is done now using Public Key Infrastructure and hierarchical Certificate Authorities

4. Results

To test the proposed grid scenario a testbed was created consisting of three nodes. The first one is called GWU and is located in USA at George Washington University. The other two nodes are called ASU1 and ASU2 and are located in Egypt at Ain Shams University. GWU contains both a signatures database and an eight node cluster for distributed signature analysis. ASU1 consists on only of an eight node cluster for distributed signature analysis. ASU2 contains only a signatures database. The three nodes where linked using the Globus ToolKit 3. The distributed hand written signature verification algorithm [9] is installed on the clusters at GWU and ASU1. A grid enabled application [10] was developed that enables the users of the system to locate available processing nodes with appropriate algorithms, locate and search the databases for a specific person's signatures sets, select the best signatures set and processing node for the user, upload the suspected signature, and finally return the result of analysis back to the user. The system user can be located anywhere in the world, all what is needed it appropriate infrastructure installed, the grid enabled application, and appropriate credentials to use the system.

The system total execution time was measured with respect to three parameters: The location of the user, the location of the signatures database node, and the location of the analysis node, each of which can be – in this testbed – either in Egypt or in USA. According to the user QoS, the processing is done on the nearest available analyzing node to the signatures database node. The following scenarios can occur independent on the user location:

1. Signatures are available on GWU and GWU is chosen for analysis.
2. Signatures are available on GWU and ASU1 is chosen for analysis.
3. Signatures are available on ASU2 and ASU1 is chosen for analysis.
4. Signatures are available on ASU2 and GWU is chosen for analysis.

EGYPT		**Database**		
		ASU1	**ASU2**	**GWU**
Processing	**ASU1**	X	0.27	1.27
	ASU2	X	X	X
	GWU	X	1.31	0.51

Table 1. Execution time in minutes of experiments when user is in Egypt.

USA		**Database**		
		ASU1	**ASU2**	**GWU**
Processing	**ASU1**	X	0.45	1.26
	ASU2	X	X	X
	GWU	X	1.14	0.25

Table 2. Execution time in minutes of experiments when user is in USA.

Table 1 and Table 2 summarize the results obtained from running the grid application, the first when the user is in Egypt, and the second when the user is in USA when the analysis is completed without any errors. The columns represent the node on which the signatures set was found and the rows represent the node on which processing was done. A cell with value X means that this case is not applicable due to the assumptions that the signatures do not lie on all nodes and that not all nodes can do analysis. The average execution time when the user was in Egypt was found to be 0.84 minutes and that when the user is in USA was found to be 0.775 minutes. This difference in the averages is attributed to the connection speed in Egypt site, as it is based on a ADSL line with upload to download ratio is 1:4.

Executing the same algorithm on a sequential machine took about 34 seconds. The Grid enables the access to more powerful resources – the cluster in this case – to perform computations. From this point of view it is fare to compare the Grid enabled system performance that took 49 seconds in average to the sequential single machine system performance. It can be noticed that the delays due to communication was recovered by faster computation resulting in approximately the same performance.

5. Conclusions

The Grid infrastructure provided a seamless way to efficiently link and access different distributed computational resources. Providing an ideal development environment to create innovative applications such as the hand written signature verification system presented here. It enabled the efficient use of the available resources – databases, computations, and algorithms – by an intelligent Grid enabled application that used the underlying infrastructure to efficiently satisfy the desired user's QoS. The system was also secure by keeping the genuine signatures sets at secure databases away from the users. And also by using the Grid Security Infrastructure for authenticating users and transferring signature images. The different tests performed on the system showed good performance compared with single machine systems proving the applicability and advantages of using Grid paradigms for future applications.

6. References

[1] Foster I. , and C. Kesselman (Eds). “The Grid: Blueprint for a New Computing Infrastructure”, *Morgan Kaufmann*, 1999.

[2] I. Foster, and C. Kesselman, “The Globus Project: A Status Report”, *Proc. IPPS/SPDP '98 Heterogeneous Computing Workshop*, pp. 4-18, 1998.

[3] I. Foster, C. Kesselman, G. Tsudik, and S. Tuecke, “A security architecture for computational grids”, *Fifth ACM Conference on Computers and Communications Security*, November 1998, pp. 83-91.

[4] I. Foster, C. Kesselman, J. Nick and S. Tuecke, “The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration”, *Globus Project*, 2002. www.globus.org/research/papers/ogsa.pdf.

[5] I. Foster, C. Kesselman, and S. Tuecke, “The Anatomy of the Grid. Enabling Scalable Virtual Organizations”, *International Journal of Supercomputer Applications*, 2001.

[6] K. Czajkowski , I. Foster, C. Kesselman, S. Martin, W. Smith, and S. Tuecke, “A Resource Management Architecture for Metacomputing Systems”, *Technical report, Mathematics and Computer Science Division*, Argonne National Laboratory, Argonne, Ill., 1997.

[7] K. Czajkowski, S. Fitzgerald, I. Foster, and C. Kesselman, “Grid Information Services for Distributed Resource Sharing”, *Proc. 10 th IEEE Symp. On High Performance Distributed Computing*, 2001.

[8] M. Baker, R. Buyya, and D. Laforenza, “Grids and Grid Technologies for Wide-Area Distributed Computing”, *Software Practice and Experience*, 2002

[9] M. Tolba, I. Taha, M. Al Shandawely, “Building a Grid-Enabled Distributed System to Solve Signature Verification Problem Based on Improving QoS”. *Second International Conference on Intelligent Computing and Information Systems*, Cairo, Egypt, March 2005.

[10] M. Tolba, M. Abdel-Wahab, I. Taha, A. Anbar, “Fault Tolerant Scheduling for Grid Enabled Signature Verification System”. *Second International Conference on Intelligent Computing and Information Systems*, Cairo, Egypt, March 2005.

[11] The Globus Project™ Home Page. http://www.globus.org/ .

[12] W. Allcock, A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and S. Tuecke, “The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets” *The Journal of Network and Computer Applications*, 2001.

[13] W. Roush, A. Goho, E. Scigliano, D. Talbot, M. Waldrop, G. Huang, P. Fairley, E. Jonietz, and H. Brody, “10 Emerging Technologies That Will Change The World”, *Technology Review*, MIT, 106:33--49, Feb 2003.