DISTRIBUTED ENCRYPTION USING JPVM

SARUCHI
CSE/IT Department Lovely Professional University, Phagwara, Punjab

ganpati.saruchi@gmail.com
+91-9464055884

1.1 Introduction

1.1.1 An Introduction to encryption

Encryption transforms the original information called clear text or plain text into transformed information called code text, simply cipher or cipher text which usually has the appearance of unintelligible and random data. The transformed information in its encrypted form is called the cryptogram. Encryption transforms the message in such a way that its contents are hidden from unauthorized readers (Lehtinen R, Russell D and Gangemi G T, 2006). Encryption is designed to keep messages secret. When encryption is used to send messages it is reversible. The sender uses an encryption algorithm to convert the original message called clear text into a coded equivalent called cipher text. At the receiving end the cipher text is decoded back into clear text.

1.1.2 Types of encryption

In the scenario of real world there are different encryption types that are used often to maximize integrity of data in adjacent with each other. The types of encryption are:

i. **Encryption of Secret key:** The symmetric form of encryption is the secret key encryption. The same key which is used to encrypt the data when using the secret key encryption is also used to decrypt the data. This encryption type is the quickest encryption form and is efficient extremely at big data streams encryption such as documents, big conversions, files, etc. The algorithms of secret key are also known as block ciphers, because the encryption is done 1 block at a time (Strebe M, 2006; Mir N F, 2007). The block size is determined by the algorithms. However block ciphers can be reversed potentially because every time at the same time the block cipher will encrypt similar data. So 8 bytes of similar unencrypted block will obtain similar encrypted output each time. To determine the encryption key this output consistency can permit a capable hacker. Through initialization vector usage this problem is solved. The initialization vector is used to encrypt the first stream block and then the first blocks portion which is used to encrypt the second stream blocks. This mingling of blocks will protect similar blocks of 8 byte from developing predictable results of encryption because that block is encrypted every time. In the block stream it is being encrypted with the previous block portion. The secret key encryption downside is that where both the parties require having access to secret key.

ii. **Encryption of Public key:** An asymmetric encryption algorithm is referred to as Public key encryption because the encrypted key is not used to decrypt data. Using public key encryption to transform the data between the parties with the receiver of public key the sender will encrypt the data. Then with the private key the receiver can decrypt the data. The converse form is also applied for sending a message in another direction (Held G, 2004). Over any channel in this way the public key can be transmitted in spite of how the channel might be protected. The below figure shows an example of encryption of public key:

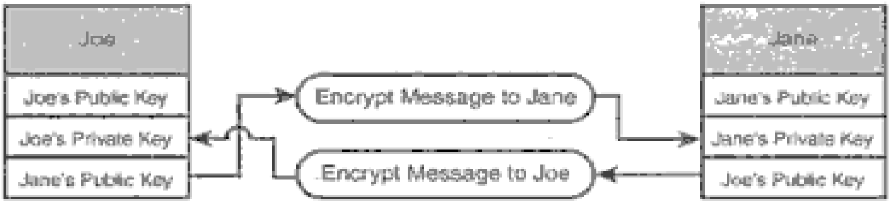

Figure 1: Example of Public Key Encryption

Source: Hoffman K (2006), Microsoft Visual C# 2005 Unleashed, Sams Publishing, New Delhi

Here the problem is that the encryption of public key is somewhat slow and for decrypting and encrypting few numbers of data is optimized. It does not work on huge number of data but the block ciphers works. So after message would be impractical by using encryption of public key to send message.

iii. **Hashing:** A reduced representation of big number of binary data is referred to as hashing. It is decreased down to fixed length bytes when a set of data is hashed. The hash benefits is that when using correct algorithm of hashing it is impossible statistically to have 2 data sources to produce the similar value for hash. For any arbitrarily big data set there is only one distinct value for hash that represents the data set (Sud R and Edelman K, 2004). The data can be detected if the data has been meddled using hashes. If a message is sent to hash of the message or to someone else then the receiver can hash the message on the end using similar algorithm of hashing and compare it against the hash which is transmitted. Since if the hashes did not match then the message must be modified when it was transferred. This is not secure because if anybody deflects original message then they can simply rehash the changed message and send it to the recipient who must monitor the unmodified message. In originality usually the hash is encrypted using Public Key Cryptography Standards so that the receiver can decrypt the hash using private key.

iv. **Digital signatures:** A human signature is evidenced aside forgery that the document was signed by a person who claims to be originator. The same can be said for a document or an electronic message. With a digital signature the person can sign that message which serves as proof for one who developed the document. The first thing done by a person when the person signs a message digitally is to develop the hash of the message. Using private key this hash is encrypted. This permits anyone with their public key to decrypt the hash and rehash the message on machine and compare the two messages to monitor the authenticity of the signature. Using their private key the person can also encrypt the hash because they do not want to hesitate people's list who verifies the signature authenticity. Therefore the person must be able to verify any messages authenticity sent by people for whom they have public keys (Stallings W, 2007).

1.2 Review of Literature

1.2.1 What is distributed encryption?

Distributed encryption is a special form of folder or file encryption where the keys are managed centrally with the engine of distributed encryption. It is used to encrypt folders or files for individuals or groups that moves around various systems. There are a bundle of various technical approaches, but the product is on the system as long as the person is using, and has access to the major server, and the person doesn't need to manage the keys manually. To encrypt the file or folder the person can right-click the file and chose the key/group they would like to use. In distributed encryption the options include BitArmor, Vormetric, WInMAgic, Utimaco and PGP (Website, Securosis.com).

1.2.2 Distributed encryption using JPVM

Java parallel virtual machine is a PVM like library of object classes implemented purely in Java to achieve portability. The major goal is to enable a system to utilize the available resources of computing in a heterogeneous system. It permits explicit passing of message parallel programming in Java. However programs written from java parallel virtual machine cannot be ported to Java virtual machine. Experiments were concluded to measure the development of communications and tasks overhead (Bode A, Dongarra J, Ludwig T and Sunderam V, 1996). The communication and task development overhead is greater which implied that Java Parallel Virtual Machine is most suitable for coarse grain parallelization. The below figure shows the Java Parallel Virtual Machine:

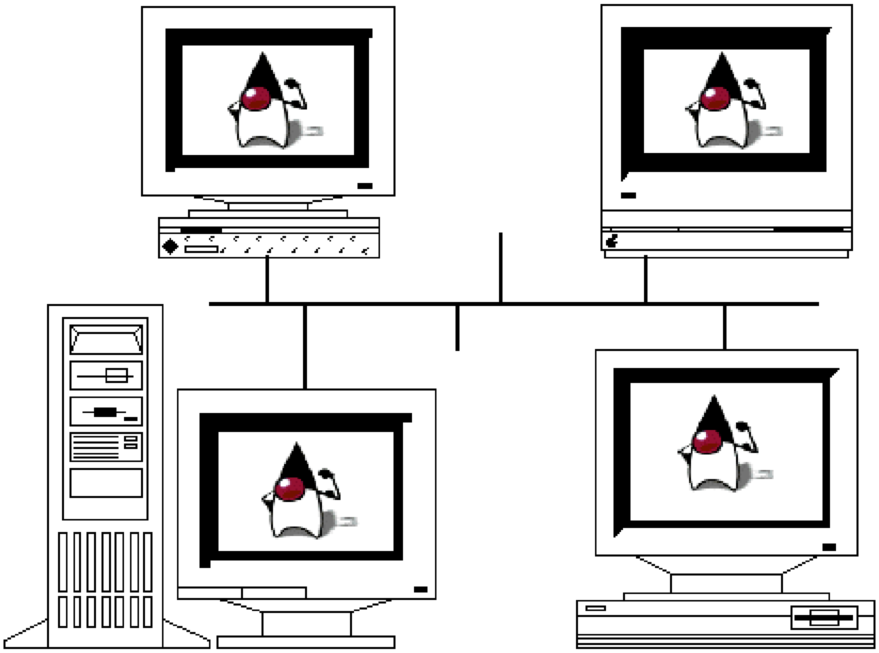

Figure 1: Distributed Encryption using JPVM

Source: Website, cs.virginia.edu

The Java Parallel Virtual Machine environment serves as an intermediate layer between the slaves and the master. The master sends the messages to the slaves by its threads ordering to the Java Parallel Virtual Machine demons of the slaves to initialize the correct analysis procedures. In the next Java Parallel Virtual Machine message input data for these procedures are sent (Royo J D and Hasegawa G, 2005). The result of slaves work is a Java Parallel Virtual Machine message with the description of terminal markings or a message describing errors (Karatkevich A, 2007; Desprez F, 2007).

The Java Parallel Virtual Machine is somewhat different from tk PVM and Java PVM in that it is not a Parallel virtual machine interface for another language but rather another implementation of Parallel Virtual Machine itself. In fact Java Parallel Virtual Machine is not actually a full implementation of Parallel Virtual Machine but actually a Parallel Virtual Machine like library (Wyrzykowski R, 2002). The main disadvantage of this approach is that Java Parallel Virtual Machine programs cannot interact with Parallel Virtual Machine programs written using standard Parallel Virtual Machine. The advantage of this system is that it is written entirely in Java and thus has some hope of retaining the safety and portability promised by the Java language. Distributed encryption using JPVM is an approach by which a system is developed using Java Parallel Virtual Machine for performing Distributed encryption operation.

1.3 Significance of the study

While the existing systems for performing distributed encryption operation are based either of MPI or PVM, this study intends to find a new approach of performing the distributed encryption operation in a better way by implementing it in Java Parallel Virtual Machine, there by making the system more scalable, fault tolerant, flexible and efficient.

1.4 Aims and Objectives of the Study

1.4.1 Aim of the study:

This study aims at using Java Parallel Virtual Machine in developing a system that would perform distributed encryption.

1.4.2 Primary objective of the study

To implement distributed encryption technology using Java Parallel Virtual Machine

1.4.3 Secondary objectives of the study:

i. To study the basic concepts and types of encryption
ii. To understand the concept of Java Parallel Virtual Machine
iii. To analyze in detail the process of distributed encryption
iv. To propose a system that would perform distributed encryption using JPVM

1.5 Research Methodology

A Research methodology is a framework or blueprint for conducting the marketing research project. It details the procedures necessary for obtaining the information needed to structure or solve marketing research problems. Although a broad approach to the problem has already been developed, the research methodology specifies the details of implementing the approach. This research makes use of qualitative approach to analyze the primary data collected.

1.5.1 Sources of Data

People who use the newly developed system are the primary sources of this research. Online Web Portals, news articles, magazines, information portals, books and weekly issues are the major secondary sources for the study.

1.5.2 Research Design

A research design is the planning and executing a research project from identifying the problem through to reporting and publishing the results. By contrast at its most specific level the design of the study refers to the way a researcher guards against and tries to rule out alternative interpretation of results. Pennink B and Jonker J (2010) says that the idea of situating the researcher in the empirical world illuminates the way in which research design can be thought of as providing a bridge between conceptualizing and operationalizing research. Research design is about making a basic plan for a research project. According to Bordens (2006) a research design establishes the general framework of a study addressing each phase of the investigative process. Expert researchers design their studies and then implement these designs with flexibility as they respond to situations that arise as their projects progress. Thus a research design specifies the logical structure of a research project and the plan will be followed in its execution. This study makes use of ***descriptive research***.

1.5.3 Tools of Data Collection & Analysis

Both primary and secondary data are involved in the study.

Primary data is to be collected through questionnaires. A total of 10 respondents would be interviewed to share their experiences about the performance of the new system after deploying the same.

Secondary data in this research is collected by referring the research papers related to the proposed topic that are already in existence and available in the form of published or unpublished reports. For the purpose of this study several secondary sources such as Online Web Portals, news articles, magazines, information portals, books and weekly issues are to be referred.

1.5.4 Shortcomings in existing system

The systems existing at present for distributed encryption work with either Message Passing Interface (MPI) or Parallel Virtual Machine (PVM) technologies or combine the concepts of both. However, these two technologies have certain limitations. Increasing the scalability of the system through new resource addition is dynamic in case of PVM, however, not dynamic in case of MPI (Elts Ekaterina, 2004). Another flaw is that Fault tolerance one of the important features needed for distributed computing is supported by PVM and not by MPI (Lusk E and Gropp W, 2002). One of the major drawbacks of the existing system is that it does not support interoperability since MPI implementations are not interoperable.

1.6 Outcome of the study

The present study intends to create a system using Java Parallel Virtual Machine to perform distributed encryption that overcomes all the flaws in existing system thereby providing scalability, interoperability and fault tolerance.

1.7 Significance of the proposed system

Java Parallel Virtual Machine is an implementation of Parallel Virtual Machine in Java. Java Parallel Virtual Machine can work on a Heterogeneous environment which is one of the greatest strengths for distributed computing. As JPVM is developed in java, it can be used in any Operating System unlike MPI or PVM that are platform dependant. This system is easy to maintain and it is possible for the user to tune its performance by adding or removing resources dynamically.

1.8 Innovation in the proposed system

While the present systems support only manual resource addition this system provides configurable resource management for performing encryption in distributed environment. This offers better maintainability due to automatic generation of alert mails in case of a failure of a process. Another innovative feature of this system is that it has a plug and play facility for using encryption algorithms.

1.9 Benefits of the proposed system

i. This system is made up a complex resource intensive processor.

ii. Encryption algorithm implementations can be easily done by distributing the task to several clusters of nodes. This reduces time of execution.

iii. In general if a large text file is encrypted in single node it will take a long time and block other processes. Using JPVM makes it easy to encrypt large text files by splitting data and sending them to different nodes of cluster and achieve greater performance.

1.10 Limitations of the proposed system

This system has only one major limitation. It is tough to decide the number of nodes needed for executing particular algorithm in distributed manner at first shot itself. Hence it is required to configure, every time, the number of nodes needed based on observations for achieving better performance.

References

[1] Lehtinen R, Russell D and Gangemi G T (2006), Computer Security Basics, O'Reiley Media Inc., USA.
[2] Strebe M (2006), Network Security Foundations, John Wiley & Sons, USA.
[3] Mir N F (2007), Computer and Communication Networks, Pearson Education, New Delhi.
[4] Held G (2004), Virtual private networking: a construction, operation and utilization guide, John Wiley & Sons, USA.
[5] Sud R and Edelman K (2004), SECUR Exam Cram 2 (Exam Cram 642-501), Que Publishing, USA.
[6] Stallings W (2007), Business Data Communications, Pearson Education, New Delhi.
[7] Karatkevich A (2007), Dynamic analysis of Petri net-based discrete systems, Springer, USA.
[8] Desprez F (2007), Algorithms and tools for parallel computing on heterogeneous clusters, Nova Publishers, New York.
[9] Bode A, Dongarra J, Ludwig T and Sunderam V (1996), Parallel virtual machine, EuroPVM '96: third European PVM conference, Munich, Germany, October, 1996 proceedings, Springer, Germany.
[10] Royo J D and Hasegawa G (2005), Management of multimedia networks and services, Springer, Germany.
[11] Wyrzykowski R (2002), Parallel processing and applied mathematics, Springer, Germany.
[12] Pennink B and Jonker J (2010), The Essence of Research Methodology, Springer, New York.
[13] Bordens (2006), Research Design and Methods, Tata McGraw Hill, New York.
[14] Lusk E and Gropp W (2002),Goals Guiding Design: PVM and MPI, Mathematics and Computer Science Division, Argonne National Laboratory.
[15] Elts Ekaterina (2004), Comparative analysis of PVM and MPI for the development of physical applications on parallel clusters, Saint-Petersburg State University.
[16] Website Available at http://securosis.com/tag/distributed+encryption, accessed on 27th August 27, 2011.