International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-4 Number-1 Issue-14 March-2014

Securing Cloud Infrastructure for High Performance Scientific Computations Using Cryptographic Techniques

G K Patra[1], Nilotpal Chakraborty[2]

Abstract

In today's scenario, a large scale of engineering and scientific applications requires high performance computation power in order to simulate various models. Scientific and Engineering models such as Climate Modeling, Weather Forecasting, Large Scale Ocean Modeling, Cyclone Prediction etc require parallel processing of data on high performance computing infrastructure. With the rise of cloud computing, it would be great if such high performance computations can be provided as a service to the users, reducing the costs and overhead in maintaining such a system. But before such a high performance cloud computing model can be introduced, it is very important to address the security issues and challenges to ensure its long term sustainability. Proper security assessments and techniques must be present before such a system can be implemented.

Keywords

High Performance Computing, Cloud Computing, Cloud Security, Functional Encryption, Fully Homomorphic encryption

1. Introduction

High Performance computing is a generic term that refers to the computing infrastructures possessing large number of processors with the capability to perform parallel operations to achieve efficient and high throughput that would have been impossible with normal desktop computers [1]. In the earlier days, the term high performance computing generally used to refer to supercomputers, which were thought of monsterly large complicated machines used to solve problems no one really understood. But in today's world, High Performance Computing can eventually differ from just a supercomputer.

Manuscript received January 10, 2014.

G K Patra, Principal Scientist, CSIR- Fourth Paradigm Institute, Council of Scientific & Industrial Research, Bangalore, India.

Nilotpal Chakraborty, M.Tech student, School of Future Studies and Planning, Devi Ahilya University, Indore, India.

Though HPC is using supercomputers, they can be built to use the same hardware and infrastructure found in web servers and even desktop workstations. Whereas supercomputers were meant to solve problems of high complexity only, HPC apart from solving complex problems can be used for small and medium sized problems as well. They perform it with the help of what is known as clusters of small computers. Each individual computer is configured to form clusters, having one to four processors with each processor typically having two to four cores.

The primary motivation behind HPC is that individual small computers or clusters can work together efficiently to solve problems that a single computer cannot do [2]. There has to have an efficient communication system among the nodes or clusters and a job assignment module that can assign specific jobs to the individual nodes. Thus HPC amalgamates the computing power of various computers in order to form a larger one that can be used to solve large problems by assembling the computing power of individual nodes. In today's scenario, a large number of scientific and engineering applications such as climate modeling, environmental modeling, ocean modeling, weather prediction etc need to produce results very efficiently. HPC, with the help of parallel processing capability, helps such models to simulate and obtain results in a short amount of time.

Though HPC seems to be a vibrant technology, its implementation is easier said than done. Developing an HPC infrastructure requires huge financial investments along with potential manpower to maintain the system. In CSIR-Fourth Paradigm Institute (earlier Centre of Mathematical Model and Computer Simulation, Bangalore), we have one of the World's largest supercomputer (ranked 3rd in India, 99 in the World [3]) Cluster Platform 3000 BL460c Gen 8, Xeon E5-2670 8C 2.60GHz. It consists of 1088 nodes with each node having two eight core processors and 64GB of primary memory. For maintaining the temperature of the system, there is dedicated water cooling system and dedicated technically skilled manpower employed to look after the system. Developing and maintaining such a dedicated HPC system would not be possible for

many of the organizations, who nonetheless may have the requirement of using HPC. Thus there is a need to reach out to those individuals and organizations and provide them the HPC services just like any other services provided through the cloud computing.

Cloud computing is a computational model that provides a shared configurable pool of computing resources that can be accessed in a convenient way through the internet. Users of cloud based services pay only as per their usage of resources provided by Cloud Service Providers (CSP) and they need not worry about the underlying infrastructure of the system. Cloud computing introduces a cost effective, multi-tenant, ubiquitous mode of computing that hides all the internal complexities from the user and providing efficient services. Users use them, just like using the bank ATM machine, without need to understand or worry about how it is actually functioning. But as the users are excited about cloud because of its enormous features and benefits, they are also skeptical about the security. As the users have no control over the functioning of cloud, the security of it remains to a big question.

This paper is organized in the following way: Section 2 discusses the basic notions of Cloud computing and its security challenges, section 3 talks about Security Measurements undertaken in Cloud based environment, Section 4 discusses the existing cryptographic techniques for mitigating security issues in cloud. In section 5, we talk about our contribution and research work performed on securing cloud infrastructure.

2. Cloud Computing and its Security Challenges

Cloud computing has been one of the buzz word around for the last couple of years in the world of computing. It has been concretely defined by NIST as in [4]—

"It is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction".

Today, cloud computing, with the help of its various service models viz Software as a Service, Platform as a Service, and Infrastructure as a Service, offers a wide range of computing services in a convenient, on-demand and multi-tenant way to a number of users concurrently. The Cloud architecture appears to be a black box to the user, meaning the users have no idea about where exactly the data resides or how the applications are actually executing. All the internal functioning and complexities are hidden from the general users, performing computation in an easy and less overhead manner. Cloud computing provides a number computing resources such as storage, applications, data, platforms in a leased basis. Users make use of the resources and pay according to the usage.

The model that is used for various services in Cloud computing can be devised to provide HPC services as well. Buying and installing an HPC architecture posses a great deal of overhead. It needs a huge amount of investment to install a supercomputer and its maintenance also requires dedicated skilled manpower. For any small or medium scale organization, it would impose great difficulty in maintaining such a huge system. Moreover, the system may not be utilized properly as they might have a requirement for a small fraction of time. Such failure to utilize the resources can be addressed if all these services can be provided through the cloud. Users would get their own login credentials to access the HPC services, submit their job over the console interface and the job would then be submitted to the HPC by the cloud. After processing, cloud would get back the result to the user. In this way, any user can access the HPC services for their models, without need to worry about the infrastructure and overhead to maintain the system in-house.

Though cloud offers such a beautiful and elegant model of computing, there is a sheer need to address all the issues and challenging before migrating the system to the cloud, such as security. Security threats possess great difficulty in the growth of cloud computing. Users are excited because of the enormous advantages they can make out of cloud computing but are skeptical too due to the security issues. Cloud appears to be a black box to any user and he is completely unaware of the security measurements being taken inside the system. Data over the cloud is actually stored on a remote server may be very far from the user and any security attack on the remote system can also be vulnerable to the privacy and secrecy of user data. Before the local systems can fully be migrated to the cloud, all such issues need to be taken care of efficiently in order to gain users trust and providing reliable computing infrastructure. Once its security issues are properly addressed and mitigated, it will eventually start gaining trust of users and thus would grow enormously.

3. Security Threats in Cloud Computing

Cloud computing promises to help organizations and their IT departments are more agile, efficient, and able to cost-effectively deliver new services that enable their businesses to thrive. But the promise of the cloud cannot be fulfilled until IT professionals have more confidence in the security and safety of the cloud. We know that IT concerns with cloud computing security are major barriers to business adoption of the cloud. But before the IT industry can address these concerns, better understanding of them is required.

Many security and privacy threats, such as malware or the risk of a malicious insider, appear to be omnipresent aspects the information technology landscape today, and must be addressed as part of a larger national and international cyber security agenda. The security challenges faced by organizations wishing to use cloud services are not radically different from the traditional security issues and threats. The same internal and external threats are present and require proper risk mitigation and disastrous management policies in order to protect privacy and security.

To identify the top most security threats impending in cloud computing, Cloud Security Alliance conducted a survey of industry experts to compile professional opinion on the greatest vulnerabilities within cloud computing. In this most recent edition of this report, experts identified the following nine critical threats to cloud security [5] (ranked in order of severity)—

- *Data Breaches*: A data breach is a security incident in which an unauthorized person accesses and modifies sensitive, protected or confidential data. The cloud service provider must ensure the best level of security authentication and authorization procedure to ensure data protection. If a multitenant cloud service database is not properly designed, a flaw in one client's application could allow an attacker access not only to that client's data, but every other client's data as well.
- *Data Loss*: For both consumers and businesses, the prospect of permanently losing one's data is terrifying. Of course, data stored in the cloud can be lost due to reasons other than malicious attackers. Any accidental deletion by the cloud service provider, or worse, a physical catastrophe such as a fire or earthquake, could lead to the permanent loss of customers' data unless the provider takes adequate measures to backup data. Furthermore, the burden of avoiding data loss does not fall solely on the provider's shoulders. If a customer encrypts his or her data before uploading it to the cloud, but loses the encryption key, the data will be lost as well.
- *Account Hijacking*: Account hijacking is not a new threat to computing. It is a type of identity theft in which the attacker uses stolen account information to carry out malicious or unauthorized activities. Typically account hijacking is carried our through phishing, sending spoofed emails to the user, password guessing or a number of other hacking techniques. In many cases, an email account is linked to a person's social networks and financial networks etc. and by impersonating the account; a hacker can gain access to these confidential data for illegitimate activity.
- *Insecure APIs*: Users of cloud services access their data through some interfaces as provided by the service providers. The security and availability of general cloud services is dependent upon the security of these basic APIs. From authentication and access control to encryption and activity monitoring, these interfaces must be designed to protect against both accidental and malicious attempts to circumvent policy.
- *Denial of Service*: Denial of Service (DOS) attacks are nothing new and they've been a thorn in the sides of data center managers and IT staff for more than a decade now. Through DOS, a hacker doesn't need to attack the entire infrastructure anymore. They can simply choose the most resource intensive app that the user is running on the cloud and use simple low band width attacks to take out that service.
- *Malicious Insiders*: Cloud computing as a process is governed, managed, and maintained by site administrators. By default, they hold the key to managing all the data, files and privileged company resources and files. These administrators sometimes because of some personal differences can leak out the important data of a client or can distribute confidential financial or official data of the organization.
- *Abuse of Cloud Services:* One of cloud computing greatest benefits is that it allows even small organizations access to vast

amounts of computing power. Any organization irrespective of its size, can rent computing resources based on its needs and requirements. However, not everyone wants to use this power for good. Using the enormous computational resources provided by the cloud, an attacker can attack a high performance machine in a convenient amount of time, which otherwise would have taken years to accomplish.

- *Insufficient Due Diligence*: ever since its introduction, the demand for cloud computing has always been increasing and it is mainly due to its promising cost cutting technology and increasing operational efficiency. But without fully understanding the CSP functioning and environment, migrating to cloud would lead to chaos for the organizations.
- *Shared Technology Issues*: Cloud service providers deliver their services in a scalable way by sharing infrastructure, platforms, and applications. A defensive in-depth strategy is recommended and should include compute, storage, network, application and user security enforcement, and monitoring, whether the service model is IaaS, PaaS, or SaaS. The key is that a single vulnerability or wrong configuration can lead to a compromise across an entire provider's cloud.

4. Techniques to Ensure Secured Cloud HPC Services

Security and privacy have always been a challenge and is of primary concern with the computing resources and with the invention of cloud computing, the need to secure the data stored in the cloud becomes much more challenging. As cloud appears to be a black box, the user of the cloud is completely unaware of the security of the data residing in the cloud servers. Moreover, what if the cloud data centers are under attack? Whether any potential attack on the cloud service provider's systems would cause adverse affect on user data? Thus we can understand that there are a number of security issues to be addressed before any organization or individual migrate to cloud environment.

Data security over the cloud can eventually be addressed if cryptographic schemes are implemented [6]. But before traditional ciphers converts a plain text into a piece of encrypted data that is of no use unless decrypted. But this actually does not help in gaining the full benefits of cloud computing, as cloud is not meant for storage purpose only; it is also used for performing computations on the stored data. Thus we need schemes that allow performing computation in the encrypted data. Such a scheme is known as fully homomorphic encryption scheme that allows arbitrary number of operations to be carried out on the encrypted data.

Researchers around the world have spent many years into devising a cryptographic technique that can eventually allow various operations to be performed on the cipher texts, have failed to obtain one. In 2009, Craig Gentry, an IBM researcher have shown the first possible fully homomorphic encryption construction that can evaluate arbitrary operations [7]. Gentry first obtained a "somewhat homomorphic" scheme, supporting only a limited number of cipher text multiplications due to the fact that cipher text contains a certain amount of noise which increases with every multiplication and that decryption fails when noise size passes a certain bound. As a result, in the somewhat homomorphic scheme, the functions that can be homomorphically evaluated on cipher texts are polynomials of small, bounded degree. The second step in Gentry's framework consists of squashing the decryption procedure so that it can be expressed as a low degree polynomial in the bits of the cipher text and the secret key. Then, Gentry's key idea, called bootstrapping, is to evaluate this decryption polynomial not on the cipher text bits and the secret-key bits (which would yield the plaintext), but homomorphically on the encryption of those bits, which gives another cipher text of the same plaintext. If the degree of the decryption polynomial is small enough, the noise in the new cipher text can become smaller than it was the original cipher text, so that this new cipher text can be used again in a subsequent homomorphic operation (either addition or multiplication). Using this cipher text refresh procedure the number of permissible homomorphic operations becomes unlimited and one obtains a fully homomorphic encryption scheme.

Though Gentry's work has opened the possibility of fully homomorphic encryption theoretically, implementing such a scheme in practice following his work is very inefficient due to the time it takes to evaluate homomorphically. For the past five years, researchers are working on this and a number of FHE variants have been introduced [8] [9] [10] [11]. But it is still in its infancy to be implemented in real life scenario. Obtaining a practical efficient FHE scheme

is an open challenge. The pictorial description of FHE is shown in figure-1.

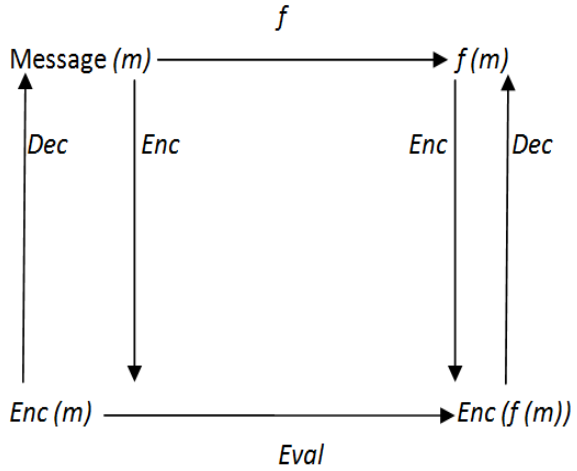

Figure 1: Fully Homomorphic Encryption

Along with FHE, recently researchers have come with another notion of cryptographic scheme called Functional Encryption [12]. In traditional public-key cryptosystems, the corresponding cipher text of a plain text is intended to be decrypted only by a single recipient of the encrypted data. But in certain scenario, such as cloud computing, an encrypted message may be directed to a group of people, without knowing a specific individual. In such a communication system, functional encryption lets a user to decrypt only a specific functionality of the cipher text, without revealing any more information about the original data. Thus it introduces a new form of cryptographic encapsulation on the cipher texts. Functional encryption may be depicted by figure-2.

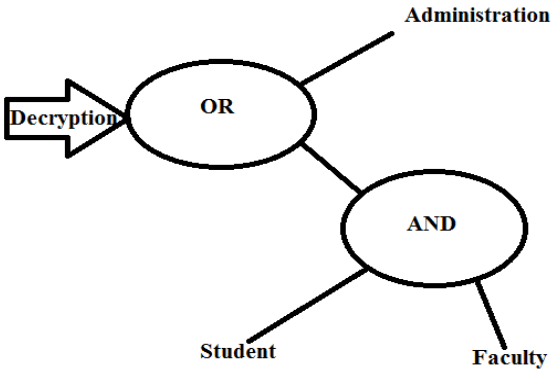

Figure 2: Functional Encryption

5. Our Contribution

CSIR Fourth Paradigm Institute, Bangalore has been one of the leading organizations in the country for high performance computing, having the 3rd largest supercomputer in India. Such infrastructure is being used regularly for various scientific and engineering problems by various Indian scientists. But maintaining such a system has great difficulty as it incurs a huge amount of financial investment along with dedicated manpower. Because of this high overhead, it has been observed that many organizations with small budget cannot take the advantage of such heavy computing strength. Thus there's a need to provide such services over to them in an efficient and cost-effective manner.

Through this work, we propose the introduction of cloud computing in providing high performance computing infrastructure so as to reach people who are otherwise deprived of using HPC. As per their requirements, users will submit their jobs through the web based cloud terminal available to them and cloud then submits the job to HPC for computations. After the computation is over, the same is handed back to the user through cloud interface. With the current ongoing developments on Cloud Computing and the available infrastructures, it is not hard to implement such an HPC service in practice, although the security implications of such systems have to be properly addressed.

HPC services provided via Cloud computing would, if introduced, would require to mandate the security issues and threats for a reliable and efficient computing environment. As part of this project work, we have performed research work on computing in an encrypted domain where we have come across various advanced cryptographic techniques such as fully homomorphic encryption and functional encryption. We have started the work with surveys on fully homomorphic encryption and have analyzed Gentry's scheme extensively to understand the security and inefficiency of the scheme. We have observed that due to the bootstrapping stage for reducing noise associated with the cipher text, it is very slow to actually process the whole scheme. But again, this noise reduction is necessary in order to allow arbitrary number of operations to be performed on the encrypted data. Thus there's a need to look forward to some variant scheme.

As part of our FHE implementation, we have followed the FHE scheme proposed by vanDijk et. al. [8] and have implemented the same in C programming. The security of this scheme is based on approximate GCD based problem, which states that given a nearest multiple of a prime number with some noise, it is difficult to get back the prime number. The scheme proposes the symmetric key

version of FHE which we have converted to asymmetric key as well to let anonymous users to carry out computations on the cipher texts, having access to the correct public key. We have also identified some of the applications that require no arbitrary number of computations, and where the noise can be controlled up to a certain limit always. Currently we are working on devising such a scheme that can handle the noise in an efficient way without affecting the running time of the encryption/ decryption or evaluation procedures. We are focusing on the FHE construction that is based on Learning with Errors (LWE) that requires no bootstrapping and squashing the decryption circuit, making the whole encryption/ decryption and Homomorphic operation process, a complex-free one.

With the target of securing cloud infrastructure for HPC services, we have also worked on functional encryption, which provides a fine-grained control of access to encrypted data [13]. It is a public key cryptographic scheme that provides restricted keys to users, with the help of which they can only perform some specific computations on the cipher texts. The underlying infrastructure of functional encryption scheme is based on fully homomorphic encryption and existing schemes such as Attribute based encryption and Identity based encryption have been categorized as specific cases of functional encryption. As part of our study and implementation, we have followed the scheme as in [14] and have implemented the same in C programming. The scheme generates specific private keys based on the policy imposed, and lets the user to decrypt any file only when the private key suffices the decryption policy. This scheme will help scientists in securing their model outputs and specifying a set of people to whom the same can be shared.

6. Conclusion and Future Work

In today's era, high performance computing is an integral part of any research and development oriented organizations and activities. Various scientific and engineering problems require programs to run in parallel for faster and efficient results, which in turn, requires Super Computers. But due to the high cost and maintenance overhead, installing an HPC infrastructure becomes very difficult. Cloud computing can be one of the enabler of providing such HPC services over the web, that can easily reach out to a wide range of users. It would be very efficient and cost effective and less overhead for a user as he is now free of maintaining the system on his own. But with the introduction of cloud, the security of such services needs to be taken care of. Security threats and challenges have affected the growth of cloud computing and they need to be properly addressed. With the advent cryptographic techniques such as fully homomorphic encryption and functional encryption, it has now become possible to mitigate all such security problems over the cloud. But efficiency of such schemes remains to be an open question. Optimizing the existing FHE schemes and devising a scheme that would be efficient enough to be practical remain to be a future work.

Acknowledgment

Nilotpal is thankful to the SPARK program of CSIR Fourth Paradigm Institute, Bangalore for allowing him to carry out his final year project at the institute. This work is partially supported by the project ARiEES, funded by CSIR, India, under the 12th Five Year Plan.

References

[1] D Eadline; High Performance Computing for Dummies, Sun and AMD Special Edition, Wiley Publishing Inc, 2009.

[2] V Eijkhout; Introduction to High Performance ScientificComputing,2011,Availableonline: http://www.tacc.utexas.edu/~eijkhout/istc/istc.htm

[3] List of top 500 super computers, November 2013 survey: www.top500.org/list/2013/11/?page=1, accessed in January, 2014.

[4] Peter Mell, Timothy Grance; The NIST Definition of Cloud Computing; NIST Special Publication 800-145, 2011.

[5] "The Notorious Nine: Cloud Computing Top Threats in 2013", Top Threats Working Group, Cloud Security Alliance, February 2013.

[6] G K Patra, Nilotpal Chakraborty; Securing Cloud Computing Environment with the help of Fully Homomorphic Encryption, Journal of Computer Technology & Applications, ISSN: 2229-6964, Vol 4 Issue 3, December 2013.

[7] C. Gentry. Fully homomorphic encryption using ideal lattices. In *STOC '09*, pages 169-178, ACM, 2009.

[8] Marten van Dijk, Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. Fully homomorphic encryption over the integers. In Advances in Cryptology - EUROCRYPT'10, volume 6110 of Lecture Notes in Computer Science, pages 24-43. Springer, 2010.

[9] N. P. Smart and F. Vercauteren. Fully homomorphic encryption with relatively small

key and ciphertext sizes. In *Public Key Cryptography - PKC'10*, volume 6056 of *Lecture Notes in Computer Science*, pages 420-443. Springer, 2010.

[10] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. Fully homomorphic encryption without bootstrapping. In Innovations in Theoretical Computer Science (ITCS'12), 2012.

[11] JS Caron, T Lepoint, M Tibouchi; Batch Fully Homomorphic Encryption over the Integers, Cryptology ePrint Archive, Report 2013/036, 2013.

[12] Dan Boneh, Amit Sahai, Brent Waters, Functional Encryption: Definitions and Challenges, Proceedings of Cryptography Conference, 2011.

[13] S. Agarwal, S Gorbunov, V. Vaikuntanathan, H. Wee; Functional Encryption: New Perspective and Lower Bounds, Cryptology ePrint Archive, Report 2012/468, 2012.

[14] S. Agarwal, S. Agarwal, S. Badrinarayanan, A. Kumarasubramanian, M. Prabhakaran, A. Sahai; Function Private Functional Encryption and Property Preserving Encryption: New Definitions and Positive Results,Cryptology ePrint Archive, Report 2013/744, 2013.

G K Patra received his Doctoral degree and Masters Degree in Electronic Sciences, from Berhampur University, Orissa in 2002 and 1994 respectively. Since 1997, he has been working as a scientist in CSIR Centre for Mathematical Modelling and Computer simulations (renamed now as CSIR Fourth Paradigm Institute). In addition he is also an Associate Professor at Academy of Scientific and Innovative Research (AcSIR). His current interest includes cryptography and High Performance Computing.

Nilotpal Chakraborty received his B.Tech degree in Information Technology from Assam University, Silchar. He is currently pursuing M.Tech in Systems Management in Devi Ahilya University, Indore and doing his final year project at CSIR Centre for Mathematical Modelling and Computer simulations (renamed now as CSIR Fourth Paradigm Institute), Bangalore. His current research area includes Cryptography, Information Security, Cloud Computing and Big Data.