Securing the Clouds

Mohammed Imtiaz Ahmed **Mohammed Bakhtawar Ahmed** **Debojit Das**

Abstract

Cloud computing is all the rage. "It's become the phrase du jour," says Gartner senior analyst Ben Pring [16]. While some consider it as a future derived technology that will not only help organisations to gain more profit but will also have a positive impact on the environment, others consider it as a past derived technology which is a refined version of Timesharing model from late 1960's and nothing new. Whatever the words be, Cloud computing has gained a lot of attention in the recent years and has a wide scope in the IT industry and the big IT giants are all set to provide cloud services to the clients, as they require. Beside its wide possibilities, one threat that cloud faces is of its security issues. As in cloud computing, the data resides in the third party data centers, and the data is always vulnerable to attacks and changes. The changes or modification of data can be in from of a third party attack where a unauthorised person makes unwanted modification in the data or it can be a insider attack where an authorised person makes unauthorised changes in data. In this paper we identify some areas where cryptography can help a rapid adoption of cloud computing. Although secure storage has already captured the attention of many cloud providers, offering a higher level of protection for their customer's data, we think that more advanced techniques such as searchable encryption will become popular in the near future, opening the doors of the Cloud to customers with higher security requirements.

Keywords: Cloud Computing, Searchable Encryption, Secure Storage, Cloud Computing - Security

1. Introduction

"*Cloud computing is a model for enabling suitable, on-demand network access to a shared pool of organize able computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.*" [15]

Ref: (Lee Badger; High priority requirement to further USG agency cloud computing adaptation, NIST, November 2011)

2. What Cloud Computing isn't:

Even though cloud computing can incorporate some computing paradigm such as Utility computing, Grid computing, Autonomic computing, or Platform virtualisation; it is not synonymous to them. For example cloud computing is not the same as utility computing, as it does not always employ the metered service pricing of utility computing. Cloud computing is not similar to Grid computing, as unlike grid computing which focuses all its computing resources in one single large task, cloud computing provides a distributed

virtual machine over a network. Although cloud computing is sometimes referred to as client server architecture in which the cloud serves as the server, but unlike traditional server which is a specific machine on a specific location, computation running on cloud are split anywhere running on different machines.

3. Alternative Views

A number of prominent people view cloud computing as a total hype and a past derived technology. In an online video (http:/www.techcentral.ie/article.aspx?id=13775), oracle CEO Larry Ellison bluntly stated, "What the hell is cloud computing? ... When I read articles on cloud computing, it is pure idiocy… when is this idiocy going to stop?"

Information n security expert Bruce Schneier, in his June 4, 2009 online news letter Schneier on security (www.schneier.com/blog/archives/2009/06/cloud_computing.html), says "this year's overhyped IT concept is Cloud computing ... but, hype aside, cloud computing is nothing new. It's modern version of time sharing model from 1960's. It's what Gmail and hotmail has been doing for so many years, and IT outsourcing, network infrastructure, security monitoring is a form of cloud computing"

Another interesting article on cloud computing on Information week article titled "HP on cloud: The World is Cleaving in two" (http:/informationweek.com/news/services/business/showAricle.jhtml?articleid= 213422906), Russ Daniels from Hewllet Packard said, "The idea that we are one day going to throw a switch and move everything out to one small number of data centers, located next to low cost power source is nonsensical. It's not going to happen. Cloud computing is not the end of IT" [14].

4. Cloud Computing Security Fundamentals

Security is a prime concern when entrusting organisations' critical information to geographically dispread cloud platform is not under the direct control of the organisation. Three factors that support information assurance are confidentiality, integrity, and availability.

Confidentiality: It refers to the prevention of intentional or unintentional unauthorised disclosure of information. Confidentiality in cloud computing is related to the areas of intellectual property rights, covert channels, traffic analysis, and encryption.

Integrity: The concept of integrity requires the following three principles:

- Modifications are not made to data by unauthorised person.
- Unauthorised modifications are not made to data by authorised person.
- The data is internally and externally consistent.

Availability: It refers to reliable and timely access to data or cloud computing resources by appropriate personnel's. Availability Guarantees that the system will work properly when needed.

5. Security Design Principles

Initially computer codes were not written with security in mind; but because of increased frequency and sophistication of attack against information system, modern software design methodologies include security as prime objective.

In 1974, Saltzer and Schroeder of University of Virginia addressed the protection of information stored in a computer system by focusing on hardware on software issues that are necessary to support information security, which is effective even today. The paper presented following 11 design Principles.

i. Least Privilege
ii. Separation of duties
iii. Defense in depth
iv. Fail safe
v. Economy of mechanism
vi. Complete medication
vii. Open design
viii. Least common mechanism
ix. Physiological acceptability
x. Weakest link
xi. Leveraging existing components

Least Privilege: This principle states that an individual, process, or other type of entity should be given the minimum privilege and resources for the minimum period of time required to complete the task. This method reduces the risk of unauthorised access to the sensitive information.

Separation of Duties: It requires the completion of one specific and sensitive activity or access to sensitive information to depend on satisfaction of plurality of condition. For example an access may require signature of more than two individuals.

Defense in Depth: It is the application of multiple layer protection where in a layer will provide protection if the subsequent layer fails. The Information Assurance Technical Frame work Forum (IATFF) an organisation sponsored by National Security Agency (NSA) has produced a document titled "International Assurance Technical Framework" (IATF) that can be downloaded from (www.niap-ccevs.org/cc-scheme/ITAF_3.1-chapter_03-ISSEP.pdf)

Fail Safe: Fail safe means if a cloud system fails it should fail in such a state that the security of system and its data are not compromised. One implementation of this philosophy would be to define the system default to such a state in which no user or process is provided access to the data.

Economy of Mechanism: Economy of Mechanism promotes simple and comprehensible design and implementation of protection mechanism, so that unintended access path does not exist or can be readily identified and eliminated.

Complete Medication: In Complete medication, every request by a subject to access an object in computer system must undergo a valid and effective authorisation process. The Medication must not be suspended or become capable of bypass.

Open Design: It promotes that any method or procedure that is to be used to secure the system should be made open such that others in the same profession can check its reliability and find loop holes in it. There has always been a debate to whether or not to disclose the security algorithm, in most of the cases it has been found that disclosing an algorithm can help improve its efficiency except for some organisations such as National Security Agency (NSA), which employs some of the world's best Mathematicians and Cryptographers.

Least Common Mechanism: It implies that a minimum number of protection mechanisms should be common to multiple users, as shared access path can be source to unauthorised information exchange.

Physiological Acceptability: It refers to the ease of use and intuitiveness of the user interface that controls and interacts with the cloud access mechanism.

Weakest Link: As an old saying goes "*A chain is only as good as its weakest link*". Similarly, security of cloud system is only as good as its weakest component. Thus, we must identify the weakest mechanism in the security chain and defense layers.

Leveraging Existing Components: In many instances the security component of the system may not work properly. In addition, there may be difficulty in getting configured with other system, thus greatly increasing the security threat of the whole system. Thus, the security mechanism must be checked properly for the optimum working of all the configurations.

6. The Problem Scenario

With the cloud services providing online delivery of services over the internet, security is the prime concern for all the so called potential customers. The wide adoption of cloud computing is raising several concerns about treatment of data in the cloud. Advantages of cloud storage are enormous:

i) Ubiquitous access: anywhere, anyhow, anytime access to your data,

ii) High reliability,

iii) Resilience and

iv) Scalability,

v) Cost efficiency.

But, unfortunately, to date, several security and legal risks should be considered:

i) Unauthorised access,

ii) Sensitive data disclosure,

iii) IPR protection,

iv) Communication threats and loads in transferring data,

v) Data integrity.

As stated in [8][9], cloud data security is the most worrying issue of cloud technology and before enterprises or public authorities will fully outsource their data management to cloud vendors, replacing their internal facilities, security has to improve to an acceptable level. With the confidential and sensitive data stored on the cloud, there's always a risk of unauthorised access to the data. The threat of misuse of the data can either come from an unauthorised user or from any authorised person of the service provider. The main concern around data storage is the protection of information from unauthorised access. In several usage scenarios the risk of data being disclosed, lost, corrupted, or stolen is unacceptable. Until data is stored on resources owned, controlled, and maintained by the data owner, the possibility of unauthorised access is reduced by any physical countermeasure or trust in authentication / authorisation mechanism put in place by him / herself (e.g. physically located in room at the establishment, behind closed doors and installing network firewalls and ACLs at software level).

Things radically change when moving from resources fully controlled by the data owner to resources administrated by third party entities like public clouds. Resources that sit outside the user's domain are resources not owned and not controlled by the user and even trusting the resources' provider, the risk that someone (e.g. an employee of the resources' vendor) can access and disclose/corrupt data is considerable. In the literature this risk is usually known as insider abuse or insider threat [10][11][12]. This is the major risk that, presently, is preventing the large adoption of cloud-based solutions by enterprises. Before companies move their data to the cloud, benefitting from the cloud storage advantages, all issues deriving from storing data on un-owned and un-trusted resources must be addressed, including the inconsistencies with this new model, put there by the legal frameworks.

7. The Possible Solutions

Secure Storage: The secure storage approach aims to avoid insider threats using encryption techniques to protect data from unauthorised access. The core concept of secure storage is the encryption of data in the trusted environment before sending it outside to the un-trusted cloud storage resource. There are a wide range of encryption algorithms at the cutting edge which have been proved to be secure that can be used to perform encryption/decryption operations (e.g. AES, Serpent, Blowfish). Theoretically both symmetric and asymmetric algorithms can be used, but, since the latter are much slower than the former, for performance reasons symmetric algorithms are preferred. The usage of encryption as a technique to secure data guarantees

the confidentiality of data and helps to detect any corruption in data. The main issue in the secure storage approach is the management of encryption keys. In fact, once data is encrypted, keys become the true bits to protect! If keys were stored in the un-trusted environment along with data, an attacker could have at his/her disposal both data and the keys to decrypt the data, with disastrous consequences. Keys are stored on a keystore that can be implemented either on a i) portable device (e.g., an usb pen-drive) owned by the user who can plug it anywhere (within the trusted environment) or ii) in a specialised server which sits somewhere in the trusted environment. Two good examples of commercial secure storage services which encrypt data client-side before transmitting it outside the user's machine (considered reasonably trusted) are:

1. Spideroak; that is a cloud based backup and file synchronisation service which uses client-side encryption and implements a so called zero knowledge system where keys (neither the master passwords nor encryption keys) are never transmitted to the service's provider, and
2. GoldKey; Cloud Storage Security Key with the peculiarity of storing keys on the GoldKey Token: a portable device very similar to an usb pen-drive instead of the user's machine. In contrast, an example of an ineffective secure storage service is Dropbox. It offers its users an online space to store data using Amazon's S3 as a back-end storage service. Users' data is encrypted before being stored on S3, but it has to be considered that Dropbox encrypts data in its servers keeping the keys in its servers. This, in fact, nullifies the added value of encryption since, unless for particular use cases, users consider Dropbox' servers as trusted as S3 servers.

A basic architecture of a secure storage system is presented in fig. 1: when the user wants to store data on an un-trusted resource: i) encrypts data at an encryption point (it may be either the user's machine or a service offered within the trusted environment), ii) encrypted data is sent to the storage service and iii) keys are stored on the keystore. Inverse procedure is applied when data must be accessed: it is transferred from the un-trusted storage to the encryption point, here keys are retrieved, the decryption takes place, and original data is sent back to the user.

Searchable Encryption: An interesting problem related with searchable encryption is Private Information Retrieval (PIR) [1][2]. The problem here is for the user to retrieve information from a database without revealing any information about the requested data to the server. Searchable encryption goes one step further by allowing the user not only to retrieve information privately but also to search it.

The first searchable encryption scheme was defined in 2000. It makes use of symmetric encryption and provides:

- Query isolation for searches
- Controlled searching
- Hidden queries

Those three properties guarantee that the server is not able to learn anything more about the plaintext than the search result, it needs the user's authorisation to search for an arbitrary word, and the user need not reveal the word they are searching for. There are some other symmetric schemes for searchable encryption that enhance security and efficiency [3].

Another approach for searchable encryption is to use asymmetric encryption. The first scheme for searchable encryption that makes use of public key cryptography is the Public-Key Encryption with Keyword Search (PEKS) scheme, proposed by Boneh et al. [4]. This scheme has been enhanced (in [5][6]). More advanced solutions also allow searching with wildcards [7]. Most of those schemes use hidden vector encryption (HVE). We can see HVE as an identity-based encryption where both the encryption and the decryption key are derived from a vector.

Out of the two described solution strategies Searchable encryption can prove to be a more efficient method for providing securities to cloud, as in secure storage the security of the data is totally dependent on the secrecy of the encryption key. If the key is leaked the security of the data will be compromised. Here we propose an encryption mechanism that can be used for searchable encryption through RSA algorithm.

Example: Consider a scenario where the cloud customer wants to store personal data in the third party cloud environment. As the data is at a risk of unauthorised access, one option for the customer to protect the data is to first encrypt the message and then store the data. One problem with this method is that if the customer later wants to retrieve the data to search through any related content, he will be unable to do so as the data stored is no more a plain text and is in encrypted from. To overcome this problem we use the searchable encryption technique. An example of this technique is discussed here. Suppose Alice wants to store some data in the third party cloud, the best way by which she can ensure the data safety is by encrypting the data. First Alice has to find the keys to encrypt the data as follows:

1. Choose two distinct prime numbers p and q.

   o For security purposes, the integers p and q should be chosen at random, and should be of similar bit-length. Prime integers can be efficiently found using a primality test.

2. Compute $n = pq$.

   o n is used as the modulus for both the public and private keys. Its length, usually expressed in bits, is the key length.

3. Compute ö (n) = (p – 1) (q – 1), where ö is Euler's totient function.

4. Choose an integer e such that $1 < e <$ ö (n) and greatest common divisor gcd (e, ö (n)) = 1; i.e., e and ö (n) are coprime. e is released as the public key exponent.

5. e having a short bit-length and small Hamming weight results in more efficient encryption.

6. Determine *d* as $d \equiv e^{-1} \pmod{\varphi(n)}$, i.e., *d* is the multiplicative inverse of *e* (modulo $\varphi(n)$).

Alice can now use her public key (n, e) to encrypt the message and store the message in the third party cloud. The encryption can be done in following way:

$$c \equiv m^e \pmod{n}.$$

Where *C* is the encrypted cipher text and *m* is the actual message.

Once the encrypted data is stored in the cloud, Alice can now implement the searchable encryption method to search through her personal data without enabling the server to know what she is searching for.

For Example: If Alice wants to search a particular content related to data stored in the cloud, she can enable this by first encrypting the data to search using the same public key which was used to encrypt the data stored in the cloud, as the keys will be same, the data to be searched will be encrypted to same cipher text as the actual data. The encrypted data to be searched can then be sent to the server for search process without the knowledge of the server as to what it is searching for. Once the search is complete the server will send this data to Alice who can then decrypt the search content using the following calculation:

$$m \equiv c^d \pmod{n}.$$

Where *m* is the actual message and *C* is the cipher text.

8. Conclusion

We have presented in this paper the recent advances in crypto that we foresee will add a new layer of security to Cloud Computing and boost its adoption. As we have shown in the paper, most cryptographic primitives are ready to be used. We only need to convince Cloud Providers to implement them or produce efficient implementations that could ease its inclusion in open source Cloud Computing platforms.

References

1. Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. 1998. Private information retrieval. J. ACM 45, 6 (November 1998), 965-981.

2. Kushilevitz, E.; Ostrovsky, R., "Replication is not needed: single database, computationally-private information retrieval," Foundations of Computer Science, 1997. Proceedings, 38th Annual Symposium on, vol., no., pp.364-373.

3. Curtmola, R., Garay, J.A., Kamara, S., Ostrovsky, R.: Searchable symmetric encryption: improved definitions and efficient constructions. 13th ACM Conference on Computer and Communications Security (CCS '06), pp. 79–88, 2006.

4. Boneh, D., Di Crescenzo, G., Ostrovsky, R., Persiano, G.: Public key encryption with keyword search. *In* EUROCRYPT 2004. LNCS 3027, pp. 506–522. Springer, Heidelberg, 2004.

5. Giovanni Di Crescenzo and Vishal Saraswat. Public key encryption with searchable keywords based on Jacobi symbols. *In* Proceedings of the 8th International Conference on Progress in Cryptology (INDOCRYPT'07), Springer-Verlag, Berlin, Heidelberg, 282-296, 2007.

6. Rhee, H. S., Park, J.H., Susilo, W., Lee, D. H.: Improved searchable public key encryption with designated tester. *In* ASIACCS, pp. 376–379. ACM, New York, 2009.

7. Boneh, D., Waters, B.: Conjunctive, subset, and range queries on encrypted data. *In* TCC 2007. LNCS 4392, pp. 535–554. Springer, Heidelberg, 2007.

8. Yanpei Chen, Vern Paxson and Randy H. Katz, "What's New About Cloud Computing Security?" Technical Report No. UCB/EECS-2010-5, http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-5.html, Jan. 20, 2010.

9. RSA, The Role of Security in Trustworthy Cloud Computing

10. Ebenezer A. Oladimeji, Security threat Modeling and Analysis: A goal-oriented approach, 2006

11. Ristenpart, Thomas and Tromer, Eran and Shacham, Hovav and Savage, Stefan, Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds, 2009

12. Isaac Agudo David Nuñez , Gabriele Giammatteo , Panagiotis Rizomiliotis , Costas Lambrinoudakis , Cryptography goes to the Cloud, Research and Development Laboratory, Engineering Ingegneria Informatica S.p.A. Roma – Italy, June 2010.

13. Wikipedia, RSA algorithm, http://en.wikipedia.org/wiki/RSA_%28algorithm%29, accessed on 23-jan-2013.

14. Ronald, Russel, Cloud Security, Wiley India Pvt Ltd., vol 2 June 2010

15. csrs.nist.gov/groups/SNS/computing/Cloudcomputing-v25.ppt

16. http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031 (Accessed on 31-jan-2013)

About Authors

Dr. Mohammed Imtiaz Ahmed, Librarian, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh.
E-Mail: imtiazexplores@gmail.com.

Mr. Mohammed Bakhtawar Ahmed, Columbia Institute of Engg and Tech, Raipur, Chhattisgarh, India
E-mail: Bakhtawar229@gmail.com

Mr. Debojit Das, Columbia Institute of Engg And Tech, Raipur, Chhattisgarh, India
E-Mail: Debojit.das2707@gmail.com,