Volume : 3 | Issue : 5 | May 2014 • ISSN No 2277 - 8179

Research Paper

Computer Science

KEYWORDS :

Data Security in Cloud Computing Using Homomorphic encryption

Miss.Sumit Yadav

Usha Verma

Chhavi Bhardwaj

ABSTRACT

Cloud computing is an emerging paradigm which has become today's hottest research area due to its ability to reduce the costs associated with computing. In today's era, it is most interesting and enticing technology which is offering the services to its users on demand over the internet. Since Cloud Computing stores the data and disseminated resources in the open environment, security has become the main obstacle which is hampering the deployment of Cloud environments. Even though the Cloud Computing is promising and efficient, there are many challenges for data security as there is no vicinity of the data for the Cloud user. To ensure the security of data, we proposed a method by implementing homomorphic encryption. Security issues and issues regarding homomorphic algorithm is also introduced.

1. INTRODUCTION

Cloud computing is a technology that keep up data and its application by using internet and central remote servers. Cloud computing can be considered a new computing paradigm with implications for greater flexibility and availability at lower cost. Because of this, cloud computing has been receiving a good deal of attention lately.

Definition of cloud computing is based on five attributes: multitenancy (shared resources), massive scalability, elasticity, pay as you go, and self-provisioning of resources.

Multitenancy (shared resources): Unlike previous computing models, which assumed dedicated resources (i.e., computing facilities dedicated to a single user or owner), cloud computing is based on a business modelin which resources are shared (i.e., multiple users use the same resource) at the networklevel, host level, and application level.

Massive scalability: Although organizations might have hundreds or thousands of systems, cloud computingprovides the ability to scale to tens of thousands of systems, as well as the ability tomassively scale bandwidth and storage space.

Elasticity: Users can rapidly increase and decrease their computing resources as needed, as well asrelease resources for other uses when they are no longer required.

Pay as you go: Users pay for only the resources they actually use and for only the time they require them.

Self-provisioning of resources: Users self-provision resources, such as additional systems (processing capability, software,storage) and network resources.

Interest in the cloud is growing because cloud solutions provide users with access tosupercomputer-like power at a fraction of the cost of buying such a solution outright. Moreimportantly, these solutions can be acquired on demand; the network becomes thesupercomputer in the cloud where users can buy what they need when they need it. Cloudcomputing identifies where scalable IT-enabled capabilities are delivered as a service tocustomers using Internet technologies.

Cloud computing is expected to be a significant growth driver in worldwide IT spending. Infact, cloud services are expected to grow at a compound annual growth rate (CAGR) of 27%and reach $42 billion by 2012; spending on non-cloud IT services is expected to grow at aCAGR of 5%, according to IDC.

2. SPI FRAMEWORK FORCLOUD COMPUTING

A commonly agreed upon framework for describing cloud computing services goes by theacronym "SPI." This acronym stands for the three major services provided through the cloud:software-as-a-service (*S*aaS), platform-as-a-service (*P*aaS), and infrastructure-as-a-service(*I*aaS).

The Cloud Services Delivery Model

A cloud services delivery model is commonly referred to as an SPI and fallsinto three generally accepted services

i. The Software-As-a-Service Model

Traditional methods of purchasing software involved the customer loading the software ontohis own hardware in return for a license fee. The customercould also purchase a maintenance agreement to receive patches to the software or othersupport services.

In a SaaS model, the customer does not purchase software, but rather rents it for use on asubscription or pay-per-use model. Software as a Service (SaaS) is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet.

ii. The Platform-As-a-Service Model

Platform as a Service (PaaS) is an outgrowth of Software as a Service (SaaS), a software distribution model in which hosted software applications are made available to customers over the Internet. PaaS has several advantages for developers. With PaaS, operating system features can be changed and upgraded frequently. Geographically distributed development teams can work together on software development projects. Services can be obtained from diverse sources that cross international boundaries. Initial and ongoing costs can be reduced by the use of infrastructure services from a single vendor rather thanmaintaining multiple hardware facilities that often perform duplicate functions or suffer from incompatibility problems. Overall expenses can also be minimized by unification of programming development efforts.

PaaS solutions are development platforms for which the development tool itself is hosted inthe cloud and accessed through a browser. With PaaS, developers can often build webapplications without installing any tools on their computer, and can then deploy thoseapplications without any specialized system administration skills.

iii. Infrastructure as a Service (IaaS)

Infrastructure as a Service is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running and maintaining it. The client typically pays on a per-use basis.

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

Cloud Deployment Models

Public Cloud. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

Private Cloud. The cloud infrastructure is operated solely for a single organization. It may be managed by the organization or a third party, and may exist on-premises or off-premises.

Community Cloud. The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, or compliance considerations). It may be managed by the organizations or a third party and may exist on-premises or off premises.

Hybrid Cloud. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

3. SECURITY CONCERNS

Privacy and Confidentiality:

Once the client host data to the cloud there should be some guarantee that access to that data will only be limited to the authorized access. Inappropriate access to customer sensitive data by cloud personnel is another risk that can pose potential threat to cloud data. Assurances should be provided to the clients and proper practices and privacy policies andprocedures should be in place to assure the cloud users of the data safety.

The cloud seeker should be assured that data hosted on the cloud will be confidential.

Data integrity:

With providing the security of data, cloud service providers should implement mechanisms to ensure data integrity and be able to tell what happened to a certain dataset and at whatpoint. The cloud provider should make the client aware of what particular data is hosted on the cloud, the origin and the integrity mechanisms put in place.

For compliance purposes, it may be necessary to have exact records as to what data was placed in a public cloud, when it occurred, what virtual memories (VMs) and storage it residedon, and where it was processed. When such data integrity requirements exists, that the origin and custody of data or information must be maintained in order to prevent tampering or to prevent the exposure of data beyond the agreed territories (either between different servers or different networks).

Data location and Relocation:

Cloud Computing offers a high degree of data mobility.Consumers do not always know the location of their data. However, when an enterprise has some sensitive data that is kept on a storage device in the Cloud, they may want to knowthe location of it. They may also wish to specify a preferred location (e.g. data to be kept in India). This, then, requires a contractual agreement, between the Cloud provider and theconsumer that data should stay in a particular location or reside on a given known server. Also, cloud providers should take responsibility to ensure the security of systems (includingdata) and provide robust authentication to safeguard customers' information.

Another issue is the movement of data from one location to another. Data is initially stored at an appropriate location decide by the Cloud provider. However, it is often moved from one place to another. Cloud providers have contractswith each other and they use each other's resources.

Data Availability:

Customer data is normally stored in chunk on different servers often residing in different locations or in different Clouds. In this case, data availability becomes a major legitimate issue as the availability of uninterruptible and seamless provision becomes relatively difficult.

Storage, Backup and Recovery:

When you decide to move your data to the cloud the cloud provider should ensure adequate data resilience storage systems. At a minimum they should be able to provide RAID (Redundant Array of Independent Disks) storage systems although most cloud providers will store the data in multiple copies across many independent servers.

In addition to that, most cloud providers should be able to provide options on backup services which are certainly important for those businesses that run cloud based applications so that in the event of a serious hardware failure they can roll back to an earlier state.

4. HOMOMORPHIC ENCRYPTION

Need for homomorphic algorithm

By cloud computing we mean: The Information Technology (IT) model for computing, which is composed of all the IT components (hardware, software, networking, and services) that are necessary to enable development and delivery of cloud services via the Internet or a private network.

This definition doesn't mention any security notion of the data stored in the Cloud Computing even being a recent definition. Therefore we understand that the Cloud Computing is lacking security, confidentiality and visibility. To Provide Infrastructure (IaaS), Platform Service (PaaS) or Software (SaaS) as a Service is not sufficient if the Cloud provider does not guaranty a better security and confidentiality of customer's data.

By convention, we consider as Cloud Computing any treatment or storage of personal or professional information which are realized outside the concerned structure (i.e outside the company), to secure the Cloud means secure the treatments (calculations) and storage (databases hosted by the Cloud provider).

Cloud providers such as IBM, Google and Amazon use the virtualization on their Cloud platform and on the same server can coexist a virtualized storage and treatment space that belong to concurrent enterprises.

The aspect of security and confidentiality must intervene to protect the data from each of the enterprises.

Secure storage and treatment of data requires using a modern aspect of cryptography that has the criteria for treatment such as, the necessary time to respond to any request sent from the client and the size of an encrypted data which will be stored on the Cloud server.

Here arises two main questions: How to be sure that even if the data-centers of the Cloud Computing provider were attacked, my data won't be stolen or reused? And how can my data remain confidential and invisible even to my Cloud provider?

One proposal is to encrypt data before sending it to the cloud provider, but to execute the calculations the data should be decrypted every time we need to work on it. Until now it was impossible to encrypt data and to trust a third party to keep them safe and able to perform distant calculations on them. So to allow the Cloud provider to perform the operations on encrypted data without decrypting them requires using the cryptosystems based on Homomorphic Encryption.

HOMOMORPHIC ENCRYPTION

Homomorphic Encryption systems are used to perform operations on encrypted data without knowing the private key (without decryption), the client is the only holder of the secret key.

When we decrypt the result of any operation, it is the same as if we had carried out the calculation on the raw data.

Definition: An encryption is homomorphic, if: from Enc(a) and Enc(b) it is possible to compute Enc(f (a, b)), where f can be: +, ×, ⊕ and without using the private key.

Among the Homomorphic encryption we distinguish, according to the operations that allows to assess on raw data, the additive Homomorphic encryption (only additions of the raw data) is the Pailler and Goldwasser-Micalli cryptosystems, and the multiplicative Homomorphic encryption (only products on raw data) is the RSA and El Gamal cryptosystems.

History of the Homomorphic encryption

In 1978 Ronald Rivest, Leonard Adleman and Michael Dertouzos suggested for the first time the concept of Homomorphic encryption. Since then, little progress has been made for 30 years. The encryption system of ShafiGoldwasser and Silvio Micali was proposed in 1982 was a provable security encryption scheme which reached a remarkable level of safety, it was an additive Homomorphic encryption, but it can encrypt only a single bit. In the same concept in 1999 Pascal Paillier was also proposed a provable security encryption system that was also an additive Homomorphic encryption. Few years later, in 2005, Dan Boneh, Eu-Jin Goh and KobiNissim invented a system of provable security encryption, with which we can perform an unlimited number of additions but only one multiplication.

Additive Homomorphic Encryption

A Homomorphic encryption is additive, if:

$$Enc(x \oplus y) = Enc(x) \otimes Enc(y)$$

$$Enc\left(\sum_{i=1}^{l} mi\right) = \prod_{i=1}^{l} Enc(mi)$$

Suppose we have two ciphers C1 et C2 such that:

$C1 = gm1.\ r1n \bmod n2$

$C2 = gm2.\ r2n \bmod n2$

$C1.C2 = gm1.\ r1n\ .\ gm2.\ r2n \bmod n2 = gm1+ m2\ (r1r2)n \bmod n2$

So, Pailler cryptosystem realizes the property of additive Homomorphic encryption.

An application of an additive Homomorphic encryption is electronic voting: Each vote is encrypted but only the "sum" is decrypted.

Multiplicative Homomorphic Encryption

A Homomorphic encryption is multiplicative, if:

$$Enc(x \otimes y) = Enc(x) \otimes Enc(y)$$

$$Enc\left(\prod_{i=1}^{l} mi\right) = \prod_{i=1}^{l} Enc(mi)$$

Suppose we have two ciphers C1 et C2 such that:

$C1 = m1e \bmod n$

$C2 = m2e \bmod n$

$C1.C2 = m1em2e \bmod n = (m1m2)e \bmod n$

Result (C1 × C2) to the client.

If the attacker RSA cryptosystem realize the properties of the multiplicative Homomorphic encryption, but it still has a lake of security, because if we assume that two ciphers C1, C2 corresponding respectively to the messages m1, m2, so:

$C1 = m1e \bmod n$

$C2 = m2e \bmod n$

The client sends the pair (C1,C2) to the Cloud server, the server will perform the calculations requested by the client and sends the encrypted intercepts two ciphers C1 et C2, which are encrypted with the same private key, he/she will be able to decrypt all messages exchanged between the server and the client. Because the Homomorphic encryption is multiplicative, i.e. the product of the ciphers equals the cipher of the product.

RSA Algorithm:

RSA is widely used Public-Key algorithm. RSA stands for Ron Rivest, Adi Shamir and Len dleman, who first publicly described it in 1977. In our proposed work, we are using RSA algorithm to encrypt the data to provide security so that only the concerned user can access it. By securing the data, we are not allowing unauthorized access to it.

User data is encrypted first and then it is stored in the Cloud. When required, user places a request for the data for the Cloud provider, Cloud provider authenticates the user and delivers the data.

RSA is a block cipher, in which every message is mapped to an integer. RSA consists of Public-Key and Private-Key. In our Cloud environment, Pubic-Key is known to all, whereas Private-Key is known only to the user who originally owns the data. Thus, encryption is done by the Cloud service provider and decryption is done by the Cloud user or consumer. Once the data is encrypted with the Public-Key, it can be decrypted with the corresponding Private-Key only.

RSA algorithm involves three steps:

1. Key Generation
2. Encryption
3. Decryption

1) Key Generation:

Before the data is encrypted, Key generation should be done. This process is done between the Cloud service provider and the user.

Steps:

1. Choose two distinct prime numbers a and b. For security purposes, the integers a and b should be chosen at random and should be of similar bit length.
2. Compute n = a * b.
3. Compute Euler's totient function, Ø(n) = (a-1) * (b-1).
4. Chose an integer e, such that 1 < e <Ø(n) and greatestcommon divisor of e , Ø(n) is 1. Now e is released asPublic-Key exponent.
5. Now determine d as follows: d = e-1(mod Ø(n)) i.e., d ismultiplicate inverse of e mod Ø(n).
6. d is kept as Private-Key component,so that d * e = 1 mod Ø(n).
7. The Public-Key consists of modulus n and the publicexponent e i.e, (e, n).
8. The Private-Key consists of modulus n and the privateexponent d, which must be kept secret i.e, (d, n).

2) Encryption:

Encryption is the process of converting original plain text (data)

into cipher text (data).

Steps:
1. Cloud service provider should give or transmit the Public-Key (n, e) to the user who want to store the data with him or her.
2. User data is now mapped to an integer by using an agreed upon reversible protocol, known as padding scheme.
3. Data is encrypted and the resultant cipher text(data) C is C = me (mod n).
4. This cipher text or encrypted data is now stored with the Cloud service provider.

3) Decryption*:*
Decryption is the process of converting the cipher text(data) to the original plain text(data).

Steps:
1. The cloud user requests the Cloud service provider for the data.
2. Cloud service provider verify's the authenticity of the user and gives the encrypted data i.e, C.
3. The Cloud user then decrypts the data by computing, m = Cd (mod n).
4. Once m is obtained, the user can get back the original data by reversing the padding scheme.

Example:we are taking some sample data end implementing RSA algorithm over it.

Key Generation:
1. We have chosen two distinct prime numbers a=61 and b=53.
2. Compute n=a*b, thus n=61*53 = 3233.
3. Compute Euler's totient function, Ø(n)=(a-1)*(b-1), Thus Ø(n)=(61-1)*(53-1) = 60*52 = 3120.
4. Choose any integer e, such that 1 < e < 3120 that is coprime-to 3120. Here, we chose e=17.
5. Compute d , d = e-1(mod Ø(n)), thus d=17-1(mod 3120) = 2753.
6. Thus the Public-Key is (e, n) = (17, 3233) and the Private-Key is (d, n) = (2753, 3233). This Private-Key is kept secretand it is known only to the user.

Encryption:
1. The Public-Key (17, 3233) is given by the Cloud serviceprovider to the user who wish to store the data.
2. Let us consider that the user mapped the data to an integerm=65.
3. Data is encrypted now by the Cloud service provider byusing the corresponding Public-Key which is shared by boththe Cloud service provider and the user.

C = 6517(mod 3233) = 2790.

4. This encrypted data i.e, cipher text is now stored by theCloud service provider.

Decryption:
1. When the user requests for the data, Cloud service providerwill authenticate the user and delivers the encrypted data (If the user is valid).
2. The cloud user then decrypts the data by computing, m = Cd(mod n) = 27902753(mod 3233) = 65.
3. Once the m value is obtained, user will get back the originaldata.

5. ISSUSES
Homomorphic encryption is one of the most exciting new research topics in cryptography, which promises to make cloud computing perfectly secure. With it, a Web user would send encrypted data to a server in the cloud, which would process it without decrypting it and send back a still-encrypted result.

Sometimes, however, the server needs to know something about the data it's handling. Otherwise, some computational tasks become prohibitively time consuming — if not outright impossible.

Suppose, for instance, that the task you've outsourced to the cloud is to search a huge encrypted database for the handful of records that match an encrypted search term.

Homomorphic encryption ensures that the server has no idea what the search term is or which records match it. As a consequence, however, it has no choice but to send back information on every record in the database. The user's computer can decrypt that information to see which records matched and which didn't, but then it's assuming much of the computational burden that it was trying to offload to the cloud in the first place.

6. CONCLUSION
Cloud Computing is still a new and evolving paradigm where computing is regarded as on-demand service. Once the organization takes the decision to move to the cloud, it loses control over the data. Thus, the amount of protection needed to secure data is directly proportional to the value of the data.

Security of the Cloud relies on trusted computing and cryptography.

Thus, in our proposed work, only the authorized user can access the data. Even if some intruder (unauthorized user) gets the data accidentally or intentionally if he captures the data also, he can't decrypt it and get back the original data from it.

Hence forth, data security is provided by implementing RSA algorithm.

REFERENCE

1. P.Kalpana, "Cloud Computing – Wave of the Future", International Journal of Electronics Communication and Computer Engineering, Vol 3, Issue 3, ISSN 2249–071X, June 2012. | 2. SubedariMithila, P. Pradeep Kumar, "Data Security through Confidentiality in Cloud Computing | 3. Environment", SubedariMithila et al, / (IJCSIT) International Journal of Computer Science and | 4. Information Technologies, Vol. 2 , 1836-1840, 2011. | 5. Zaigham Mahmood, "Data Location and Security Issues in Cloud Computing", Proceedings of International Conference on Emerging Intelligent Data and Web Technologies-2011. | 6. Vishwagupta, Gajendra Singh, Ravindra Gupta, "Advance Cryptography algorithm for improving data security", International Journal of Advanced Research in Computer Science and Software Engineering, Vol 2, Issue 1, Jan 2012. | 7. Simarjeet Kaur, "Cryptography and Encryption in Cloud Computing", VSRD International Journal of Computer Science and Information Technology, Vol.2(3), 242-249, 2012. | 8. Craig Gentry, A Fully Homomorphic Encryption Scheme, 2009. | 9. TaherElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory, 469-472, 1985. | 10. R. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public key cryptosystems. Communications of the ACM, 21(2) :120-126, 1978. Computer Science, pages 223-238. Springer, 1999. |