Designing Data Storage Security and Privacy in Cloud Computing

Mohit Marwaha
Department of CSE
Punjab Technical University, Jalandhar, India
infomaticmohit@gmail.com

Rajeev Bedi
Assistant Professor/Department of CSE
Punjab Technical University, Jalandhar, India
rajeevbedi@rediffmail.com

***Abstract*— Cloud computing has certainly created a buzz around the world of it is the next big thing after internet in the field of information technology; some say it's a metaphor for internet. It is a computing technology based on internet, Cloud computing presents a new model for IT service delivery and it typically involves over-a-network, on-demand, self-service access, which is dynamically scalable and elastic, utilizing pools of often virtualized resources. Even though the cloud continues to grow in popularity, Usability and respectability, Questions about cloud data protection, data privacy and other Security issues may continue to linger and are cited as the most substantial roadblock for cloud computing uptake. Privacy and security are the key issue for cloud storage. As promising as it is, this paradigm also brings into limelight many new challenges for data security and access control when users outsource sensitive data for sharing on cloud servers, which are not within the same trusted domain as data owners. To keep sensitive user data confidential against untrusted servers, we suggest applying cryptographic methods by disclosing data decryption keys only to authorized users. The paper analyzes the feasibility of the applying RSA encryption algorithm for data security and privacy in cloud Storage.**

Index Term : Cloud Data Storage, Cryptographic methods, RSA encryption algorithm.

I. INTRODUCTION

Cloud computing is a flexible, cost- effective and proven delivery platform for providing business or consumer IT services over the Internet. Cloud computing supports distributed service oriented architecture, multi-users and multi-domain administrative infrastructure, it is more prone to security threats and vulnerabilities. At present, a major concern in cloud adoption is its security and Privacy. Intrusion prospects within cloud environment are many and with high gains. Security and Privacy issues are of more concern to cloud service providers who are actually hosting the services. In most cases, the provider must guarantee that their infrastructure is secure and clients' data and applications are safe by implementing security policies and mechanisms. While the cloud customer must ensure that provider has taken proper security measures to protect their information. The issues are organized into several general categories: trust, architecture, identity management, software isolation, data protection, availability Reliability, Ownership, Data Backup, Data Portability and Conversion, Multiplatform Support and Intellectual Property.

II. Cloud Computing Framework

Service Models: These three are the most widely used service models of cloud computing.

1. Software as a service Software-as-a-Service (SaaS): It is also referred as software available on demand; it is based on multi-tenant architecture. Software like word processor, CRM (Customer Relation Management), etc. or application services like schedule, calendar, etc. are executed in the "cloud" using the interconnectivity of the internet to do manipulation on data. Custom services are combined with 3rd party commercial services via Service oriented architecture to create new applications. It is a software delivery for business applications like accounting, content delivery, Human resource management (HRM), Enterprise resource planning (ERP) etc on demand on pay-as-you go model [1].

2. Platform as a Service Platform-as-a-Service (PaaS): This layer of cloud provides computing platform and solution stack as service**.** Platform-as-a-Service provides the user with the freedom of application design, application development, testing, deployment and hosting as well as application services such as team collaboration, web service integration and database integration, security, scalability, storage, persistence, state management, application versioning, without thinking about the underlying hardware and software layers by providing facilities required for completion of project through web application and services via Internet.

3. Infrastructure as a Service Infrastructure-as-a-Service (IaaS): Infrastructure as a service delivers a platform virtualization environment as a service. Instead of purchasing servers, software, data center space or network equipment, clients can buy these resources as outsourced service. In other words 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.

III. Cloud Deployment Models

There are three types cloud Deployment models that widely used are:

1. Public It is referred as external cloud or multi-tenant cloud, this model represents an openly accessible cloud environment in this cloud can be accessed by general public. Customer can access resources and pay for the operating resources. Public Cloud can host individual services as well as collection of services

2. Private It is also known as internal cloud or on-premise cloud, a private cloud provides a limited access to its resources and services to consumers that belong to the same organization that owns the cloud. In other words the infrastructure that is managed and operated for one organization only so that a consistent level of control over security, privacy, and governance can be maintained.

3. Hybrid A hybrid cloud is a combination of public and private cloud. It provides benefits of multiple deployment models. It enables the enterprise to manage steady-state workload in the private cloud, and if the workload increases asking the public cloud for intensive computing resources, then return if no longer needed.

4. Community This deployment model share resources with many organizations in a community that shares common concerns (like security, governance, compliance etc). It typically refers to special-purpose cloud computing environments shared and managed by a number of related organizations participating in a common domain or vertical market [12].

IV. Issues in Cloud Data Storage.

Cloud Computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy. This unique attribute, however, poses many new security challenges which have not been well understood. In this article, we focus on cloud data storage security, which has always been an important aspect of quality of service. To ensure the correctness of user's data in the cloud.

A. Trust: Trust is defined as reliance on the integrity, strength, ability and surety of a person or thing. Entrusting your data on to a third party who is providing cloud services is an issue. Recent incidents like In April of 2012 Amazon's Elastic Compute Cloud service crashed during a system upgrade, knocking customers' websites off-line for anywhere from several hours to several days. That same month, hackers broke into the Sony PlayStation Network, exposing the personal information of 77 million people around the world. And in June a software glitch at cloud-storage provider Dropbox temporarily allowed visitors to log in to any of its 25 million customers' accounts using any password or none at all. These issues have certainly created doubts in mind of cloud consumers and damaged the trust ability of Consumers [4].

B. Privacy: Different from the traditional computing model, cloud computing utilizes the virtual computing technology, users' personal data may be scattered in various virtual data center rather than stay in the same physical location, even across the national borders, at this time, data privacy protection will face the controversy of different legal systems. On the other hand, users may leak hidden information when they accessing cloud computing services. Attackers can analyze the critical task depend on the computing task submitted by the users [9].

C. Security: Cloud service providers employ data storage and transmission encryption, user authentication, and authorization. Many clients worry about the vulnerability of remote data to criminals and hackers. Cloud providers are enormously sensitive to this issue and apply substantial resources to mitigate this problem.

D. Ownership: Once data has been relegated to the cloud, some worry about losing their rights or being unable to protect the rights of their customers. Many cloud providers address this issue with well-skilled user-sided agreements. According to the agreement, users would be wise to seek advice from their favourite legal representative [10].

E. Performance and Availability: Business organizations are worried about acceptable levels of performance and availability of applications hosted in the cloud.

F. Legal: There are certain apprehensions for a cloud service provider and a client receiving the service like location of the cloud provider, infrastructure and physical location of the data and outsourcing of the cloud provider's services etc.

G. Multiplatform Support: More an issue for IT departments using managed services is how the cloud-based service integrates across different platforms and operating systems, e.g. OS X, Windows, Linux and thin-clients. Usually, some customized adaption of the service takes care of any problem. Multiplatform support requirements will ease as more user interfaces become web-based.

H. Intellectual Property: A company invents something new and it uses cloud services as part of the invention. Is the invention still patentable? Or there can be issues like cloud service provider can make claim for that invention or leak the information to the competitor.

I. Data Backup*:* Cloud providers employ redundant servers and routine data backup processes, but some people worry about being able to control their own backups. Many providers are now offering data dumps onto media or allowing users to back up data through regular downloads.

J. Data Portability and Conversion*:* Some people have concerns like, switching service providers; there may be difficulty in transferring data. Porting and converting data is highly dependent on the nature of the cloud provider's data retrieval format, particular in cases where the format cannot be easily revealed. As service competition grows and open standards become established, the data portability issue will ease, and conversion processes will become available supporting the more popular cloud providers. Worst case, a cloud subscriber will have to pay for some custom data conversion.

These are certain areas in which cloud computing requires to excel and solve problem related to it. Out of all the problems Security, Privacy and Intellectual property put the major threats on growth of cloud computing that are needed to be worked upon.

V. PROBLEM DEFINATION

We have discussed some serious issues in cloud computing in the above section where Security, Privacy and trust are major concern. Entrusting a third party by user is never easy and incidents like data thefts or data manipulation and hacking of data makes it more vulnerable. So what our system purpose is if security of data is taken care by user itself it will eliminates all three issues of data Security Privacy and Trust. As user will be able to encrypt and decrypt his data on his own with the interference of the cloud service provider.

VI. OVERVIEW OF OUR APPROACH

Our goal is to build up a repository to facilitate the data integration and sharing across cloud along with preservation of data confidentiality. For this we will be using an encryption technique to provide data security on data storage [16].

A. Objectives

In This paper we propose the following objectives

1. To develop a system those will Provide Security and Privacy to Cloud Storage

2. To Establish An Encryption Based System For protecting Sensitive data on the cloud and Structure how owner and storage Service Provider to operate on encrypted Data

3. To Create a System where the user store its data on the cloud the data is sent and stored on the cloud in encrypted form As in normal cases in cloud computing when a user login to the cloud and they store data on cloud storage device the data stored on the server cloud is not much secure as it can be readable to anyone which have permission to access and Leaving data vulnerable,

4. To Develop a retrieval System in which the data is retrieved by the user in encrypted form and is decrypted by the user at its own site using a public and private key encryption both the keys working at the user level.

VII. PROPOSED METHODOLOGY

- First of all we develop a web cloud application using java language and in DotNet and implement RSA Algorithm to secure data. We can develop web application using Eclipse IDE and in DotNet Framework.
- Then deploy web application for cloud.
- We shall test web application using Apache tomcat 7.0.22 web Server.
- To test the web application on cloud we can use the Azure Emulator. Azure version which we can use is 4.0 and to store the data we can use the SQL Server 2008 R2
- Our Work is to provide Data security on cloud to user which can upload there data on cloud Servers.
- We can implement a mechanism to improve the cloud data security.
- In this we can use RSA encryption Mechanism to improve cloud security.
- We can develop web cloud Application which can upload data on cloud server in encrypted form and
- The private key to decrypt data will be send to the user email account with decryption Software.

VIII SIMULATION SETUP

We will make a web Application using either java's netbeans or asp.net that will work as the front end of our cloud system then use Windows Azure Emulator to provide infrastructure as a service as the storage service. Windows Azure Emulator works as the backend attached with SQL Server to Store information of the user. System Generated email is used to send the Email to the user.

RSA is an algorithm for public-key cryptography is used for encryption that is based on the presumed difficulty of factoring large integers, the factoring problem. RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman. We have specifically used RSA algorithm for the reason being it's the only asymmetrical encryption algorithm.

RSA involves a public key and a private key. The public key can be known to everyone and is used for encrypting messages. Messages encrypted with the public key can only be decrypted using the private key. The keys for the RSA algorithm are generated the following way:

1. Choose two distinct prime numbers *p* and *q*.
    - For security purposes, the integer's *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 $n = pq$.
    - *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 $\varphi(n) = (p-1)(q-1)$, where φ is Euler's totient function.
4. Choose an integer *e* such that $1 < e < \varphi(n)$ and greatest common divisor of $(e, \varphi(n)) = 1$; i.e., *e* and $\varphi(n)$ are coprime.
    - *e* is released as the public key exponent.
    - *e* having a short bit-length and small Hamming weight results in more efficient encryption - most commonly 0x10001 = 65,537. However, small values of *e* (such as 3) have been shown to be less secure in some settings.[4]
5. Determine *d* as:

$$d \equiv e^{-1} \pmod{\varphi(n)}$$

i.e., *d* is the multiplicative inverse of *e* mod $\varphi(n)$.

- This is more clearly stated as solve for d given (de) = 1 mod $\varphi(n)$
- This is often computed using the extended Euclidean algorithm.
- *d* is kept as the private key exponent.

By construction, d*e= 1 mod $\varphi(n)$. The public key consists of the modulus *n* and the public (or encryption) exponent *e*. The private key consists of the modulus *n* and the private (or decryption) exponent *d* which must be kept secret. (*p*, *q*, and $\varphi(n)$ must also be kept secret because they can be used to calculate *d*.)

- An alternative, used by PKCS#1, is to choose *d* matching $de \equiv 1 \mod \lambda$ with $\lambda = \text{lcm}(p-1, q-1)$, where lcm is the least common multiple. Using λ instead of $\varphi(n)$ allows more choices for *d*. λ can also be defined using the Carmichael function, $\lambda(n)$.
- The ANSI X9.31 standard prescribes, IEEE 1363 describes, and PKCS#1 allows, that *p* and *q* match additional requirements: be strong primes, and be different enough that Fermat factorization fails.

This algorithm will generate two keys the public key will be used by user when he will upload the data and immediately a email will be sent to user with the private key to decrypt the data when user want to use the data he will login in his account choose the file and open. An executable file will run that will ask for the private key when user will provide the private key he can work on his data.

IX. RESULTS

1. The system provides Security and Privacy to Cloud Storage

2. The RSA Encryption Based System protecting Sensitive data on the cloud and give owner rights to encrypt and decrypt data on his will.

3. In the Created System the user store its data on the cloud the data is sent and stored on the cloud in encrypted form, an Email containing the Decryption is sent by the system and when the user want to decrypt the data he uses that decryption key with the executable file that is loaded with key to decrypt his data.

4. With this system we are able to solve problems of Data Security, Privacy, Trust and Intellectual Property Rights

X. CONCLUSION

Our research indicates that that Security and Privacy are the major issues that are needed to be countered, efforts are being made to develop many efficient System That can Provide Security and privacy at the user level and maintain the trust and intellectual property rights of the user. Our method States Encryption is one such method that can provide peace of mind to user and if the user have control over encryption and decryptions of data that will boost consumer confidence and attract more people to cloud platform.

References

[1]. http://en.wikipedia.org/wiki/Cloud_computing.
[2]. Rich Maggiani, solari communication. "Cloud computing is changing how we communicate".
[3]. Randolph Barr, Qualys Inc, "How to gain comfort in losing control to the cloud".

[4]. Greg Boss, Padma Malladi, Dennis Quan, Linda Legregni, HaroldHallHiPODS,www.ibm.com/developerworks/webspere/zones/hipds.
[5]. http://www.roughtype.com.
[6]. Tharam Dillon, Chen Wu, Elizabeth Chang, 2010 24th IEEE International Conference on Advanced Information Networking and Applications, "Cloud computing: issues and challenges".
[7]. June13, 2009, http://server.zol.com.cn/183/ 1830464.html.
[8]. Elinor Mills, January 27, 2009. "Cloud computing security forecast: clear skies".
[9]. Jianchun Jiang, Weiping Wen, "Information security issues inCloudcomputingenvironment,NetinfoSecurity,doi:10.3969/j.issn.1671-1122.2010.02.026.
[10]. Jianchun Jiang, Weiping Wen, "Information security issues in cloud computing environment", Net info Security,doi:10.3969/j.issn.1671- 1122.2010.02.026. of virtual machines" In Proc. Of NSDI'05, pages 273-286, Berkeley CA, USA, 2005. USENIX Association.
[11]. Eucalyptus Completes Amazon Web Services Specs with Latest Release.
[12]. Open Cloud Consortium.org.
[13]. July 27,2009. Available from http://fx.caixun.com/.
[14]. Jack Schofield. Wednesday 17 June 2009 22.00 BST, http://www.guardian.co.uk/technology/2009/jun/17/cloud-computing jack- schofield.
[15]. Gartner. "Seven cloud-computing security risks".
[16]. Ranjita Mishra "A Privacy Preserving Repository for Securing Data across the Cloud".

Authors Profile

Mohit Marwaha received the **B.Tech.** degree in Information Technology from the Beant College of Engineering and Technology, Gurdaspur, Punjab Technical University, Jalandhar, India, in 2008. Currently pursuing **M.Tech.** in Computer Science engineering (cloud computing) in Punjab Technical University, Jalandhar, India. My research interest includes Cloud Computing, Distributed computing and Network security

Rajeev Bedi received the **B.Tech.** degree in Computer Science and Engineering from Punjab Technical University, Jalandhar, in 2000 and completed **M.Tech.** in Computer Science engineering from Punjab Technical University, Jalandhar, in 2008. Currently doing PhD from CMJ University Shillong. My research interest includes Cloud Computing and Distributed computing.