Data Integrity and Security in Cloud

Ku.Swati G. Anantwar, Prof. Karuna G. Bagde

Abstract -- **Cloud computing has been envisioned as the next-generation architecture of IT enterprise. The important concerns that need to be addressed in cloud computing is to assure the customer of the integrity i.e. correctness of his data in the cloud. Our aim is to provide a safe and secured Cloud Computing Environment whose main objectives are ,to achieve data and Network security. In this scheme , we focus on the software as well as the platform as a service of cloud. Here the computational time and network bandwidth is minimized by using integrity and encryption methodology on the data in cloud . It should be noted that this scheme applies not only to static storage data but to the dynamic also .**

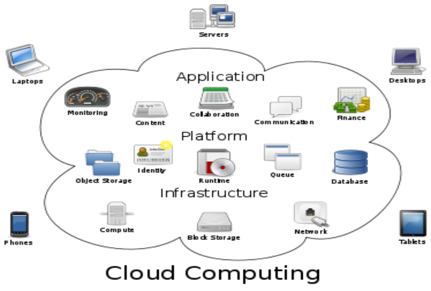
Index Terms— **Cloud computing, Encryption, data integrity**

I. Introduction

A. *Cloud Computing*

Cloud Computing, to put it simply, means "Internet Computing." The Internet is commonly visualized as clouds; hence the term "cloud computing" for computation done through the Internet. With Cloud Computing users can access database resources via the Internet from anywhere, for as long as they need, without worrying about any maintenance or management of actual resources. Besides, databases in cloud are very dynamic and scalable.

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



Cloud Computing brings with it many benefits to the end user. These include:

1.Access to a huge range of applications without having to download or install anything

2.Applications can be accessed from any computer, anywhere in the world

3.Users can avoid expenditure on hardware and software; only using what they need

4.Companies can share resources in one place

5.Consumption is billed as a utility with minimal upfront costs

6.Scalability via on-demand resources

There are several differences from traditional hosting to cloud hosting. The main differences are:
1.Cloud Computing is sold on demand

2.The service is managed by the provider

3.User can determine the amount of service they take

4.Users can log on to the network from any computer in the world .

B. *Cloud Service Models*

Types of cloud service
There three main types of cloud service provided:

1) Software as a Service (SaaS) :

Clients can use the software provide by provider. Which usually need not to install and it is usually a one to many service. Like Gmail, search engine.
The typical cloud application we used is Google Doc, a powerful web application similar to Microsoft office. It is free for most users and has paid version for the company who want more features.
Apart from the software we often use in office, there are some more powerful Cloud service like Jaycut and Pixlr.
Example: Yahoo!, Gmail, Google Docs, etc.

Jaycut is a free online application implemented with Flash; you can upload movies and edit it. Jaycut is a very powerful and can do almost all basic effects same as other desktop application. After finishing your work, it can compile the final product with a magical high speed and give you a download link. You can deliver the movies by passing the links to others and need not to worry about the storage problem.

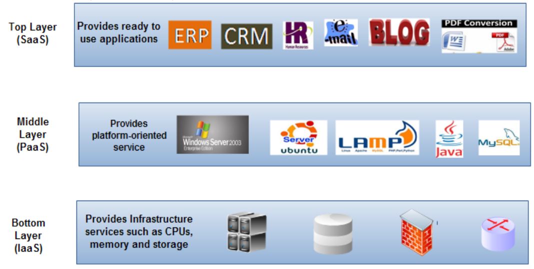
Jaycut, web video editor

2) Platform as a Service (PaaS):

Clients can run their own applications on the platform provided; General platforms are Linux and Windows.

3) Infrastructure as a Service (IaaS Infrastructure-as-a-Service (IaaS).

The IaaS service model is the lowest service model in the technology stack, offering infrastructure resources as a service, such as raw data storage, processing power and network capacity. The consumer can the use IaaS based service offerings to deploy his own operating systems and applications, offering a wider variety of deployment possibilities for a consumer than the PaaS and SaaS models. "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)". Example: Amazon (S3, EC2), Windows Azure, etc.


Top Layer (SaaS)
Provides ready to use applications
ERP
CRM
HR
e-mail
BLOG
PDF Conversion
Middle Layer (PaaS)
Provides platform-oriented service
Windows Server 2003
Server ubuntu
LAMP
Java
MySQL
Bottom Layer (IaaS)
Provides Infrastructure services such as CPUs, memory and storage

.

In practical usage, PaaS and IaaS are actually very similar with the difference that whether the image is provided by user or not. If you use the image provided, it is PaaS, otherwise, it is IaaS.

Cloud deployment models

Regardless of which delivery model is utilized, cloud offerings can be deployed in four primary ways, each with their own characteristics. The characteristics to describe the deployment models are; (i) who owns the infrastructure; (ii) who manages the infrastructure; (iii) where is the infrastructure located; (iv) and who accesses the cloud services.

Cloud Computing is a web processing with large amount of resource. The user of the cloud can obtain the service thought network (in both internet and intranet). In other words, users are using or buying computing service from others. The resource of the cloud can be anything IT related. In general, cloud provides application, computation power, storage, bandwidth, database and some technologies like Map Reduce. As the resource pool is very large, user can increase the application on cloud to any scale. It is fully under users‟ control.

Types of cloud resources

There are four main type of cloud:

1) **Public cloud**: The cloud computing resource is shared outside, anyone can use it and some payment maybe need.

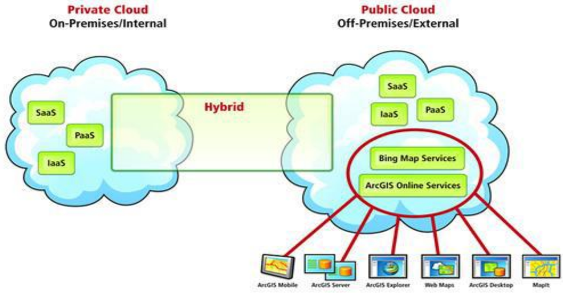
Eample: Google Aps, SQL Azure, etc.

Fig:Pre-installed image provided by AWS

2) **Private cloud**: It is opposite to public cloud, private cloud's resource is limit to a group of people, like a staff of a company etc.

3) **Hybrid cloud**: this is a mixture of previous two clouds, some cloud computing resource is shared outside but some don‟ t.

4) **Community cloud**: this is a special cloud to make use of cloud computing‟ s features. More than one community shares a cloud to share and reduce the cost of computing system.



Cloud Computing has many benefits, however there are also some associated risks with using cloud computing. These include:

1. Users do not physically possess storage of their own data, which leaves the responsibility and control of data storage with the provider.

2. Users could become dependent upon the cloud computing provider.

3. With data held externally, business continuity and disaster recovery are in the hands of the provider.

4. Data migration issues when changing cloud provider.

5. What happens if your cloud provider goes out of business?

II. RELATED WORK

Cloud computing provides the way to share distributed resources and services that belong to different organizations or sites. Since cloud computing share distributed resources via the network in the open environment, thus it makes security problems important for us to develop the cloud computing application.

Sravan Kumar R and Ashutosh Saxena [1] have have worked to facilitate the client in getting a "proof of integrity of the data " which he wishes to store in the cloud storage servers with bare minimum costs and efforts.

Meiko Jensen et al. [2] have shown that to improve cloud computing security, the security capabilities of both web browsers and web service frame works, should be strengthened. This can best be done by integrating the latter into the former.

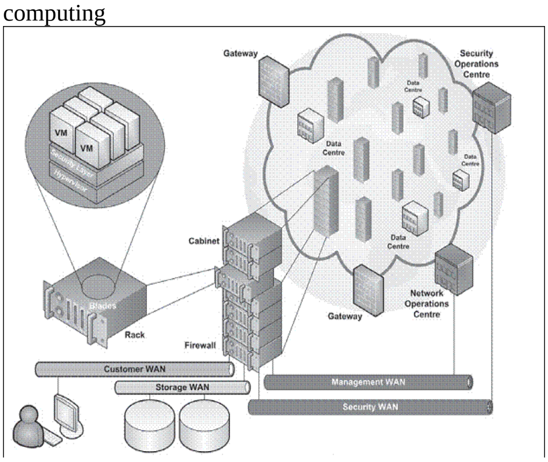
M. Jensen et al. [3] focus on special type of Denial of Service attacks on network based service that relies on message flooding techniques, overloading the victims with invalid requests. They describe some well known and some rather new attacks and discuss commonalities and approaches for countermeasures.

Armbust M Fox et al. [4] discuss that resources should be virtualized to hide the implementation of how they are multiplexed and shared.

Kaufman, L. M. [5] have shown that to ensure CIA (Confidentiality, Integrity and Availability) of the information, the service provider should offer tested encryption schema, stringent access controls and scheduled data backups.

Sloan, K. [6] has explored and demystified the technologies involved in cloud computing in which he discusses about the challenges posed in security of cloud computing. According to him, security components could be added to the security layer and be delivered as Security as a Service.

Figure shows the security architecture of cloud computing



Hewitt, C. [7] To guarantee the privacy of information hosted on servers in cloud, the information could be encrypted which can only be decrypted at the client level with a key. Again this is only reliable if the data can be quickly decrypted at the client level as it might need high processing power. The multi-core processors which are evolving will make this possible and provide greater integration of information.

Providing security for cloud computing requires more than authentication using passwords and confidentiality in data transmission. Vieira, K., A. Schulter, et al. [8] have proposed a solution for intrusion detection in cloud computing. The solution consists of two kinds of analysis behavioral analysis and knowledge analysis.

Mowbray, M. and S. Pearson [9] has proposed a client based privacy manager to eliminate the fear of data leakage and loss of privacy in cloud computing. In the paper, they have presented a scenario of salesforce.com which can undergo a security threat; theft of sales data and various ways that an intruder can gain knowledge based on the un-encrypted data. The threats include the collection of personal information and getting inappropriate access to the information. Based on this scenario a set of requirements was derived which include the minimization of personal and sensitive data used in cloud and maximising security protection of data. Finally the overall architecture for client-based privacy data manager has been depicted.

[10] describes Amazon Web Services' (AWS) physical and operational security processes for network and infrastructure under Amazon Web Services (AWS) management. It also gives service specific security implementations for Amazon Web Services (AWS).

To guarantee the privacy of information hosted on servers in cloud, the information could be encrypted which can only be decrypted at the client level with a key. Again this is only reliable if the data can be quickly decrypted at the client level as it might need high processing power. The multi-core processors which are evolving will make this possible and provide greater integration of information (Hewitt, C., 2008).

Providing security for cloud computing requires more than authentication using passwords and confidentiality in data transmission. Vieira, K., A. Schulter, et al. (2009) have proposed a solution for intrusion detection in cloud computing. The solution consists of two kinds of analysis behavioral analysis and knowledge analysis.

Mowbray, M. and S. Pearson (2009) has proposed a client based privacy manager to eliminate the fear of data leakage and loss of privacy in cloud computing. In the paper, they have presented a scenario of salesforce.com which can undergo a security threat; theft of sales data and various ways that an intruder can gain knowledge based on the un-encrypted data. The threats include the collection of personal information and getting inappropriate access to the information. Based on this scenario a set of requirements was derived which include the minimization of personal and sensitive data used in cloud and maximizing security protection of data. Finally the overall architecture for client-based privacy data manager has been depicted.

III. ANALYSIS OF PROBLEM

The first problem is establishing trust in remote execution. Essentially, a cloud is a distributed computing architecture in which the client's computation runs on a remote host in a data center. In a cloud system, a customer must gain assurance that the base system executes his or her cloud instance while protecting its integrity and secrecy, tantamount to running on the customer's own machine.

The second problem is protecting the execution of one cloud instance from other instances on the same base system or infrastructure. Once again, the base system must enforce this requirement. Base systems and cloud services are shared resources that any cloud customer can leverage, so they must ensure isolation during interactions with customers.

The third problem is protecting the execution of a cloud instance from external adversaries. Often, a cloud instance needs network access to communicate with the customer and interact with storage. In both cases, the cloud architecture can simplify administration because the base system knows the location of both storage and the customer. However, if the instance needs open communication to the Internet, then it faces risks similar to those of a typical desktop or server. Customers must ensure that their instance's security configuration, including firewall, access control, and authentication, protects the instance from external attackers, as they would for a normal server.

IV. PROPOSED WORK

One of the important concern in cloud computing is that there is need to be addressed to assure the customer of the integrity i.e. correctness of his data in the cloud. As the data is physically not accessible to the user the cloud should provide a way for the user to check if the integrity of his data is maintained or is compromised. We provide a scheme which gives a proof of data integrity in the cloud which the customer can employ to check the correctness of his data in the cloud. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service level agreement (SLA). The proof of data integrity protocol just checks the integrity of static as well as dynamic data i.e. if the data has been illegally modified or deleted.

Here we are Implementing Cloud Environment(PaaS) using virtualizations and isolations in Java (EJB , OSGi containers, class loaders etc) for that the following tools are also require

- For Intrusion detection and prevention SNORT
- For H/W security Firewalls
- S/W security using Antivirus
- User Data Security by encryption with 3DES

V. DESIRED IMPLICATION

In our project we are going to use 3DES algorithm for Data Encryption, SHA1 algorithm for Data Integrity , and Multilevel security for securing the data.

A.Data Encryption:

For Data Encryption can be achieved using many algorithms but experimentally it is proved that TDES is best among all.

As the DES algorithm is reasonably secure and fast. There is no feasible way to break DES, however because DES is only a 64-bit (eight characters) block cipher, an exhaustive search of 2^{55} steps on average, can retrieve the key used in the encryption. For this reason, it is a common practice to protect critical data using something more powerful than DES.

A much more secure version of DES called Triple-DES (TDES), which is essentially equivalent to using DES three times on plaintext with three different keys. It is considered much safer than the plain DES and like DES, TDES is a block cipher operating on 64-bit data blocks. There are several forms, each of which use the DES cipher three times.

Data security is maintained on 3 layers

In this scheme we are using Secure Hash Algorithm (SHA1) algorithm for data integrity. We are trying to obtain and verify a proof that the data that is stored by a user at remote data storage in the cloud is not modified by the archive and thereby the integrity of the data is assured. We propose this scheme for software as a service as well as platform as a service. With this scheme we are using different security levels such as authentication, encryption and decryption and data recovery We are providing encryption to meta data only and for storing actual data

This scheme will reduce the computational and storage overhead of the client as well as minimize the computational overhead of the cloud storage server, this will also minimized the size of the proof of data integrity so as to reduce the network bandwidth consumption.

B. Data Integrity:

There are 3 Secure Hash Algorithms SHA0 ,SHA1,SHA2, Out of these we are using SHA1 algorithm because.SHA-1 forms part of several widely used security applications and protocols. SHA-1 hashing is also used in distributed revision control systems . SHA-1 not for security, but for ensuring that the data has not changed,[12], and to detect data corruption or tampering. Private and commercial organizations mainly use SHA-1 [11] This was designed by the National Security Agency (NSA) to be part of the Digital Signature Algorithm.

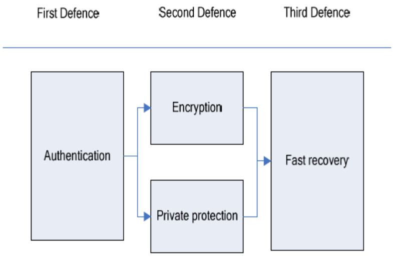
C.Layered Security :

Layer1: User authentication

Layer2 : User Data Encryption

Layer 3: User Data Recovery

Fig: multiple clouds data security model



CONCLUSION

Our scheme was developed to reduce the computational and storage overhead of the client as well as to minimize the computational overhead of the cloud storage server. this scheme applies not only to static storage data but to the dynamic also.

REFERENCES

[1] Sravan Kumar R and Ashutosh Saxena “Data Integrity Proofs in Cloud Storage” 978-1-4244-8953-4/11/$26.00 © 2011 IEEE.

[2] Meiko Jensen ,Jorg Sehwenk et al., “On Technical Security,Issues in cloud Computing ”IEEE International conference on cloud Computing,2009.

[3] M.Jensen ,N.Gruschka et al., “The impact of flooding Attacks on network based services”Proceedings of the IEEE International conference on Availiabilty,Reliability and Security (ARES) 2008.

[4] Armbrust ,M. ,Fox, A., Griffth, R., et al “Above the clouds: A Berkeley View of Cloud Computing” , UCB/EECS-2009-28,EECS Department University of California Berkeley, 2009.

[5] Kaufman, L. M. (2009)."Data Security in the World of Cloud Computing." IEEE Security and Privacy 7(4): 61-64.

[6] Sloan, K. (2009)."Security in a virtualised world." Network Security 2009(8): 15-18.

[7] Hewitt, C. (2008). "ORGsfor scalable, robust, privacy-friendly client cloud computing." IEEE Internet Computing 12(5): 96-99.

[8] Vieira, K., A. Schulter, et al. (2009). "Intrusion Detection Techniques in Grid and Cloud Computing Environment".

[9] Mowbray, M. and S. Pearson (2009). A client-based privacy manager for cloud computing, ACM.

[10] “Amazon Web services: Overview of Security processes “ September 2008.

[11]^ *a b* Schneier on Security: SHA-1 Broken

[12] ^ http://debugmo.de/?p=61 Debugmo.de "For verifying the hash (which is the only thing they verify in the signature), they have chosen to use a function (strncmp) which stops on the first null byte – with a positive result. Out of the 160 bits of the SHA1-hash, up to 152 bits are thrown away."

Ku.Swati G. Anantwar ,ME II year, CSE Branch , H.V.P.M C.O.E.T Amravati.

Karuna G. Bagde , Asst.Professor, CSE Branch ,H.V.P.M C.O.E.T Amravati

.