iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** III **Month of publication:** March 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Survey on data storage security in Cloud

Navalpreet Kaur[#1], Harsimran Singh[*2]
[#]*Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab (INDIA)*

Abstract— Cloud is studied as the future of information technology. Cloud computing assigns to online network based computing where virtual servers serve software, hardware, infrastructure, devices and platform. Users can make them use on a pay-as-you-use basis. Many organizations are strict about the acceptance of cloud computing due to their worry toward the security of the data storage. Therefore, issues related to the security of data in the cloud have become very critical. In this paper, various data storage security issues in cloud have been discussed. Also, some already existing schemes for responding the problem of data storage security have also been discussed. Finally, a comparative study has been carried out among the schemes with respect to data storage security issues.
Keywords— Cloud computing, Data storage, security, Database, servers

I. INTRODUCTION

Now-a-days, Cloud Computing is the most popular concept in Information Technology (IT). The Cloud Computing derived from the word cloud like structure that was used to represent online network. It is the approach of computing where thickly scaled IT related capabilities are provided as a service across the online network to various external customers that billed by consumption. It is a new technology, where users need not to have their own hardware, software, storage space etc. All things will be served by the cloud itself. Google, Microsoft, Yahoo, IBM and Amazon have started serving cloud computing services. Amazon is the lead in this field. Cloud Computing serves the facility to access shared resources and common infrastructure providing services on demand over the network to perform operations. But the location of the users from where to access the data is not known. With the help of Cloud Computing, users can access their databases from anywhere in the world only if they connected to the internet. Today's world depends on cloud computing to store their public data as well as personal data. That data may be required by them or others at any instant of time. As a result, data security in cloud computing has required lots of attention from the research society.

II. OVERVIEW

Cloud Computing is a phrase that involves introduced services over the online network. Cloud Computing serves to access shared resources and common infrastructure providing services on demand over the network to carry out operations that meets changing business requirement. Cloud Computing allows customers and businesses to access applications without installation and can access their personal files at any computer with online access. The location of physical resources and devices being accessed that are mostly not known to the user. It also provides facilities for users to develop, use and manage their applications in the cloud, which requires virtualization of resources that maintains it. There are a little differences between old hosting and cloud computing. Cloud Computing is use on demand. All the services in the cloud are handled by the provider. Users can fulfil the amount of services they take and can log on to the internet from any computer in the world.

A. Delivery Models

The architecture of Cloud Computing can be categorized according to the three types of delivery models [1][2]:

1) Infrastructure as a Service (IAAS): Consumers are allocated computing resources to run virtual machines that consist of operating systems and applications that are provided as an on-demand service. The best example of IAAS is Amazon.com's Elastic Compute Cloud (EC2) service. The requirements of security beyond the basic infrastructure are carried out mainly by the cloud consumer.
2) Platform as a service (PAAS): Cloud consumers are allowed to write applications that run on the service provider's environment. It is a model of service delivery where the computing platform is served as an on-demand service upon which applications can be developed and used. Google Apps engine is an example of PAAS. The requirements of security are split between the cloud provider and the cloud consumer.
3) Software as a service (SAAS): Cloud consumers are provided with various software applications that run over the online network. Google Docs programs are an example of SAAS. The requirements of security are carried out generally by the cloud provider.

B. Deployment Models

Deployment models broadly characterize the management and disposition of computational resources for delivery of services to consumers, as well as the differentiation between classes of consumers. Different deployment models are [2] [3] [4][5]:

1) Public Cloud: A public cloud is the infrastructure and computational resources that it contains the general public over the online network. It is purchased and managed by a cloud provider that delivering cloud services to consumers and, by definition, is external to the consumers' organizations.
2) Private Cloud: A private cloud is the computing environment that is managed only by specific organization. It may be

operated by the organization or by a third party, and may be hosted within the organization's data center or outside of it.

3) Community Cloud: A community cloud is the combination of public and private clouds. It is quite similar to a private cloud, but the infrastructure and computational resources are only two or more organizations that have common privacy, security, and regulatory considerations, rather than a single organization.
4) Hybrid Cloud: Hybrid cloud is the complex model among all the deployment models. They involve two or more clouds (private, community, or public). Each cloud remains a unique entity, but is bound to the others through consistent that enable application and data portability among them.

III.DATA SECURITY

Data searching is done by many cloud consumers, which can give progress to serious security concerns in cloud environment. In this section, different issues related to data security in cloud computing have been described.
There is a critical need to securely store, manage, share and scrutinize enormous amounts of data. It is very important that the cloud should be secured enough to maintain the security of data. Exact physical localization of user data in virtual cloud atmosphere is among some of the prime challenges in cloud computing. The major security challenge with clouds is that the owner of the data may not have complete knowledge of where their data are stored. Data security involves encrypting the data as well as ensuring that suitable policies are imposed for sharing those data. There are numerous security issues for cloud computing. Some of the major data security issues are [1]:

1) Data Integrity: It is very important to maintain the integrity of data. The stored data in the cloud storage may suffer from huge damage occurring during the transition operations from or to the cloud storage provider. The risk of attacks from both inside and outside the cloud provider exists and should be considered.
2) Data Intrusion: Data Intrusion is another security issue that may occur with a cloud provider. If any intruder can advantage to access the account password, then he/she will be able to do anything and may be unwanted changes to the account's private documents. Undesirable changes of user data may begin due to intrusion.
3) Service availability: Service availability is another major concern in cloud services. It is mentioned in some cloud providers licensing agreement, that the service may be unavailable anytime due to some unexpected reason. If all the important documents of business are stored on the cloud and the cloud suddenly goes down, will it be coming back up with all of our important documents perfectly? It is also important to know whether the company with whom user is storing his critical information is economically safe and will not suddenly be lost taking all of the important information with it.
4) Confidentiality: Confidentiality of data is another security issue in cloud computing. The data should be kept secured and should not be disclosed to anyone at any cost. The users do not want their confidential data to be exposed to any service provider. But it is not always possible to encrypt the data before storing it in cloud.

5) Non- Repudiation: Non-repudiation is a major concern for data security. It assures the transmission of message between parties and gives the guarantee that someone cannot reject something. Non-repudiation is often used for signatures, digital contracts, and email messages. It ensures that a party cannot reject the genuineness of their signature on a document or the sending of a message that they originated.

IV.MITIGATION ALGORITHMS

In this section, data security algorithms have been discussed. Users are benefited by the cloud service providers, which may give growth to new security risks. Users store the data in cloud that are not aware of the physical location of the data. Attacks on the data storage can directly affect the security of the user's important data that include application data or sensitive data. Many cloud service providers serve storage as a service. They take the data from the users and store them on large data centers. The data stored in the cloud may get lost or change due to some security cracks that may be caused by various factors. Several algorithms for ensuring data storage security have been proposed.
Pradeep Kumar et al. [6] proposed a scheme for security in cloud computing using Hidden Markov model (HMM) [7] and Clustering. This scheme performs intrusion detection based on the probability of the behaviour. Clustering is used for serving only those data that are needed by the user. Cluster gives a solid view of data storage in cloud environment. It is used for fraud detection and is utilized to reduce the data seeking time [8]. In [6], HMM model auditors the user behaviour continuously and informs the administrator with the help of filter network. Firewall gateway is used to block the SSH service, which is used for increasing access to the server from anywhere in the world. It simply rejects the packets that come from any IP address which is not in knowledge. It is the duty of the system administrator to keep the data safe in cloud. The system administrator immediately takes the action if any malicious activity occur that are related to data mining is find by the filter network after getting input from the HMM. In this scheme, the cloud computing environment has an item attached with it. It can help the proposed model to backup as well as recover the data if any harm occurs. It can be easily attached with any other cloud environment if needed. As a result, the load on the servers of a cloud could be decreased. So, an intruder can be easily captured even if he has the stolen ID and/or password. This environment has classified the data into two kinds. One kind includes normal data searching from the database. On the other hand, Second kind involves searching sensitive data that must be kept protected from unauthenticated access.
Shuai Han et al. [9] proposed a third party auditor scheme in cloud computing for ensuring data storage security. Third party auditor (TPA) gives trustful authentication for the user who stores their data in the cloud. TPA is capable than the cloud

consumers and makes every data in control to be accessed. Users cannot fully depend whether their data is safe on the cloud providers. TPA can analyses and clarify the data stored in the cloud on behalf of the users upon request. It can provide log information to the users. The authors have built a new architecture for cloud storage, where the third party auditor and the cloud service provider have been combined together. The traditional network architecture [10] consists of three entities which are users, Cloud service provider and TPA. Users have large data files to be stored in the cloud. They are active participants and can be individual consumers or organizations. Cloud service provider has sufficient storage space as well as computation resources for maintaining the data stored by the users. In this scheme RSA has been used to encrypt the data flow between servers in the advance cloud service provider. It uses Bilinear Diffie-Hellman algorithm for exchanging keys. Users and cloud service provider can communicate with each other by using a message header without a third party auditor. Each of the cloud storage servers can add, identify and update the message header for users whose copy will be sent to the trustful organization server for the first time. A pair of keys is allocated to each user for accessing the cloud. In this scheme, users add a message header before sending it to the cloud. The data packets are encrypted with the allocated keys using RSA algorithm. Trustful organization servers which are managed by trustful organization's that perform as a supervisor for every access keys in cloud service provider, comprises of few number of servers. The users and cloud service provider cannot get any authentication information from trustful organization without a certain module. The Trustful organization server is responsible for maintaining all the keys which are stored in cloud storage servers.

According to Mohammed A. Alzain et al. [11], shifting from single cloud to multi-cloud is very important for ensuring the security of user's data. Authors suggested that, there are three main security factors of data storage (data integrity, data intrusion and service availability) that needs to be considered as the major concern for cloud computing. They have introduced a new model called Multi-clouds Database Model (MCDB). A technique named Shamir's secret sharing algorithm [12], which is based on polynomial interpolation has been incorporated in the scheme. According to the algorithm [12], if a data D is shared into n pieces, in that D is easily reconstruct able from k pieces, but even complete knowledge of k-1 pieces admits absolutely no information about D. The authors have suggested that Cloud Computing should not end with one cloud. In their work, they have compared Amazon cloud service that is single cloud with their proposed multi-clouds model. This model assures the security and privacy of data in multi-clouds using multi shared technique instead of single cloud. The data is depicting among various clouds by using secret sharing approach. The operations between the clients and the cloud service providers are managed by Database Management System (DBMS). Data is being stored by cloud service providers after being divided by MCDB. Division of the data depends upon the number of cloud service providers.

V. COMPARISON

This section presents a comparison among the data storage security algorithms, discussed in the previous section, with respect to the issues discussed in section III.

TABLE I
VARIOUS SCHEMES AND THEIR SUPPORT OF FEATURES

Security Risks	[6]	[9]	[11]
Data Integrity	Yes	No	Yes
Data Intrusion	Yes	Yes	Yes
Service Availability	Yes	No	Yes
Data Confidentiality	Yes	Yes	Yes
Non-Repudiation	Yes	Yes	No

Table 1 details of the various security risks that have been resolved ([6],[9],[11],[13]) by yes in the field of the table. In [11], few major security issues such as, data integrity, data confidentiality, service availability and data intrusion have been solved. Data integrity is maintained in [11]. In [11], data confidentiality is maintained by storing the data in multiple cloud service providers and clouds respectively by using Shamir's secret sharing approach [12]. Intrusion can be prevented in [11], because if the hacker hacked the password from one cloud service provider, they still have to hack the third cloud service provider for its password. It is very hard for a hacker to retrieve the password from all the cloud service providers. Service availability is guaranteed in [11], as the data is distributed in different cloud service providers and clouds respectively. As a result, the risk of data loss gets minimized. The problem of non- repudiation have not been taken care of in the schemes [11].

In [9], data intrusion problem have been solved, as, only the authenticated users can have access to the information. The users are given permission using the authentication modules which are given by the trustful organizations. Any intruder cannot have access to the valuable data without proper permission. Confidentiality is achieved as the data file is encrypted with user secret key. After encryption, the data file is sent for storing in the cloud service provider. Service availability and data integrity problems are not considered in this scheme. If the CSP goes down, then the data cannot be recovered. The risk of attacks exists in the Cloud storage provider. The data stored in the cloud may suffer from any damage occurring during transition from or to the cloud storage provider. But, no such measures have been taken in this scheme to ensure the integrity of data. Non-

repudiation is not possible here, because when a user sends a request, then the user's information is added in cloud storage server's user list.

In [6], data integrity is maintained by using HMM model which monitors the user behavior continuously and informs the system administrator. As a result, the data is safe in cloud environment. HMM model can be used to detect intrusion based on the probability of behavior [7]. Firewall gateway is also used in this scheme [6] to prevent data intrusion. Firewall gateway can block any anonymous port with a live IP address which is not inside its knowledge. As a result, intrusion is not possible in this scheme. Data can be kept confidential using this scheme, as every traffic update is sent to the administrator via email notification. Service availability can be solved with the help of plug-in, which helps the environment to connect with another network. Non- repudiation is not possible as the firewall gateway contains the list of authenticated IP addresses.

VI. CONCLUSIONS

Data security is a very vital issue in cloud computing. In cloud data storage system, users store their data in the cloud cannot possess the data locally. Users are not aware of the physical location of their data. It is not clear how safe their data is and ownership of data is also unclear when these services are used. Cloud computing companies say that the data stored are completely safe. But, it is too early to comment on the reliability issues claimed by them. The stored data may suffer from damage that occurs during data transition operations from or to the cloud provider. Data are not always safe when they are stored inside cloud providers. For addressing these issues several algorithms have already been proposed and there is a huge scope for work in the area of data security in cloud computing.

REFERENCES

[1] Mohammed A. AlZain, Eric Pardede, Ben Soh, James A. Thom, "Cloud Computing Security: From Single to Multi-Clouds", Proc of the 45th Hawaii International Conference on System Sciences (HICSS), IEEE, pp. 5490-5499, 2012.

[2] Pengfei You, Yuxing Peng, Weidong Liu, Shoufu Xue , "Security Issues and Solutions in Cloud Computing", in Proc. of 32nd International Conference on Distributed Computing Systems Workshops(ICDCSW), pp. 573-577, 2012.

[3] Zaigham Mahmood,"Data Location and Security Issues in Cloud Computing", Proc. of International Conference on Emerging Intelligent Data and Web Technologies (EIDWT), pp. 49 – 54, 2011.

[4] T.M Bhraguram , M.S Sumesh, "Cyber Security Information Exchange Based on Data Asset De-coupling factor in Cloud Computing", IEEE Recent Advances in Intelligent Computational Systems (RAICS), pp. 089 – 095, 2011.

[5] Pavithra S., Badi Alekhya, "Implementing Efficient Monitoring And Data Dynamics For Data Storage Security in Cloud Computing", IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), pp. 139-143, Vol. 2, No. 1, 2012.

[6] Pradeep Kumar, Nitin, Vivek Sehgal, Kinjal Shah, Shiv Shankar Prasad Shukla and Durg Singh Chauhan, "A Novel approach for Security in cloud computing using Hidden Markov model and Clustering", *Proc. of Information & Communication Technologies (WICT)*, pp. 810-815, 2011.

[7] Abhinav Shrivastava, Amlan Kundu, S Surat, A.K. Majumdar "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable And Secure Computing, vol. 5, issue 1, pp. 37-48, 2008.

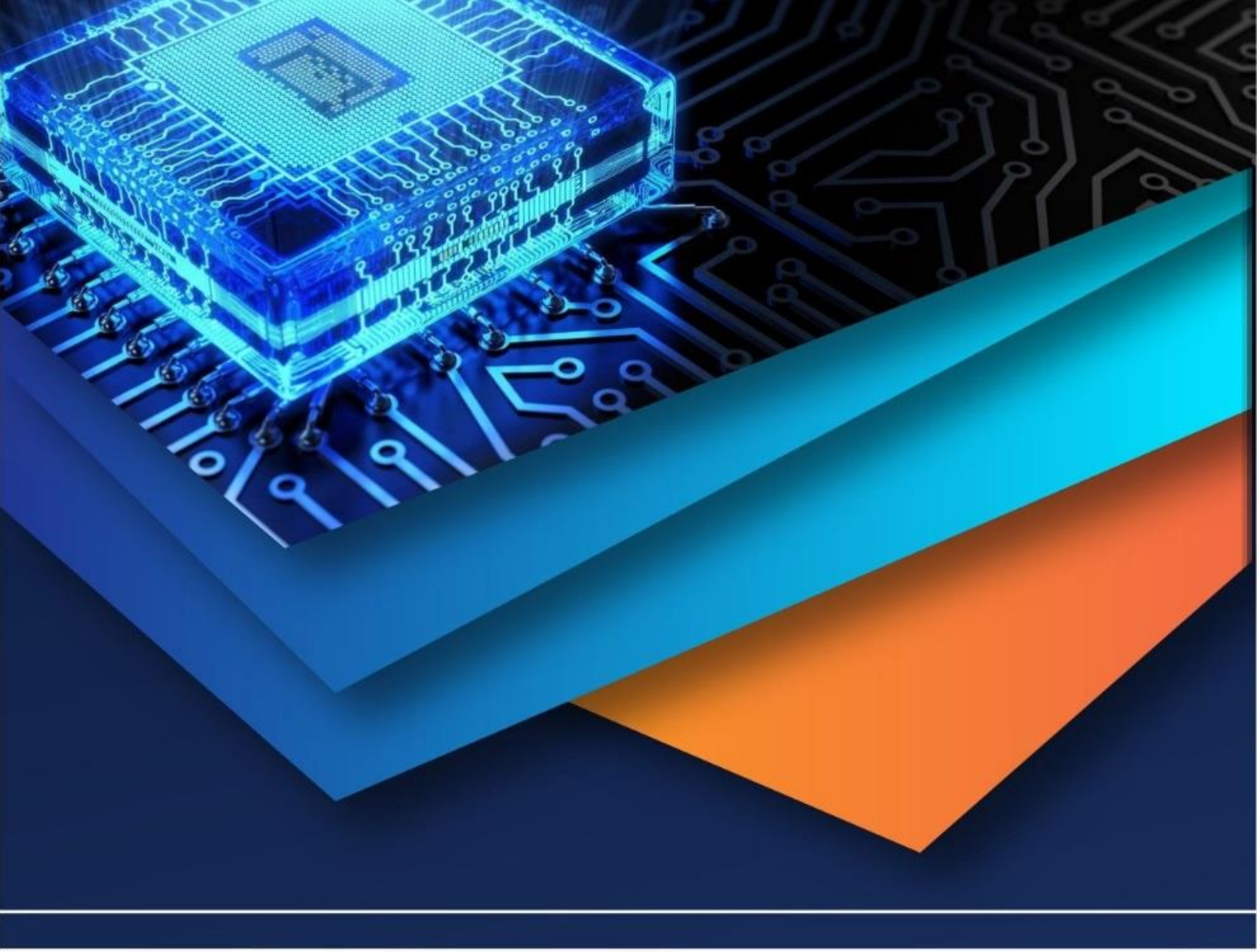
[8] Jeffrey W. Seifert, "Data Mining: An Overview", CRS Report for Congress, pp.1-16, 2004.

[9] Ian Foster, Yong Zhao, Ioan Raicu and Shiyong Lu, "Cloud Computing and Grid computing 360-Degree Compared", Grid Computing Environments Workshop, 2008, GCE'08 , pp. 1-10,2008.

[10] Qian Wang, Cong Wang, Jin Li, Kui Ren, Wenjing Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing", Proceedings of the 14th European conference on Research in computer security, pp. 355-370,2009.

[11] Mohammed A. Alzain, Ben Soh and Eric Pardede, "MCDB: Using Multi Clouds to ensure Security in Cloud Computing", Proc. of the 2011 IEEE 9th International Conference on Dependable, Autonomic & Secure Computing (DASC), pp. 784-791, 2011.

[12] A. Shamir, "How to share a secret", Communications of the ACM, Vol. 22, Issue 11, pp. 612-613, 1979..



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY