# Recovery based TPA in Cloud for Providing Security to Outsourced Data using CloudSim

Noorjahan.C
Department of Computer Science and Engineering
CEMP, Alappuzha
Kerala

Huda Noordean
Department of Computer Science and Engineering
CEMP, Alappuzha
Kerala

***Abstract*—Nowadays Cloud Computing became an inevitable thing in human's life. Both individual users and companies seek help of cloud for storing their data. When storing the data in the cloud user will have a fear about whether their data safe or not in cloud. So integrity of the data stored in the cloud is serious issue. To check the integrity of the data we propose a scheme that uses a trusted Third Party Auditor (TPA), who can verify the integrity of the data on demand. In the proposed system TPA does verification job for the user and also for Cloud Storage Server (CSS). So TPA helps users to find out the original data without any online burden to users and helps the CSS to replace the corrupted data with the original data***

**Keywords—** ***Cloud Security; CloudSim; Data Availability; Data Integrity; Third Party Auditor*.**

## I. INTRODUCTION

In recent years the cloud [1][2] has become an unavoidable thing in human's life. People have many data including sensitive data which faces security issue and also space requirement in user's personal computer or laptop. And also, when the users store their data in personal computer they always fear about whether their data will lose due to virus attack or any other problem. Because of these reasons the users always seek for a place, where their data are safe and available all time. So the cloud become popular when users want a secure data storage space. When the users started to use cloud for their data storage purpose, another big problem arises about the security of the data that uploaded in the cloud. Users of cloud fear about their data inside the cloud, ie whether their data can be assessed by any other people including the cloud service provider. So the user wants to ensure the integrity of the data in cloud.

For checking the integrity [3] - [6] of the data the user can send an integrity verification request to cloud and can verify the integrity. But it will be a burden to users, because this process is very time consuming. So many existing system uses a Third Party Auditor(TPA) to verify the integrity of the data. This TPA will be external to cloud, and trusted by both the user and Cloud Service Provider. When user want to verify the integrity of the data, he can send request to the TPA. Then TPA will send the request to Cloud Server for the user's data. Cloud Server send requested data to TPA. Then TPA will verify whether the data are modified or not and send result to user. The figure -1 shows the system architecture [3].

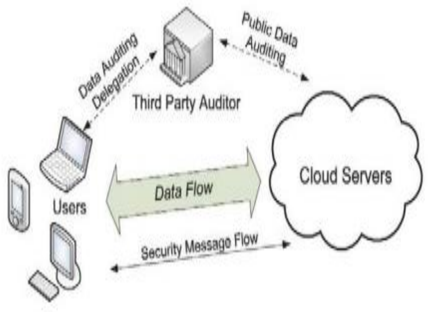


Fig. 1 System Architecture

In almost all existing systems TPA sends the data auditing result to the user. TPA has no further involvement in user's data. But in our scheme TPA helps users to retrieve original data. When the user's data are modified TPA will send result to the user. At that time user can further request to the TPA for data recovery. Thus TPA helps the user to get the correct data without any online burden. The verification of TPA also helps Cloud to replace their corrupted data with original data. This system is implemented using the CloudSim [7], which is a frame work for modelling and simulating cloud.

## II. RELATED WORKS

First, Big data [8] is the term that is used for large data storage. This Big data cannot handle with ordinary Data Base Management System. Today many researches are conducting on this big data management. There are many applications for big data storage as social network, health care, Education, manufacturing, real-time streaming data, Media, big sensor data, industry and data in the Internet, etc. So to deal with these data cloud is introduced as a more comfortable and efficient platform

**Armbrust et al 2010** [9] discusses very clearly about cloud computing. Cloud is a term used for a Datacenter which can provide both hardware and software as our demand. These cloud can be used by any registered user as a pay-as-you-go

**978-1-5386-1887-5/17/$31.00 ©2017 IEEE**

fashion. So by using the cloud, we can reduce hardware cost and software cost. Even though cloud has many advantages it has many disadvantages also. Issues such as security of data in cloud, availability of data, bugs in large distributed system etc. are main problems in cloud.

**Minqi Zhou et al** [10] discusses about the main security and privacy issues related to the cloud such as availability, confidentiality, data integrity, control and audit. So by considering these factors we can improve the security of the cloud.

**D. Zissis 2011** [11] discusses about major security problems that is present in the cloud computing. This paper proposes a trusted third party (TTP) to increase the security of the cloud. Trusted third party is a party which allow a secure communication between cloud and user.

**Cong Wang et al** [12] in their paper discuss about public auditable secure cloud storage. They propose a public auditing Third Party Auditor, who audit user's data on demand. So user can check the security of data without online burden.

**Kan Yang and Xiaohua Jia** [13] compare some existing cloud security schemes and analyses their performance. MAC-based methods, it is a type of hash function and used for checking integrity of data. Before storing the data owner compute MAC of each data block, and when they want to check the integrity they retrieve the random blocks and recompute the MAC and compare with the MAC which is saved in their system.

**Q. Wang et al**. [14] introduce a data auditing scheme which support fully dynamic data operation such as block insertion. This scheme also supports scalable and efficient public auditing in Cloud Computing and batch auditing.

**Cong Wang et al.** [15] propose a secure cloud which support privacy preserved public auditing. This scheme also allows TPA to audit multiple users simultaneously and efficiently. To achieve privacy-preserving public auditing, they integrate the homomorphic linear authenticator with random masking technique.

**Tamal Kanti Chakraborty et al**. 2013 [16] proposes Elliptic curve cryptosystem security scheme (ECCSS) to ensure the integrity of data that is stored in the cloud. This scheme makes use of Provable Data Possession to achieve dynamic data operations. It uses proof of retrievability (PoR) scheme to check integrity of data. This system uses ECC for key generation, that is for generating public key and private key, and by using these keys user can encrypt their data before store in to cloud.

**Mehdi Sookhaka et al**. 2014 [17] discusses about Remote Data Auditing (RDA) scheme and present a taxonomy of RDA approaches. RDA means a group of protocols that checks the correctness of data in cloud frequently, efficiently and securely without retrieving the data. This paper classifies the characteristics of remote data auditing approaches into five groups, namely security objective, security requirements, auditing mode, update mode and performance metrics.

**Liu et al.,** 2014 [18] Proposed a scheme ''Public Auditing for Dynamic Data with Fine-grained Updates (FUDPA)'' which allow the public auditability and also allow data dynamics. In this scheme a trusted third party is present for auditing of the data and allows fine grained updates. TPA in this scheme can verify whether the Cloud Service Provider (CSP) successfully do users fine-grained updates. So it supports public auditing on fine grained updates.

.

**Syed Rizvi et al** [19] discuss about insider attack in TPA, Although TPA has many advantages, sometimes untrusted TPA may convert in to a malicious user and may attack the user's data. To ensure the integrity of TPA this scheme propose an auditing which is done by CSP. Auditing is done by using SLA and time-released session keys.

## III. PROBLEM ANALYSIS

The existing schemes contain three participating parties: user, Cloud Storage Server and Third Party Auditor. All schemes use the TPA to check the integrity of the data. User sends a request to TPA to check whether his data are safe in cloud. TPA send request to Cloud Service Provider for specific data and verify its integrity. After the verification if the data are safe, the TPA will send reply to user that the data are not modified, otherwise send the data are modified. In all the existing schemes user gets a reply from TPA that his data is safe or not, but TPA does not help to recover the corrupted data. If there is any modification to user's data, he should get the original data.

## IV. PROPOSED SYSTEM

Proposed system introduce a Third Party Auditor, which can verify the integrity of data stored in the cloud and also helps users to get the original data if the verification result of data auditing is negative i.e., if data are modified. So by using this type of TPA user will always get his data without any online burden. Thus Recovery Based TPA improves the availability of data stored inside the cloud.

The system mainly contains five steps.

### *A. File Processing*

In this step, user selects the file which he wants store in the cloud and split the files in to several blocks. Here the file will segment in to 's' segments and each segment contains 'l' characters. Here 'l' is a predetermined value. So number of segments s can be obtained by using the equation

$$s = \text{total file size}/\ l \qquad (1)$$

After splitting file there will be 's' file blocks, we have to process each of this file blocks. Then each file block will encrypt using RSA public key encryption algorithm. This is an asymmetric key encryption algorithm, which has two keys for both encryption and decryption, i.e. public key and private key. While encrypting the system will create a public key and private key pair for each file block. So there will be 's' public key and 's' private key pair, and these keys will save in user's system.

After encryption this scheme will create a hash value of each file block using SHA-256 algorithm. So there will be a unique hash value or signature for each file block. If any of the characters in the file block changes, its signature also changes. These hash values will store in user's system and it can be used for further integrity verification.

File processing is also done after retrieving encrypted file. The encrypted file will be decrypted using RSA algorithm. For that keys are taken from user's system. After decryption each file block will merge to get original file.

### B. File uploading

After splitting and encrypting the file, user have to upload file in to cloud. For that user will first login in to the cloud to prove his identity by using user name and password. After that user uploads his encrypted file blocks in to cloud. At that time cloud will store each file blocks in different HardDriveStorage in a datacenter. To trace which file upload in to which HardDriveStorage the cloud will keep an index file. By storing each file block in to different HardDriveStorage, cloud can improve the security of data. Cloud is also store the replicated copies of data in to multiple datacenters. By doing so cloud can ensure the availability of data. That is if any data block in any datacenter modified or lost cloud can fetch the data block from another datacenter. So the user's data will be safe at all time.

### C. File retrieving

When user wants to retrieve the file he can send request to cloud. For that user login to cloud service by using user name and password. Cloud will check user's identity and will send encrypted file blocks to user.

### D. Integrity Verification

When the user wants to verify the integrity of the data, he will send a request to the Third Party Auditor. User will send the details of file which he wants to verify and also send the hash value of the file blocks. Upon receiving the request, the TPA will send a request to the cloud for the specific file. Cloud will check whether this TPA is genuine or not by verifying its ID. After identity verification, cloud will send the encrypted block of data to the TPA.TPA will perform SHA-256 algorithm to generate hash value for the encrypted blocks of file. Then TPA compare two hash values, that is the hash value received from the user and hash value created by TPA. If both of these values are same the TPA sends reply to user that his data are safe in the cloud. If not same TPA reply as data are modified. The Verification Algorithm for Data Integrity Verification by TPA is given below

***Verification Algorithm. (TPA)***

```
Genchallenge.
1. Start.
2. User authenticates with TPA.
3. User sends the file name which he wants to verify and the hash value of file segments.
4. TPA receive the request.
Genproof .
5. TPA send request to Cloud for corresponding file blocks.
6. Cloud sends the file to TPA.
7. For i=0 to s repeat the steps 8 and 9
8. TPA perform SHA-256 algorithm and create hash value.
9. Compare this hash value with user's hash value
verify.
10. If all the hash values are same, send reply to user that his data are safe, otherwise replies that data are modified
```

### E. Data Recovery.

In existing system, the TPA's job is only to send result of data integrity verification to user. There is no further role for TPA. But in proposed system user can further request for TPA to help for retrieving the original data. When receiving the request from user TPA will send message to Cloud Storage Server(CSS) that the data which the CSS given has been corrupted and then request for another copy of data. Cloud Storage Server has multiplicative copies of data they are stored in different datacenters. So when one data modified CSS will fetch data from another Datacenter and send to TPA. TPA will again calculate the hash value and compare that with user's hash value and send result to CSS. If the reply from TPA is negative, then the CSS will fetch another copy of data from its datacenter and send to TPA for verification. This process will continue until the TPA gives a positive reply i.e., until find out the exact copy of data. If the reply is positive, then cloud will replace the all the corrupted data with original data and send the original data to the user.

The algorithm for Data Recovery is given below

***Data recovery Algorithm.***

```
1. Start.
2. Client send request to TPA to get original data from cloud.
3. TPA will send message to cloud storage server(CSS) for another copy of data.
4. CSS will fetch data from another datacenter and send to TPA.
5. TPA will again calculate the hash value
6. Compare that hash with user's hash value
7. After find out original data cloud will replace the corrupted data with original data
8. Send the original data to the user.
9. Stop
```

The Fig-2 shows flowchart of working of the Third Party Auditor

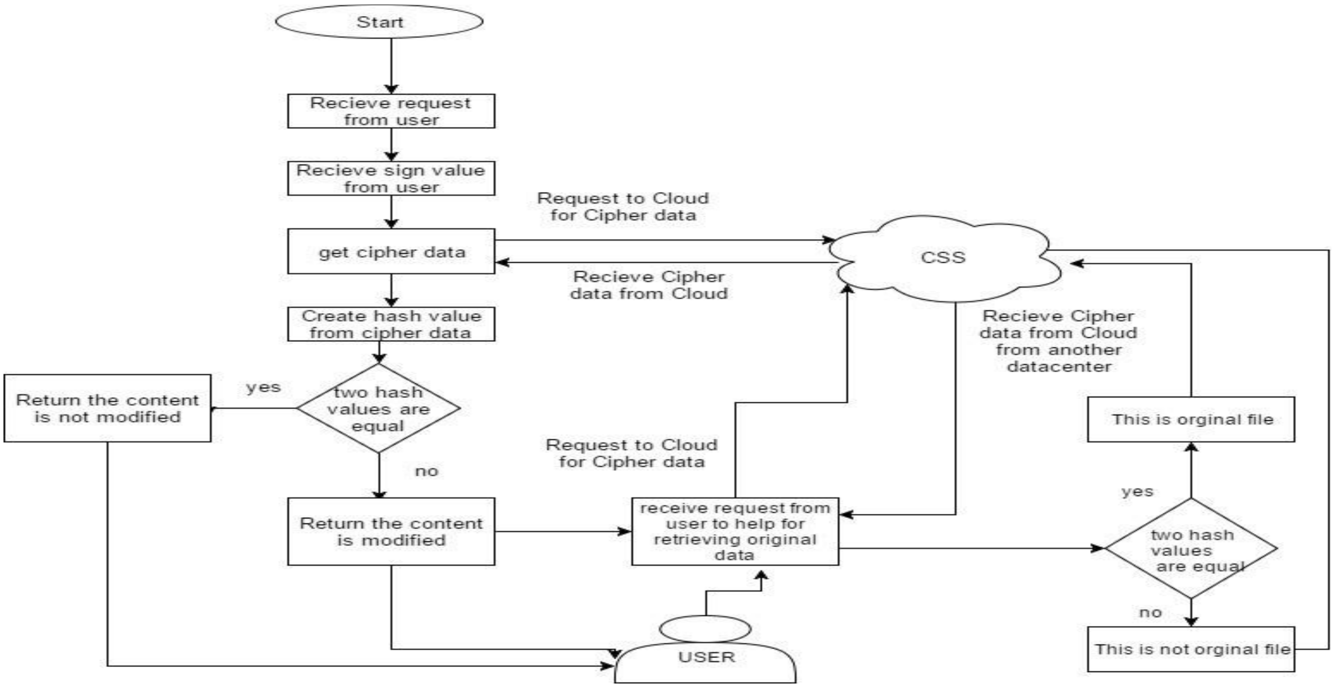

Start
Recieve request from user
Recieve sign value from user
get cipher data
Request to Cloud for Cipher data
Recieve Cipher data from Cloud
CSS
Create hash value from cipher data
two hash values are equal
yes
Return the content is not modified
no
Return the content is modified
Request to Cloud for Cipher data
receive request from user to help for retrieving original data
Recieve Cipher data from Cloud from another datacenter
This is orginal file
yes
two hash values are equal
no
This is not orginal file
USER

**Fig 2: Working of Third Party Auditor**

## V. IMPLEMENTATION

This system is implemented using CloudSim. Which is framework for developing simulating cloud architecture. By using CloudSim we can create a virtual cloud, and this cloud can do anything which actual cloud do. Implementation can be divided to five steps.

### A. *User Side File Processing And Uploading*

User select a file which he wants to upload in to the cloud. System will split file in to several blocks and encrypted it using RSA, then create a hash value and save it in user's system. The splitting, encryption and decryption are implemented using java code. After these processes the user sends each encrypted file blocks to cloud.

### B. *Cloud Side File Uploading*

In this step we have to create cloud using CloudSim. By using this toolkit, we can create any number of virtual machine and any number of datacenters. When CloudSim start we will create a number of datacenters. In that datacenters we want to store replicated copies of user's data. In reality the data centers contain many HardDriveStorage. So by using the CloudSim we can create many Hard Drive Storage. In this system an extended class of HardDriveStorage named CloudHardDriveStorage is used to store users file blocks. We can create any number of CloudHardDriveStorage, so that we can add users encrypted file in to different CloudHardDriveStorage for providing better security to user's data. In CloudHardDriveStorage cloud store the data as CloudFile.

When user requests to cloud for storing their data, Cloud will create as many CloudFile as the datablocks.ie, if there are s file blocks cloud will create s cloud file, and add data in each file block to one CloudFile. So there will be s CloudFile. After that each CloudFile should be added to CloudHardDriveStorage. For that cloud have to select a CloudHardDriveStorage, there is enough space to store the data by the instruction given below:

```
file.getSize() + hd.getCurrentSize() < hd.getCapacity())
```

hd means HardDriveStorage. If there is enough space, then the file store in to that CloudHardDriveStorage. At that time cloud also create an index file to keep track of which file stored in which CloudHardDriveStorage. Then all the CloudHardDriveStorage are added to a linked list hdList. After that this hdList added to a Datacenter. Since cloud have to make many replicated copies, there will be many CloudHardDriveStorage and many hdLists. Each hdList added in to different datacenters.

### C. *File Retrieving.*

When user request to retrieve the file, cloud access the index file and fetch the corresponding file from the corresponding CloudHardDriveStorage and send to user. User receive the data, decrypt and merge to get orginal data

### D. *Third Party Auditing.*

In CloudSim we treated TPA as a user, who can audit another user's data. When a user wants to check integrity of data, it sends request to TPA.TPA retrieve corresponding CloudFile and create its signature using SHA-256 algorithm.

Then it will compare the signatures using instruction given below.

siguser == null ? siggenerated == null : siguser.equals(siggenerated)

After comparing the signatures, TPA will send result to user.

### E. *Data Recovery*

User send request to TPA to get original data. TPA retrieve data from another datacenter.

file =hdProtected.getCloudFile(fname);

Then make a signature of the file using SHA-256, then compare it with user's signature. If not same send result to datacenter. Upon receiving the message that data has been modified, the cloud will set error check value in corresponding data center to indicate that there is a security failure in corresponding datacenter.

When get positive reply from TPA, ie, if the data is correct, then cloud will replace all the data in corrupted datacenters, reset their error check value. When user request the data, cloud send the original data to user.

Figure 3 shows flow chart of implementation of the system using CloudSim

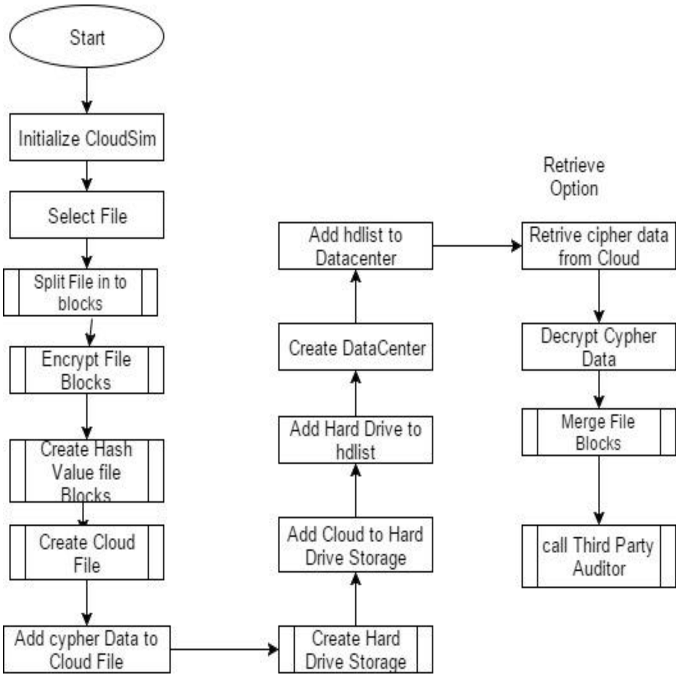
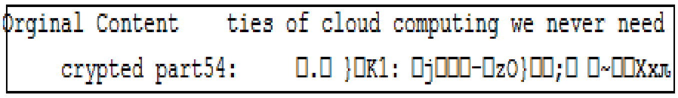


Fig.3. System Working.

## VI. Results and Discussion

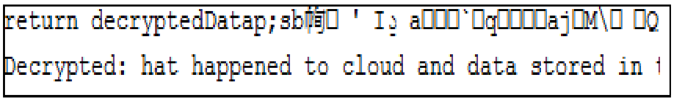
User file splitted and then encrypted using RSA encryption. Since RSA is an asymmetric encryption it is more secure. The figure-4 shows content in a file block and corresponding encrypted content



```
Orginal Content    ties of cloud computing we never need
   crypted part54:     □.□ }□K1: □j□□□-□zO}□□;□ □~□□Xxл
```

Fig. 4. Encryption

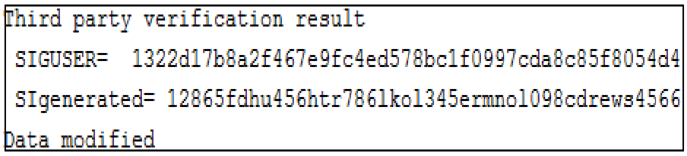
After encryption add data to cloud. When user wants to retrieve data, he retrieves it and decrypt it. Figure 5 shows decrypted data.



```
return decryptedDatap;sb□□ ' I; a□□□`□q□□□□aj□M\□ □Q
Decrypted: hat happened to cloud and data stored in
```

Fig. 5. Decryption.

If user send request to TPA for integrity verification TPA compare hash value it created using retrieved file from cloud with the hash value that user send. The figure 6 given below shows TPA verification result



```
Third party verification result
 SIGUSER=  1322d17b8a2f467e9fc4ed578bc1f0997cda8c85f8054d4
 SIgenerated= 12865fdhu456htr786lkol345ermnol098cdrews4566
Data modified
```

Fig. 5. TPA verification result

## VII. Conclusion

Cloud users are increasing day by day. Everyone wants to store their data in cloud very securely. For checking the integrity of the data an external Third Party Auditor is used. In proposed Scheme TPA verify the integrity of data on user's request. If data are modified, then the user can send request to the TPA to help for retrieving correct data. TPA send again request to cloud for the correct data, this process will continue until the verification request is positive. That is until TPA find out correct data. When TPA find out the exact data then it sends verification result to cloud. Based on this result CSS will replace all corrupted data in their datacentres with original data.This scheme help users to ensure that they will always get their data without any online burden.

## *Acknowledgment*

We thank members of Computer Science and Information Technology department of College of Engineering and Management, Punnapra, for their valuable support and feedback.

## *References*

[1] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, ''Cloud computing and emerging IT platforms: Vision, Hype, Reality for Delivering Computing as the 5th Utility,'' Future Gen. Comput. Syst., vol. 25, no. 6, pp. 599-616, June 2009

[2] .M.Armbrust,A.Fox,R.Griffith,A.D.Joseph,R.Katz,A.Konwinski, G.Lee,D.Patterson,A.Rabkin,I.Stoica,andM.Zaharia,''Aviewof cloud computing,'' Commun. ACM, vol. 53, no. 4, pp. 50-58, Apr. 2010

[3] J. Yao, S. Chen, S. Nepal, D. Levy, and J. Zic, ‘‘TrustStore: making Amazon S3 trustworthy with services composition,’’ in Proc. 10th IEEE/ACM Int’l Symposium on Cluster, Cloud and Grid Computing (CCGRID), 2010, pp. 600-605

[4] A. Juels and B.S. Kaliski Jr., ‘‘PORs: proofs of retrievability for large files,’’ in Proc. 14th ACM Conf. on Comput. and Commun. Security (CCS), 2007, pp. 584-597

[5] H. Shacham and B. Waters, ‘‘compact proofs of retrievability,’’ in Proc. 14th Int’l Conf. on Theory and Appl. of Cryptol. and Inf. Security (ASIACRYPT), 2008, pp. 90-107.

[6] S. Nepal, S. Chen, J. Yao, and D. Thilakanathan, ‘‘DIaaS: data integrity as a service in the cloud,’’ in Proc. 4th Int’l Conf. on Cloud Computing (IEEE CLOUD), 2011, pp. 308-315.

[7] Calheiros, R.N., Ranjan, R., De Rose, C.A.F., Buyya, R., “ CloudSim: A novel framework for modeling and simulation of cloud computing infrastructures and services” in Technical Report, GRIDS-TR-2009-1, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, 2009.

[8] Agrawal, D., Das, S. & Abbadi, A. E. “Big data and cloud computing: current state and future opportunities”. in Proceedings of the 14th International Conference on Extending Database Technology.

[9] .M.Armbrust,A.Fox,R.Griffith,A.D.Joseph,R.Katz,A.Konwinski, G.Lee,D.Patterson,A.Rabkin,I.Stoica,andM.Zaharia,‘‘A viewof cloud computing,’’ Commun. ACM, vol. 53, no. 4, pp. 50-58, Apr. 2010

[10] Minqi Zhou†, Rong Zhang§, Wei Xie†, Weining Qian†, Aoying Zhou, , ‘‘Security and privacy in cloud computing: a survey’’. 2010 Sixth International Conference on Semantics, Knowledge and Grids

[11] D. Zissis and D. Lekkas, ‘‘Addressing cloud computing security issues,’’ Future Gen. Comput. Syst., vol. 28, no. 3, pp. 583-592, Mar. 2011

[12] Cong Wang, Kui Ren, Wenjing Lou, and Jin Li. 2010. ‘‘ Toward publicly auditable secure cloud data storage services’’. IEEE Network 24, 4 (2010), 19–24.

[13] Kan Yang·Xiaohua Jia . ‘‘Data storage auditing service in cloud computing: challenges, methods and opportunities’’. Springer Science+Business Media, LLC 2011

[14] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, ‘‘Enabling public auditability and data dynamics for storage security in cloud computing,’’ IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 5, pp. 847-859, May 2011

[15] Cong Wang, Sherman S.-M. Chow, Qian Wang, Kui Ren, Member, IEEE, and Wenjing Lou ‘‘Privacy-Preserving public auditing for secure cloud storage’’. IEEE Transactions On Computers Vol:62 No:2 Year 2013

[16] Tamal Kanti Chakraborty, Anil Dhami, Prakhar Bansal and Tripti Singh ‘‘ Enhanced public auditability & secure data storage in cloud computing ’’ 2013 3rd IEEE InternationalAdvanceComputing Conference (IACC)

[17] Mehdi Sookhaka,n, Hamid Talebiana, Ejaz Ahmeda, Abdullah Gania, Muhammad Khurram Khanb ‘‘A review on remote data auditing in single cloud server: Taxonomy and open issues’’Journal of Network and Computer Applications 43 2014 121-141

[18] ] Liu, C., Chen, J., Yang, L. T., Zhang, X., Yang, C., Ranjan, R. & Ramamohanarao, K. 2014b. ‘‘Authorized public auditing of dynamic big data storage on cloud with efficient verifiable finegrained updates’’. IEEE Transactions on Parallel and Distributed Systems, 25, 2234 – 2244

[19] Syed Rizvi and Katie Cover, Abdul Razaque ‘‘Third-Party Auditor (TPA): a potential solution for securing a cloud environment’’. 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing