ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

Ensuring Data Integrity of Cloud Storage Services In Cloud Computing

Gunjal Yogita S.[1]

P.G. Student, Department of IT, Amrutvahini College of Engineering, Sangamner, Maharashtra, India[1].

ABSTRACT: Cloud is growing technology for enabling on demand access to a shared pool of configurable server resources with infrastructure that can be rapidly provisioned and released with minimal management efforts**.** Cloud Computing runs the application software and databases to the huge data centres in which the management of the data and services may be fully untrusted. This unique attribute poses a lot of new security challenges which have not been well understood. It focus on cloud data storage security, which has always been an important characteristics of quality of service. To assure the exactness of cloud users data in the cloud, work.In this paper we have proposed an effective and flexible distributed scheme with two relevant features and it opposing to its prior. In the existing system solutions in which information technology services are remain in the physical, logical, personal controls where cloud computing refers to the applications and services run on the distributed network using virtualized resources and accessed by the common internet protocol and networking standards. In this paper, focuses on the data integrity which is most challenging part for accurate data store .To ensure the exactness data of the cloud user in the cloud , we proposed some cloud services for the data integrity when cloud users data outsource sensitive data for sharing on cloud server. In order to address this is new challenge problem and later achieves a secure and dependable cloud storage services. In this proposed system, our scheme achieved the integration of storage correctness verification and error localization that is the identification of the misdemeaning server .For this extensive data integrity and analysis shows that the proposed scheme is highly efficient and resilient against the internal and external attacks.

KEYWORDS: Cloud user, Cloud, Cloud server, TPA, Data integrity

I. INTRODUCTION

In recent years, internet becomes necessity of the users. Users rely on remote storage instead of keeping any local copy and access as per requirement. Different cloud service provider provides a remote storage with cloud. Cloud computing is the type of computing that relies on sharing computing resources rather than having local server or personal devices to handle applications. In the cloud computing ,the name of cloud it refers to the internet so in short cloud computing means this is the internet based computing. In the cloud computing uses most of processors combine with software as a service in the computing architecture. These are transforming data centres into shared pools of computing service on a large scale. Combining a set of existing and new technology from research areas such as virtualization. In the cloud computing infrastructure resources are provided as services over the Internet. The increasing number of online services such as Amazon, Yahoo!, Google, gmail, etc.Moving data into the cloud offers great convenience to users since they don't have to care about the complexities of direct hardware management. The pioneer of cloud computing vendors, Amazon Simple Storage Service (S3), and Amazon Elastic Compute Cloud (EC2) [1] are both examples. While. These services are most important for storing and maintaining lots of valuable user data. For example uses of this storage include online backup, email, audio, video and users personal data. Transferring data into the cloud and it offers more convenience to since they don't have any care about the complexities of hardware management. There are two well known examples of cloud computing vendors that is amazon simple storage services and amazon elastic compute cloud. This cloud computing is internet based computing. provides large amount of storage space and customizable cloud computing resources ,but at the same time it eliminates the responsibility of local machines for data maintenance. For the result cloud user providers that is CSP for data integrity[3][4].

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

Juels and Kaliski Jr. [10] described in previous system a formal "proof of retrievability" model for ensuring the remote data integrity. Their scheme combines spot-checking and error correcting code to ensure both possession and retrievability of files on archive service systems.On the one side, infrastructure as a service of cloud computing are more powerful than the personal computing, extensive range of both internal and external attacks for data integrity still exits. In the currently Wang et al. [2] gave a study on many existing solutions on remote data integrity checking, and discussed their pros and cons under different design scenarios of secure cloud storage services. As remote storage fully managed by CSP, users keep faith on service provider and upload important data on cloud server without worrying about the security concern. Users are not aware about different backend security threats. To ensure data correctness on the cloud is being put at risk due to different reasons. First of all, although the infrastructures under the cloud servers are much more reliable and powerful compare to personal's computing devices but they are still facing an issue of threats for data integrity Also most of the cloud service provider behave unfaithfully towards the users for their outsource data.

Characteristics of the cloud computing:
1)On demand self services: In the on demand self services cloud client access cloud services through online control panel and manage their own computing resources.
2)Broad network access: In this broad network access we can easily access services anywhere, anytime.
3)Resource pooling: It means that customer draw from a pool of computing resources usually in the remote data centre.
4)Rapid elasticity: In this rapid elasticity means that it is computing terms for ability to provide scalability services.
5)Measured Services: In this measured services, aspect of the cloud services are controlled by the cloud provider.

Cloud Computing services:

- Software as a services: In this software as a services user can access both resource and applications .
- Infrastructure as a services: In this infrastructure as a services cloud client completely outsources the storage and resources such as hardware and software that they need
- Platform as a services: In this platform as a services cloud client give subscriber access to the components that they require to develop and operate application over the internet

Paper is organized as follows. Section II describes Cloud storage services architecture, system model adversary model.. Section III presents experimental results showing result analysis .Finally, Section IV presents conclusion.

II. RELATED WORK

The assurance that data is accurate, correct and valid. Accuracy and consistency of stored data, indicated by an absence of any alteration in data between two updates of a data record. Data integrity is imposed within a database at its design stage through the use of standard rules and procedures, and is maintained through the use of error checking and validation routines. Exact duplication of the sent data at the receiving end, achieved through the use of error checking and correcting protocols. Assurance that the data are unchanged from creation to reception.

1. Process to maintain data integrity depends on:
    a) Collection (accurate representation)
    b) Data transfer (accurate recording and transfer of data)
    c) Storage and Security (preventing loss of data)
    d) Sharing of Data
    e) Use of data (analysis)

In this paper, we propose an effective and flexible distributed storage verification scheme with explicit dynamic data support to ensure the correctness and availability of users' data in the cloud. We rely on erasure correcting code in the file distribution preparation to provide redundancies and guarantee the data dependability against byzantine servers [9], where a storage server may fail in arbitrary ways. This construction drastically reduces the communication and storage overhead as compared to the traditional replication-based file distribution techniques. In this system, we considered three different entities are as cloud user, TPA, cloud server . Ateniese et al. [10] described a PDP scheme that uses only symmetric key-based cryptography. This method has lower overhead than their existing scheme and permitted for

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

block modification and deletion, and appends to the stored file, which has also been supported in our work. However, their scheme focuses on single server scenario but it does not provide data integrity guarantee against server failures. Recently, the importance of ensuring the remote data integrity has been highlighted by the following research works under different system and security models[10][11][12] These techniques, while can be useful to ensure the storage correctness without having users possessing local data, are all focusing on single server scenario.

In fact, in the long time cloud client data in the cloud is not Prohibits the direct approval of traditional cryptographic primitives for the intension of data integrity protection. so therefore the verification of cloud storage exactness must be conducted without clear knowledge of the complete data files[5][6].Sometimes this cloud storage is not only a third party data accessed but also regularly updated by the cloud client including dynamic operations like deletion modification insertion and appending etc. This cloud computing is internet based computing provides large amount of storage space and customizable cloud computing resources, but at the same time it eliminates the responsibility of local machines for data maintenance. For the result cloud user providers that is CSP for data integrity[7].On the one side, infrastructure as a service of cloud computing are more powerful than the personal computing, extensive range of both internal and external attacks for data integrity still exits. They may be useful for quality-of-service testing [8], but does not guarantee the data availability in case of server failures There are two models are as follows : System model and adversary model. A descriptive cloud storage service network architecture is shown in fig 1.

System Model: In system model, there are three different entities can be identified as follows:

1. **Cloud User** : Cloud user has data to be stored in the cloud. It relies on the cloud for data storage and computation. Cloud user can be individual customers
2. **Cloud server:** This cloud server managed by cloud service provider(CSP) to provide data storage service. It has significant storage space and computation resources
3. **Third-Party Auditor(TPA)**: This TPA is an optional entity, This third party auditor protects from the byzantine failures, this byzantine failure means hide that errors from the cloud client for its own purpose. TPA is the system or personal having knowledge and expertise of data integrity.

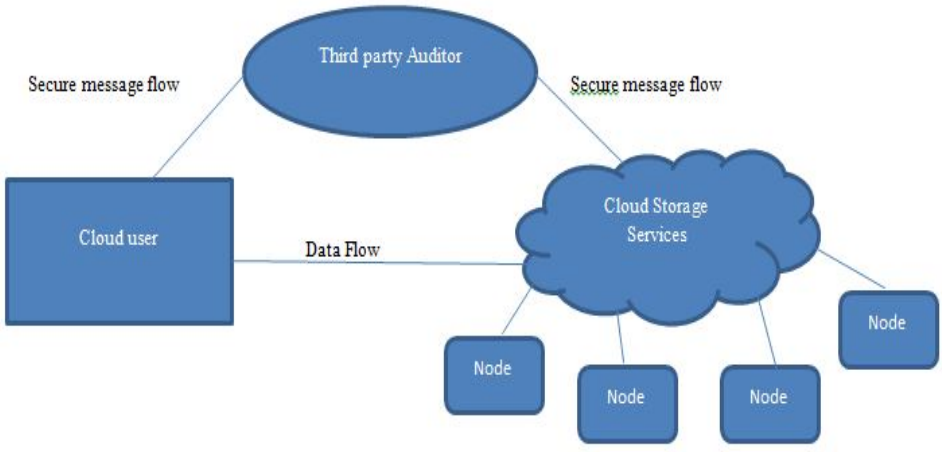


Fig1.Cloud storage services architecture

Adversary Model: In the adversary model , from cloud user viewpoint this adversary model has to capture all types of attacks toward his cloud data integrity. For reason cloud data do not reside at cloud user local site but at CSP address domain .These attack can come from two different sources:

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

1. I**nternal attacks.** In internal attacks, a cloud storage provider can be self-interested, malicious. In this internal attack malicious employees at cloud user, malicious employees at cloud provider, cloud provider itself. For example malicious attack ,it is an active attack.
2. **In external attacks:** In this external attack, data integrity attacks may come from outsiders who are beyond the control domain of cloud service provider. For example network attack, it is an passive attack.

For example cloud storage services appear outsource of sensitive data from time[9],but on the other side cloud user may not keep any local copy of outsource sensitive data, but there are various reasons for cloud service provider to behave faithlessly towards the cloud client regarding of their outsource data. In order to achieve the assurances of cloud data integrity and availability and enforce the quality of cloud storage service, efficient methods that enable on-demand data correctness verification on behalf of cloud users have to be designed. A. Juels and B.S. Kaliski[13] described the fact that users no longer have physical possession of data in the cloud prohibits the direct adoption of traditional cryptographic primitives for the purpose of data integrity protection. Hence, the verification of cloud storage correctness must be conducted without explicit knowledge of the whole data files [13][14][15]. Meanwhile, cloud storage is not just a third party data warehouse. The data stored in the cloud may not only be Considering all security aspect TPA connect to the server with predefined ports and IP address with three way full duplex socket programming connection. Based on the TPA request both system connect and communicate each other. This improves the security parameter for the communication between TPA and cloud.

S={x, e, i, o, f, DD, NDD, success, failure}
Let S be the solution perspective of the class
x= Initial state of the class Initialize ()
x= {Initialize ()} sets the default values for all variables.
Input i =(Y1,Y2)
Y1= {{U}{V}{F}{σ}}
DD=deterministic data it helps identifying the load store functions or assignment functions.
NDD=Non deterministic data of the system S to be solved.
Success-desired outcome generated.
Failure-Desired outcome not generated or forced exit due to system error.
Set of _k' cloud users U = {u1, u2, u3, u4……. ..uk}
Set of _m' cloud servers V = {v1, v2, v3, v4……. ...vm}
Set of files on cloud storage F = {f1, f2, f3, ……. fn}
Set of file tags σi = { f+t+i+u+fd+un++ic+uf+d+st+md5},
i ϵ (1, n)
f= File Name..
t= File type.
fd= File data
u=File User ID.
un= User notified.
md5= md5 hash value
I=new value in database
ic=Integrity check,
M=New value database
uf = Uploades files,
LI= List Of files.
d= Generate Challeges
st= Detail info of modified files.

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

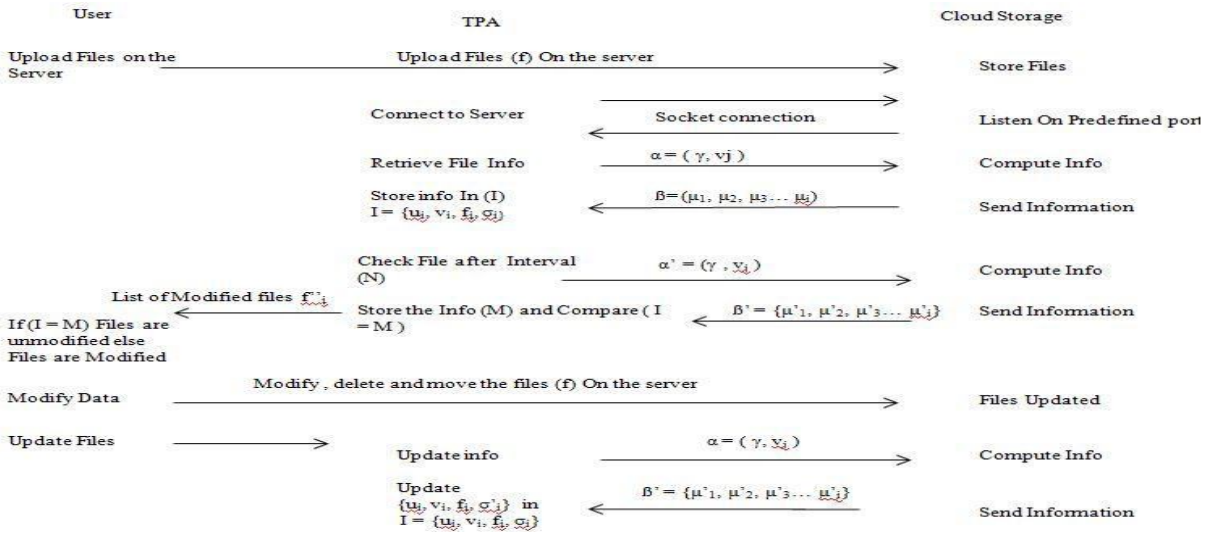


Fig 2. Proposed structure for data integrity

A. Initiate

This is first stage of algorithm. In this step TPA send a init () request to the client program install on cloud servers which include paths and user information of the files. Once receiving init request server compute the stats and an information for all files mentioned in the init request. Initialize () α = (γ, vj) Where, γ ϵ n and vj is a jth cloud server.(γ is set or path of (n) files and vj is cloud IP address) vj cloud server produces ß= (μ1, μ2, μ3, μ4… μi)Where, μi comes from (f1, f2, f3, f4, f5…fn) consists of pair (fi, σi). I = {ui, vj, fi, σi} Where, ui is user, vj cloud server and σi consist of signature tag of file fi .TPA store the received values in (I) database

B.Update

This step required when user uploads/modify the files on cloud server. Users inform TPA to update information.TPA send Query Update α = (γ , vj) Where γ ϵ n‘ and vj is a jth cloud server. n‘ updated files. Set of tags σ‘i = { f+t+i+u+g+b+d+s+p+pr+md5} i ϵ (1, n‘) where σ‘ updated files tags Number of files F = {f1, f2, f3, f4, ……. f‘n} Cloud server produces ß‘ = {μ‘1, μ‘2, μ‘3… μ‘iWhere μi comes from (f1, f2, f3, f3…fn‘) consists of pair (fi, σ‘i) TPA add/replace the ß‘ values {ui, vi, fi, σ‘i} in I = {ui, vi, fi, σi} I = {ui, vi, fi, σi}

C Data Integrity

Schedule of periodic verification phase of the scheme cloud users need to specify the scan interval of the integrity protocol. User can keep the auditing interval minimum but it will consume computing resources of the cloud and indirectly increase the cost. Basically it will be few hours to users get data integrity report as early as possible. So that user can take the necessary action to get modified file restored. Frequent auditing process would lead to a waste of network bandwidth and computing resources of TPA, CSPs, and Clients,in this phase TPA send the check request to the cloud servers, based on request client program on the cloud servers compute the file stats and information for the paths or directory in TPA request and sent result to the TPA system (M). Result of the files stats and hash are in the form of file or database. TPA system stores the new database (M). TPA system is having two databases. Initial database (I) and one after check request (M). TPA system compares the stored databases (I) and (M) and concludes the results based on the comparisons and send list of modified files to users. As TPA system contains files stat and hash value of the files it not possible to get files contents. It preserves the privacy of user‘s data from keeping copy information outside the cloud server. Also in this scheme I am utilizing fully automated TPA so it‘s not necessary to mask the results from client program. Initial values I = {ui, vi, fi, σi}Where, ui user, vi cloud sever IP, μi = (fi, σi) file name with file status.Interval to check integrity (N)
Set of file tags σi = { f+t+i+u+fd+un++ic+uf+d+st+md5},i ϵ (1, n‘)

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

where σ‘ updated files tags Number of files F = { f1, f2, f3, f4, ……. fn‘} TPA to cloud server Query Check α‘ = (γ ,vj) Produces ß‘ = {μ‘1, μ‘2, μ‘3… μ‘i} Where μi comes from (f‘1, f‘2, f‘3…….f‘n) TPA store the received ß‘ values {f‘i, σ‘i } in database (M) along with user and server details.
M = {ui, vi, f‘i, σ‘i }
TPA Search M {ui, vi, f‘i, σ‘i } in to the database I {ui, vi, fi,
σi} If M {ui, vi, f‘i, σ‘i } ϵ I {ui, vi, fi, σi}

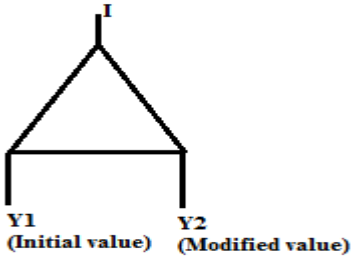


Fig3.Compare results

Success If M {ui, vi, f‘i, σ‘i } ≠ Search result I {ui, vi, fi, σi}
Results: Files modified lists (f‘i) Else M {ui, vi, f‘i, σ‘i } = Search result I {ui, vi, fi, σi}
Results: Uploaded/removed/modified files
Failure Desired results are not generated.
f= {update ( ), Check integrity ( )}

III. EXPERIMENTAL RESULTS

In this experimental results,here in the below Fig 4. login window, these is the first form when we run our project. The user should enter proper id and password, so that he will be authorized user and can access the system for its further processing. Entering the correct login id and password and then after clicking on login button, here we get users side form in which he can upload normal and CSV file. if user want to update that file then he can update it and generate challenges.

Fig 4.User login page

In below fig 5, after login user can upload any file. It may normal or CSV file.After uploading the file user can check the integrity.If there is no change in the file then he will get the message “Integrity check successful”

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

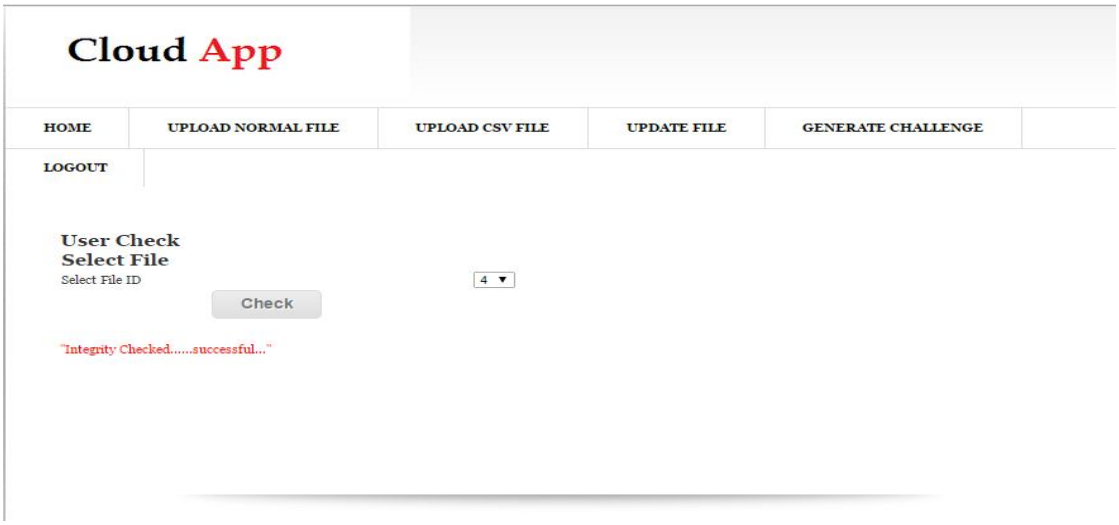
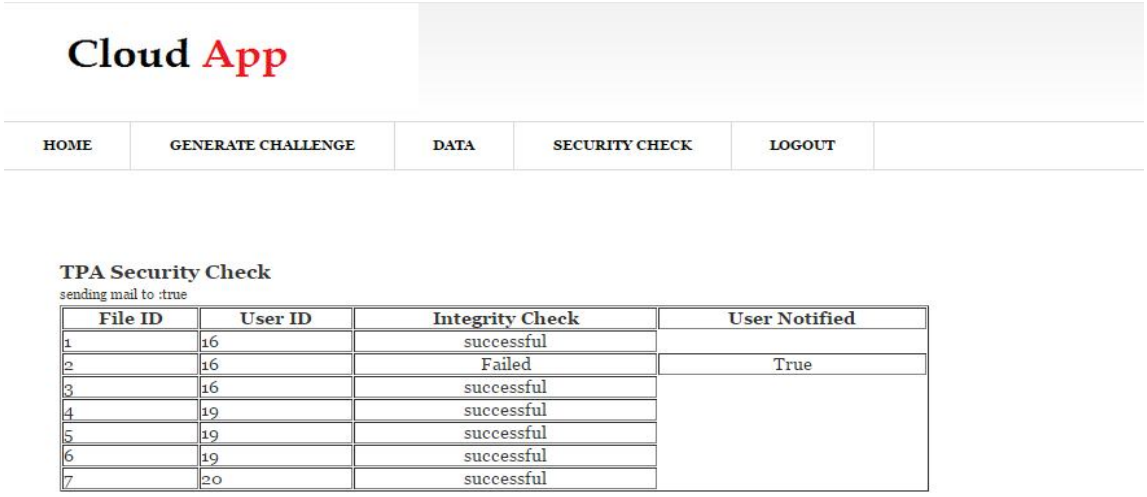


Fig 5.Snapshot of Data integrity check at users side

In the below fig 6, In the admin form admin can generate the challenge and he will check the security related user file. if there is any change in the file admin will notified to user through email message "Data integrity is failed"



File ID	User ID	Integrity Check	User Notified
1	16	successful	
2	16	Failed	True
3	16	successful	
4	19	successful	
5	19	successful	
6	19	successful	
7	20	successful	

Fig 6.Snapshot of TPA security check

IV. CONCLUSION

In this paper, we have noticed a lot of problems related to data integrity in cloud data storage, which is necessarily in a distributed storage systems. For ensuring the exactness data of the cloud user in the cloud , we proposed some cloud services for the data integrity when cloud users data outsource sensitive data for sharing on cloud server. In order to address this is new challenge problem and later achieves a secure and dependable cloud storage services. In this

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

proposed system, our scheme achieved the integration of storage correctness verification and error localization that is the identification of the misdemeaning server .For this extensive data integrity and analysis shows that the proposed scheme is highly efficient and resilient against the internal and external attacks. We have utilized file based and fully automatic and computerized TPA auditing system for efficient and avoiding new vulnerability threats, while eliminate burden of the users from auditing task.In this system, communication secure with help of socket programming. In this system cloud users able to specify specific files or directory and excludes unnecessary files which to check integrity which minimize the computing resources of CSP and TPA system. We leave the some feature extension like optimization and report the exact data made in files to users, so user data will be easily recoverable.

REFERENCES

[1] C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. 17th Int'l Workshop Quality of Service (IWQoS '09), pp. 1-9, July 2009

[2] C. Wang, K. Ren, W. Lou, and J. Li, "Towards Publicly Auditable Secure Cloud Data Storage Services," IEEE Network Magazine, vol. 24, no. 4, pp. 19-24, July/Aug. 2010.

[3] A. Juels and B.S. Kaliski Jr., "PORs: Proofs of Retrievability for Large Files," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 584-597, Oct. 2007.

[4] 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 and Distributed Systems, vol. 22, no. 5, pp. 847-859, 2011

[5] M.A. Shah, R. Swaminathan, and M. Baker, "Privacy-Preserving Audit and Extraction of Digital Contents," Cryptology ePrint Archive, Report 2008/186, http://eprint.iacr.org, 2008.

[6] T. Schwarz and E.L. Miller, "Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage," Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS '06), pp. 12-12, 2006.

[7] 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 and Distributed Systems, vol. 22, no. 5, pp. 847-859, 2011.

[8] K.D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Proc. ACM Conf. Computer and Comm. Security (CCS '09), pp. 187-198, 2009

[9] M. Castro and B. Liskov, "Practical Byzantine Fault Tolerance and Proactive Recovery," ACM Trans. Computer Systems, vol. 20, no. 4, pp. 398-461, 2002.

[10] G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Fourth Int'l Conf. Security and Privacy in Comm. Netowrks (SecureComm '08), pp. 1-10, 2008.

[11] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. 14th European Conf. Research in Computer Security (ESORICS '09), pp. 355-370, 2009.

[12] C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," Proc. 16th ACM Conf. Computer and Comm. Security (CCS '09), pp. 213-222, 2009.

[13] A. Juels and B.S. Kaliski Jr., "PORs: Proofs of Retrievability for Large Files," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 584-597, Oct. 2007.

[14] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 598-609, Oct. 2007.

[15] M.A. Shah, M. Baker, J.C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIX Workshop Hot Topics in Operating Systems (HotOS '07), pp. 1-6, 2007.