Data Integrity Auditing for Securing Cloud Storage

Wale Amol D.
M.Tech (CSE), IET
Alwar, Rajasthan
India
waleamol@gmail.com

Anil Rao
Assistant Professor (CSE),
Alwar, Rajasthan
India
anil.alw@gmail.com

ABSTRACT

Cloud servers is a platform for allowing expedient, on demand network access to a shared pool of configurable server resources (memory, networks, storage, cpu, applications, and services) that can be rapidly provisioned and released with minimal management effort or cloud service provider interactions. Cloud servers are commonly being used; however, data security is one of the major barricades to adoption in cloud storage. Users can store data and used on demand or for the applications without keeping any local copy of the data on their machine. The Cloud server's storage technologies offers the possible of massive cost savings combined with enlarged IT agility due to pay per consume. However, this technology challenges many traditional approaches to hosting provider and enterprise application design and management. Users can able to upload data on cloud storage without disturbing about to check or verify the truth. Hence integrity auditing for cloud data is more essential task to ensure users data integrity. To do this user can resort the TPA (Third Party Auditor) to check the data on the cloud storage is not sacrilegious the integrity. TPA is the expertise and having good information and capabilities which users can not able to check. TPA audit the integrity of all files stored on the cloud storage on behalf of the users and notify the results. Users should consider the auditing process will not cause new susceptibility against the users valuable and confidential data also ensures integrity auditing will not cause any resources problem.

General Terms

Cloud storage, data integrity auditing

Keywords

Auditing, Cloud, Cloud servers, Data integrity, Data privacy, Security, Storage

1. INTRODUCTION

Data integrity auditing is somewhat you need to have on cloud storages. Different threats like a hacker placing a backdoor on storage using applications; change permissions, modify files, or changing your order form to email him a copy of everyone's credit card and other information while leaving it appear to be functionally normally without any problem. By data auditing process and setting up suitable period scan reporting, this notifies user within hours of when any file was changed, altered, added or removed. It also helps establish an audit trail in the event cloud storage is negotiated. Cloud servers has been envisioned as the next-generation information technology architecture for industry, government, and research, due to configurable server multiple resources and long list of advantages: Auto-Scaling technology, on demand self-service, location self-governing, resource elasticity, dynamic resources allocation, fast, secure, ubiquitous network access, pay per consume, higher uptime and transference of risk [14].

Cloud Computing is renovation the very nature of how businesses use information technology. One elementary side of this paradigm shifting is that data is being consolidated or outsourced to the Cloud server storage. From users' perspective, including both user and enterprises, uploading data to the cloud server in a elastic on demand manner brings appealing benefits: free from the burden for storage and the security management, global data access over autonomous geographical locations, and saving of capital expenditure on maintaining security [13], hardware resources as well as maintenance, etc. whereas Cloud storage makes these features more appealing than ever, it also brings new security vulnerability towards users' valuable data. As a result, the integrity of the data in the cloud is being put at risk due to the above different reasons. Although the organizations under the cloud provider are much more commanding and secure than local computing devices, they are still facing the different internal and external threats for data integrity. Secondly, there do exist various enthusiasms for hosting provider to behave unfaithfully towards the cloud users regarding the status of their remotely stored data. In short, although subcontracting data to the cloud servers is parsimoniously attractive for long term huge data storage, cloud service provider does not provide any guarantee on data integrity and security. This drawback, if not properly addressed, could impede the successful arrangement of the cloud server's design. As users' data on remote storage, traditional cryptographic primitives for the purpose of data safety protection cannot be adopted [10] directly specifically, downloading data on native system for its integrity verification is not a practical solution due to the data transmission cost across the network and security reasons. Considering the large size of the outsourced data store and the user's limited resources capability, the work of auditing the data correctness in a cloud server environment can be expensive for the cloud server users [7], [9]. Moreover, the above of using cloud server storage should be minimized as much as possible, such that cloud user does not need to perform huge processes to use the cloud server data. For example, it is desirable that cloud users don't need to worry about the need to verify the integrity of the data before or after the data retrieval. Besides, there are may be multiple user's accesses the same cloud storage for different purpose and applications, say in an enterprise setting.

To make it ensure the data integrity and minimize the cloud server computation resources as well as online burden on cloud users', it is of critical importance that to enable public auditing process for cloud data storage, so that cloud users may resort to an independent third party auditor

(TPA) to audit the data stored over the cloud storage whenever necessary. The TPA, who has the knowledge and capabilities that users don't, can check the data integrity of all the data stored on the cloud occasionally on behalf of the cloud users, which provides a much more easier and affordable way for the users to ensure their cloud data storage integrity. Moreover, in addition to help users to appraise the risk of their subscribed cloud data services, the audit result obtained from TPA would also be beneficial for the CSP or hosting provider to improve their security related to storage platform. In a word, auditing services will play an important role for this cloud economy to become fully recognized; where users will need ways to assess the risk and gain trust in the cloud service providers or cloud storage. Currently, the notion of public auditability has been proposed in the context of ensuring remotely stored data integrity under different system and security models [8], [9], [10], [11], [12].

Auditability process allows a third party, in adding to the user himself, to verify the integrity of remotely stored data of the cloud. However, most of these schemes [8], [9], [11] don't consider the privacy protection of users' data against external auditors. Indeed, TPA may potentially reveal user data information to the auditors. This severe drawback greatly affects the security of these protocols in Cloud storage. From the perception of protecting data privacy and integrity, the users, who own the data on cloud server and rely on TPA auditing process just for the storage security and integrity of their data, do not want TPA auditing process introducing new susceptibilities of unauthorized data leakage towards their data security [12].

Also there are some legal regulations on outsourced data that is, data not to be leaked to external parties. Without properly intended auditing protocols, encryption itself cannot prevent data from "flowing away" towards TPA during the public auditing process. The reason, it does not totally solve the problem of protecting data privacy from external parties but just reduces it to the key managing. Vulnerability of unauthorized data leakage still remains a problem due to the potential exposure of decryption keys. Therefore, how to enable an auditing protocol possession data private, independent to data encryption is the problem which going to tackle in this paper.

2. PROBLEM STATEMENT

The system model careful is having cloud data storage or files storage involving three different entities. As illustrated in figure 1 [1], the cloud users who store the huge amounts of data in the form of fi les on the cloud storage. Files may be in different types such as binary files, data files, logs files, hidden files. The cloud servers, which fully accomplished by the hosting or cloud service provider for the data storage space and diverse resources like network connection, backup facilities and different level security. Third entity is TPA (Third Party Auditor) having expertise and knowledge of integrity auditing process.

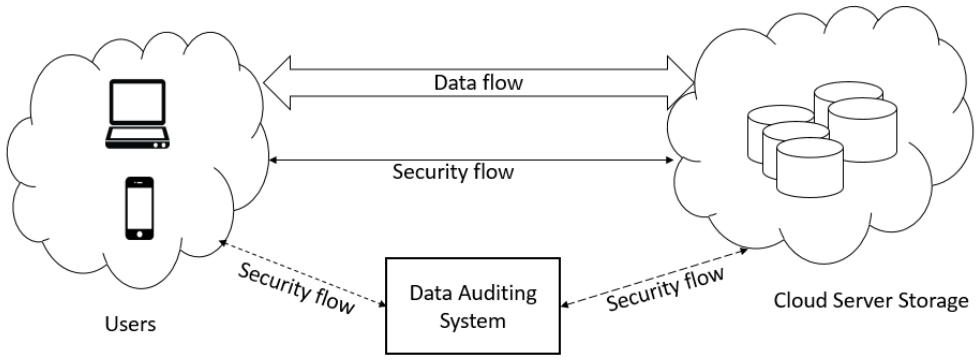

Figure 1. Cloud architecture

Cloud service provider is accountable for storage management, maintenance, ascendable location independent, higher availability, pay per consume and low cost data storage. Users upload and download data dynamically from storage device space on the cloud server for its own application purpose. Users always need to guarantees, data stored over the server is correct and maintained properly. To avoid computational resources and ensure data integrity and security of the data users resort to TPA to audit the data on behalf of user on cloud server.

User's data could be hack, altered or modified by internal or external entities. It may include software bugs, backdoors in different applications, outdated applications versions, plug-in, themes, templates, bugs in system or inexpensively encouraged hackers, malicious code and different upload forms. Cloud servers always provide better security but due to different integrity threats towards data like vulnerable functions used in application, outdated applications versions, plugins, themes, templates, bugs in system exits in application, applications from the un trusted sources which come with preloaded outdoors, hardware failure, network issue there is changes of data loss. Cloud service provider always try to conceal these details from users to their own benefits as well as uphold industry reputation the reason that's why cloud users cannot completely trust on the cloud service provider. With the help of auditing procedure user can gain trust as well as audit this data more efficiently.

3. PROPOSED WORK

This section presents integrity auditing scheme which provides a whole outsourcing solution of data. After introducing notations considered and brief beginnings, started from an overview of proposed data Integrity auditing scheme. Then, giving main scheme and show how to extent the proposed scheme to support integrity auditing for the TPA upon delegations from multiple users. Finally, the proposed how to simplify integrity auditing keeping data privacy scheme and its support of dynamic data. Figure 2 illustrate the overview of integrity auditing structure.

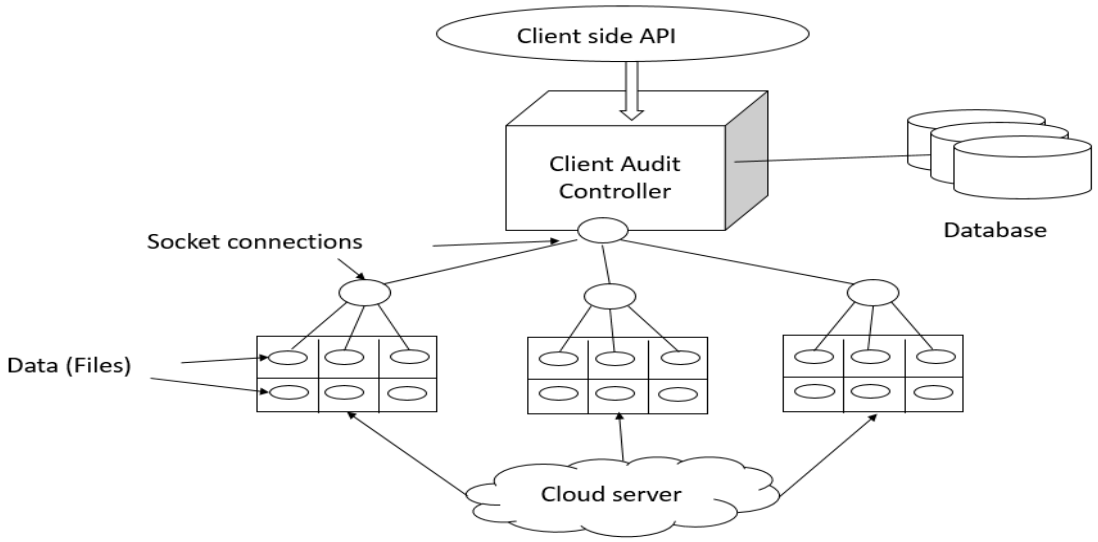

Figure 2. Integrity auditing block diagram

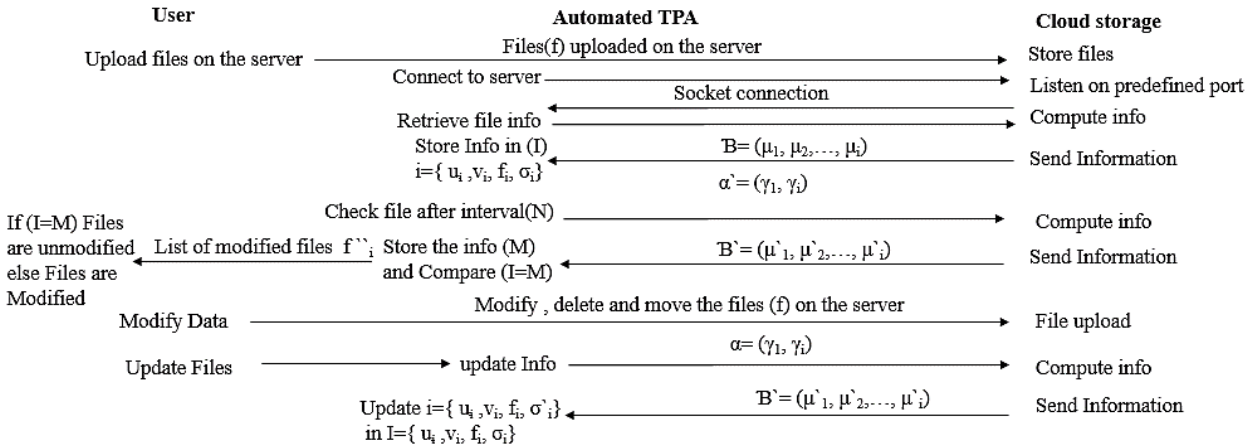

Figure 3. Auditing protocol

4. IMPLEMENTATION DETAILS

4.1 Mathematical Model

S={x, e, i, o, f, DD, NDD, success, failure} Let S be the solution viewpoint of the class x be Initial state of the class Initialize ().

x= {Initialize ()} sets the default values for all variables. Input i =(I1,I2)

I1= {{U}{V}{F}{σ}}

DD=deterministic data which benefits 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 produced or forced exit due to system errors.

Set of 'k' cloud Users U={u1,u2,u3, ……. uk}

Set of 'm' cloud servers V={v1, v2, v3, ……. vm}

Set of files on cloud server storage F={f1, f2, f3, ……. fn}

Set of file tags σi= {f+p+n+u+g+s+acl+b+selinux+md5 +sha256}, i ϵ (1, n)

p= File permissions, t= Type of File, i= File Inode number, u=File User ID,

g= File Group ID, s= File Size, b= total of File Block,

m= File Altered time, a= File Access Time (when the file was last read),

c= is the inode change time, n= Number of links for the file,

S= Check for increasing size, md5: md5 hash, sha1= sha1 hash,

f = File name, I = Initial Values in the Database, N = Interval time of auditing process,

M = New Value of database, LI= List of files, ST= Detail information of modified files,

Set of file tags σ computed based on the file types,

γ= directory path, α = query v=cloud IP address, ß= set of results, μ= consist of file stats.

[Data DATA = f+p+n+u+g+s+acl+b+selinux+md5+sha256]

[Growing files GROW=p+u+g+i+n+S+acl+selinux]

[Password and shadow files IMP =A+sha256]

[Binary and Configuration files. FIXF =A+sha256]

[Hidden file PERM = p+u+g+i+acl+selinux]

[Directories DIR = p+n+ i+u+g+acl+ selinux]

Where A= p + n + i + u + g + b + s +m + c +acl + selinux + md5 H=sha1+sha256+sha512

4.1.1 Initialize ()

TPA send Query initialize()

4.1.2 Update()

A step after the client uploads/modified the files on the cloud server. TPA send Query Update $\alpha = (\gamma , vj)$ where $\gamma \epsilon$ n' and vj is a jth cloud server. n' revised files.

Set of tags $\sigma'i$ = {f+p+n+u+g+s+acl+b+selinux+md5+sha 256},

$i \epsilon (1, n')$ 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'i} Where μi comes from (f1, f2, 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} where ui is user, vi cloud server and σi consist of signature tag of file fi.

4.1.3 Check integrity()

Initial values I = {ui, vi, fi, σi} where, ui is ith user, vi is ith cloud sever IP, μi = (fi, σi) file name with file stats.

Interim to check integrity (N)

Set of tags $\sigma'i$ = {f+p+n+u+g+s+acl+b+selinux+md5+sha 256},

$i \epsilon$ (1, n') where σ' updated files tags.

Number of files F = {f1, f2, f3, f4, ……. fn'}

TPA to cloud server Query Check $\alpha' = (\gamma , vj)$

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 particulars.

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}

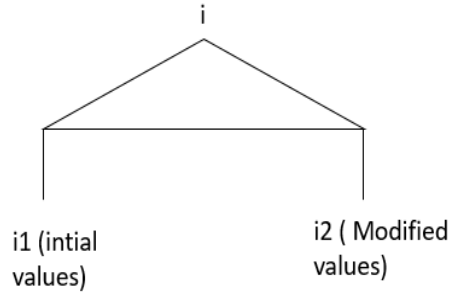

Figure 4. Results comparison

If M {ui, vi, f'i, σ'i } ϵ I {ui, vi, fi, σi}

$\alpha = (\gamma , vj)$ where, $\gamma \epsilon$ 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… μi)

Where, μi comes from (f1, f2, f3…fn) consijsts of pair (fi, σi). TPA store the received values in database (I), the sets of variables and values. I = {ui, vj, fi, σi} Where, ui is user, vj cloud server and σi consist of signature tag of file fi.

As per the Figure 4 TPA system compares the values

Success- If M{ui, vi, f'i, σ'i } ≠ Search result (I){ui, vi, fi, σi}

Results:: Files altered lists (f'i) Else M {ui, vi, f'i, σ'i } = Search result (I){ui, vi, fi, σi}

Results:: Files not altered

Failure-Desired results are not produced.

In this scheme, work based on the six phases includes Install client, connect, upload, initialize, check/compare and update.

5. REFERENCES

[1] Cong Wang, Sherman S.M Chow, Qian Wang, Kui Ren and wening Lou, "Privacy-Preserving Public Auditing for Secure cloud storage" in IEEE transaction on computers vol 62 No 2 February 2013.

[2] Cong Wang, Qian Wang, Kui Ren, Ning Cao, and Wenjing Lou "Toward Secure and Dependable Storage Services in Cloud Computing" IEEE Transaction on Services Computing vol 5 No 2 April-June 2012..

[3] Qian Wang, Cong Wang, Kui Ren , Wenjing Lou And Jin Li " Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing" IEEE transaction Paper on Parallel and Distributed Systems vol 22 No 5, pp. 847-859, May 2011.

[4] Kan Yang and Xiaohua Jia "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing" IEEE transaction on parallel distributed system, Vol 24 No 9 September 2013.

[5] Yan Zhu, Hongxin Hu, Gail-Joon Ahn and mengyang Yu "Cooperative Provable Data possession for Integrity Verification in Multicloud Storage." IEEE Transactions on parallel and distributed system, Vol 23, No. 12, pp. 2231-2244, December 2012.

[6] Shucheng Yu, C. Wang, K. Ren, and Wenjing Lou, "Achieving secure, scalable, and fine-grained access control in cloud computing," in Proc. of IEEE NFOCOM'10, San Diego, CA, USA, March 2010.

[7] Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, http://www.cloudsecurityalliance.org.

[8] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted Stores," in Proc. of CCS'07, Alexandria, VA, October 2007, pp. 598–609.

[9] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS'09, volume 5789 of LNCS. Springer-Verlag, Sep. 2009, pp. 355–370.

[10] A. Juels and J. Burton S. Kaliski, "Pors: Proofs of retrievability or large files," in Proc. of CCS'07, Alexandria, VA, October 2007, pp. 584–597.

[11] H. Shacham and B. Waters, "Compact proofs of retrievability," in Proc. of Asiacrypt 2008, vol. 5350, Dec 2008, pp. 90–107.

[12] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in Proc. Of HotOS'07. Berkeley, CA, USA: USENIX Association, 2007, pp.1–6.

[13] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep.

[14] P. Mell and T. Grance, "Draft NIST working definition of cloud computing," Referenced on June. 3rd, 2009 Online at http://csrc.nist.gov/groups/SNS/cloud-computing/index.html 2009.