Cloud Auditing: An Approach for Betterment of Data Integrity

Sanket Sandesh Shahane, Raj B. Kulkarni

Abstract*—Life has been incredibly busy. Therefore the need for data gradually increases over the time. Users require data to be stored somewhere and retrieved easily, whenever needed. Taking these factors into consideration the concept of cloud is now from the fact. It had been the place where one can store the data of different varieties and could be retrieved at any place, at any time. Also, it provides services to its users which include Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). It substantially becomes effective to a large extent in serving clients, thanks to its properties such as scalability, flexibility, transparency, cost saving ability and eco-friendliness. Each technology had been enforced for making life easier, they'll continuously persist some problems that had to be delineated with. The technology of cloud also had such issues. The field of cloud computing had been facing the issues associated with data integrity such as replaces, replay and forge attack. The continuous need had emerged out for solving such problems, to keep original user data intact. The proposed system presents a Third Party Auditor (TPA) that could solve these issues by shouldering transparency. The paper discusses methods such as Simple Auditing and Dynamic Auditing for solving above mentioned problems. Both the methods use the Tag Generation technique for doing auditing. TPA neither supports Cloud Service Provider nor did Data Owner hence act as monitoring agent. Auditing is carried through logical operations, without a single copy of an original data, to maintain confidentiality. It takes its auditing decisions based on the generated results providing opportunity to Cloud Service Provider, for improving Cloud's services. Auditor warns data owner about prohibited activities of Cloud, boosting the quality of the services. The methods of this paper help in improving auditing services to some extent, along with the above mentioned issues related to data integrity.

***Index Terms*—Cloud Auditing, Cloud Computing, Cloud Services, Cloud Service Provider, Data Owner, Third Party Auditor.**

I. INTRODUCTION

Competence, the fact that helps an organization to grow, and faster moves can be made in the parameter of its quality along with storage utilities. Most of the industries are facing the matter of secure data storage and retrieval. Also, ever increasing number of users and handling of data operations became the hectic task for the organizations. From this viewpoint, a requirement emerges to possess system that provides efficient storage choices before and handles all the possible data operations such as addition, repossession, etc.

But also, from the perspective of the individual users, there is a need of a system that handles data owner's vital documents and can make them available according to the requirement. The field of cloud computing serves the purpose for its users as well as organization. Cloud computing provides brand-new ways that supports the required storage demands. The mobile nature of cloud computing focuses on shared data and computation over a scalable grid of nodes. Various applications and services were provided by the cloud making use of common resources, internet protocols and networking standards. Different components of cloud consist of end user computers, web services and data warehouses. Applications running on the cloud were referred to as Cloud applications [1]. Cloud Applications helps the users to connect with the cloud and enjoy various characteristics of cloud [2].

A. Cloud Characteristics:

1) On-demand self-service:

Cloud has automatic provisioning of computing capabilities which has completely different services provided to users, on-demand. An Example could be taken of storage space, according to the need of the user's cloud allocates the required storage space which can be relocated on the deletion of those files from the cloud. The reallocated memory is then added to the entire storage space of cloud and could be again allocated to any different user according to its need. Thus, avoids the wastage of a storage space and efficiently works with resources at hand.

2) Broad network access:

The Cloud provides services which were available over a wide networked infrastructure. Thanks to broad network access, users from numerous regions could log into the cloud, and could take benefit of the provided services at any given time. Also, a network increases the availability of the resources and hence assures the service guarantee.

3) Resource pooling:

It pulls resources along to serve multiple consumers using a multi-tenant model which supported the resource handling mechanism of cloud to handle the extra load. Suppose a node overhead with the service requests and unable to handle, then a cloud accumulates the processing power of relatively less overhead node to handle the situation by resource pooling. Thus, the cloud provides the pooling of resources according to demand of condition and supports smooth and comfortable functioning.

4) Rapid elasticity:

The cloud provides elasticity for scaling quickly up or down as required [3]. As mentioned before, the resources could be adjusted according to the need, resulting in the extension of Cloud internal infrastructure. Also, the additional resources are often freed once there's no would like so as to save great deal of the maintenance.

Manuscript received January 15, 2014.

Sanket Sandesh Shahane, Master of Engineering, Department of Computer Science and Engineering, Solapur University, Walchand Institute of Technology, Solapur, Maharashtra, India.

Dr. Raj B. Kulkarni, Associate Professor, Department of Computer and Science Engineering, Walchand Institute of Technology, Solapur, Maharshtra, India.

5) Measured service:

The cloud has automatic control and optimization of resources that utilizes a pay-per-use model. In pay-per-use model, the users are charged according to its usage, which will be an economic issue from the user's perspective.

6) User-friendly and low cost:

Cloud substantially supported user-friendly because users requested no special training for its handling. Also, it requires no change within the current infrastructures and hence saving storage and maintenance cost.

B. Services:

The cloud provides the following services as follows [4]:

1) Software as a service:

These services would be available on demand in which service provided by cloud includes provision of different software on demand for cloud users and offered by a third-party provider to the cloud. Examples consist of online data processing and spreadsheet tools, online page delivery services like CRM services, Google Docs, etc.

2) Platform as a service:

Cloud permits customers to develop new applications using APIs and the platforms offered includes deployment platforms, configuration management and deployment tools. Examples are Google App engine, Force, Microsoft Azure, etc.

3) Infrastructure as a service:

Cloud provided virtual machines and other abstracted hardware and operating systems. Examples include Amazon EC2 and S3, Windows Live Skydrive, Enterprise Cloud, Rackspace Cloud Terremark, etc.

II. LITERATURE REVIEW

In conventional technologies, it is unable to figure out work on the outsourced data without a local copy of the entire data in which availability and integrity have been supported on hash functions and signature schemes [5],[6],[7]. Additionally, downloading has not been a sensible resolution in particular for large-sized files, due to the expensive transport communications. The power to audit the correctness of data in a cloud environment could be terrifying and costly for cloud users. Hence, it could be essential to realize public audit ability for Cloud Service Provider (CSP), so that data owners (DOs) might opt for a TPA, who has the capabilities and expertise that a common user did not have. This Audit service has been significantly essential for digital forensics and data guarantee in clouds.

Ways of memory checking called Provable data possession (PDP) and in particular, related to the concept of sub-linear authentication were quite inefficient. In alternative techniques, involving provable data possession scalable PDP (SPDP), dynamic PDP (DPDP) and compact proofs of retrievability (CPOR) files were audited without a local copy, but imposed a serious input, output and process burdens on the server. It showed that these schemes didn't support complete privacy protection and dynamic data operations. Also, it didn't give noteworthy savings in computation and communication costs, and less detection likelihood of corrupted blocks. G. Ateniese, R.C. Burns, et al. [8], bring in for provable data possession (PDP) model, permits a user who has data stored at non-trusted server, so as to confirm the authenticity of data stored at server. From the server, PDP produces possible proofs of possession by sampling random sets of blocks, which significantly scale back Input/output operating cost. G. Ateniese, R.D. Pietro, L.V. Mancini and G. Tsudik [9], made an exceptionally skilled and provably protected PDP technique supported on symmetric key cryptography, whereas not requiring any bulk encryption. Also, in distinction with predecessors, PDP technique permits outsourcing of dynamic data, i.e., competently support's dynamic operations, such as block alteration, deletion and append. It is inappropriate for third-party verification as was based upon symmetric key cryptography. C.C. Erway, A. Kupcu , C. Papamanthou and R. Tamassia [10] presents a definitional framework and economical constructions for dynamic provable data possession (DPDP), that extends the PDP model to support obvious updates to hold on data and uses a completely different version of authenticated dictionaries supported on rank data. H. Shacham and B. Waters [11], presented two approaches, first one was privately verifiable and builds pleasingly on pseudorandom functions (PRFs). The second allows for in publicly verifiable proofs and was constructed from the signature theme of Boneh, Lynn and Shacham in bilinear groups. Each solution seems different on homomorphic properties to mix a proof into one small authenticator value, however, lack dynamic operations. In the Table 1, comparisons of the above schemes with features are provided. M. Xie, H. Wang, J. Yin, and X. Meng [12], planned a probabilistic integrity audit approach that inserts a small number of tuples into the outsourced database, solely increases overhead and may not always easy to keep track of tuples which could be changed by an intruder. But due to increase in cost of computation for additional tuples and block space in the cloud, required fake tuples observance.

Table I

Scheme	Comp-Utation	Privacy	Dynamic Operations	Prob. of Detection
PDP[8]	O(1)	Yes	No	$1-(1-\rho)^{t}$
SPDP[9]	O(t)	Yes	No	$1-(1-\rho)^{t.s}$
DPDP-I[10]	O(tlogn)	Yes	Yes	$1-(1-\rho)^{t}$
DPDP-II[10]	O(tlogn)	Yes	Yes	$1-(1-\rho)^{\Omega(n)}$
CPOR-I[11]	O(t)	No	No	$1-(1-\rho)^{t}$
CPOR-II[11]	O(t+s)	No	No	$1-(1-\rho)^{t.s}$

P. Golle, S. Jarecki and I. Mironov [13] first projected the concept of enforcement of storage complexity and provided economical schemes. Unhappily, the guarantee supplied persist weakness than the one provided by PDP schemes. Syam Kumar P, Subramanian R [14], described the properties of the integrity that seems to the main concern for cloud as follows:

A. Properties of Data Integrity:

1) Completeness: Once receiving a challenge from the verifier, if server honestly computes an accurate integrity proof, in which the verifier always accepts the proof as valid.

2) Soundness: Once receiving a challenge from the verifier, the server deceitfully computes the

integrity proof by missing some data bits, the verifier accepts with negligible probability.

3) Probabilistic Detection: Once receiving a response from the server, the verifier checks response whether, valid or not, if invalid, then the verifier detects the corruptions with high probability.

Qian Wang, Cong Wang, Kui Ran, Wenjing Lou, Jin Li [15], explored the problem of providing coincident public audit ability and data dynamics for distant data integrity check in Cloud Computing. The deliberately designed formation did to fulfil these two important goals whereas potency being unbroken, kept closely in mind. Yan Zhu et al.[16] addressed the construction of the PDP scheme for hybrid clouds which have been based on homomorphic verifiable responses with hash index hierarchy and proposed a co-operative PDP theme to help dynamic scalability which resting on several storage servers. But due to homomorphic tags this scheme requires a small, constant and unavoidable amount of overhead.

III. NECESSITY OF CLOUD

Efficient storage, processing and retrieval of data were most severe issues. Everyone needs their data to be stored with higher security and minimum charges. Storage is always a responsibility of a Cloud Service Provider and ought to be efficient in terms of quality and quantity. Retrieval of data should be as easy as clicking and necessarily secured from crooked ones. The users were never concerned about the working of the system and only desire storage, security and easy access. The Cloud has an ability to furnish those requirements; however, the demands were increasing with time. The data uploaded to the cloud by user and Cloud, got to give the data to the user when needed. Even supposing, the personal system of user goes down; the stored data could be accessed from any other system. In the modern generation of technology, user should be allowed to access data from different devices like personal computer, laptop, palmtop, television, tablet, etc. The cloud would be accessible to user at any cost and solely needs internet connection and web browser. Thus, data should be easily reached even when travelling through train or vehicle and even in any other part of the world. The user should get connectivity to their data under any circumstances.

There are certain questions to be thought such as What about storage capacity of laptop or personal computer? What storage cost for? The important issue emerged that, cloud couldn't have limited capacity like hundreds or a thousand gigabytes; the cloud scales with the usage demands. Don't worry; this should be done automatically by the cloud, however, on user's request. Hence, even acquisition of one Terabyte of external hard-disk space may prove inefficient to store data. Will the question be actually resolved, by purchasing high capacity storing devices? What happens if the device got damaged or corrupted? Who will be accountable for such a fault? The user or the device provider or both are responsible. There exists a high risk of carrying such heavy storing devices that were meant to use at stable position. The question would even be raised that a data at cloud could be corrupted, broken or maybe lost and then no purpose exist to store data in the cloud. Yap! Really an expert question however the solution is lucid. The reason behind the cloud success is that, the Cloud stores and makes numerous copies of data, through replication at various data warehouses. If any replica got corrupted from the cloud, it provides the data from any other replica, hence, cloud provide valuable data even on data corruption or damage at one specific place. It is like money Banks which provide security and make sure the amount. There were numerous services provided to user other than storage and retrieval like pay per use, the most favorite and attracts many people. The feature provides the user to use any software and applications without actually purchasing them. The users only have to pay for their usage, thus the space required for the installation and money to be invested in purchasing are saved. Varieties of software were available on Cloud, round the clock and some provided free on cloud but some were charged. Free softwares are sponsored by commercial companies for selling products which increases the revenue of the Commercial Companies. Software vendors get profit from both Sponsors and Cloud users; hence beneficial to all users, software vendors, commercial firms and Cloud Service Provider.

IV. BENEFITS

The benefits of the cloud include some of the following:

1) Reduced Cost:

For IT company storage has been a critical issue. Many companies were forced to use large amounts of electricity and hardware. Maintenance of stored data demands trained personnel, thanks to Cloud that saves the major part of investment which could be invested in serving the customers. Good service leads to satisfied customers along with increase profits and growth of the company.

2) Flexibility:

It would be helpful to quickly alter with dynamic business needs which add easiness and smartness to company infrastructure. The system of assorted configurations can access cloud and many of devices like Tablet, Laptop would access the Cloud, this makes profitable to posse flexible infrastructure and supports node the addition or removal with comfort. It posse style of 'use what you want, pay only for what you use'.

3) End User Satisfaction:

The Cloud users could download software from the CSP, use tools and applications from the best companies like Microsoft, Google, etc. The users didn't need to purchase software and were relieved from purchasing high cost suits against the traditional customs. For traditional users, software consumes high storage space on hard-disk after installation and user has to use even discontented with software features, as it was purchased. On cloud user can compare numerous software features and opt for software according to need, thanks to pay as you use facility, cloud users save excess investment and enjoy flexibility of access.

4) Universally Uniform:

The facilities provided for the Cloud users are global, even the users from America, Asia, and Africa or from any continents, could download and run any software and application that were provided by cloud. There is no differentiation based on region, language and country, the access to the cloud will be unrestricted and facilities provided by cloud were uniformly global, hence users enjoy the feel of freedom and equality.

5) Scalability:

The Companies were ever increasing and needs further nodes to serve more customers. In usual, for company adding nodes is quite complex task, which requires hardware and software to be compatible and the new node has to be brought in network. In Cloud, nodes could be added without any major changes in the infrastructure by connecting new node to the internet. Cloud provides global access to highly scalable and flexible IT resources that would be turned on or off whenever business desires, thus needs no extra cost for scaling.

6) Lower Environmental Impact:

The energy required for storing data and maintaining which reduces the power usage of the computers wherever storage at large volume is required. According to Accenture Report [17], factors like Dynamic Provisioning of resources, multi-tenancy approach to minimize need of extra infrastructure, proper server utilization and energy efficient technologies for data center have enabled the Cloud computing to lower energy usage and carbon emissions. Thanks to Cloud features, organizations that can reduce carbon emissions by at least 30% per user by moving their applications to the Cloud.

7) Capacity Utilization:

In Cloud, several users were online and they may upload and download data at the diverse time. Thus, the cloud is to be available round the clock. A Cloud could nearly double capacity utilization and decreases the costs of management, generally by half [18]. In general, on-premise infrastructure runs with very low consumption, occasionally goes down up to five to ten percent of average utilization. Many customers use only a small portion of storage allotment, also the slender provisioning technology allows IT to over-allocate capacity and thereby increase utilization of storage resources [19]. Even server running at higher utilization may process added workload with similar power usage [20].

8) Concentration on core:

The companies are facing many issues to be sorted out like recruiting fresh and innovative candidates, training of new and experienced employees for increased productivity as well as efficiency, completing the projects within deadline along with precision and many more. Through a Cloud, companies need not to care about data storage and its maintenance and could directly approach Cloud for fetching it. Also, cloud handles the issues such as data security and integrity causing the company to simply concentrate on core business.

9) Protection:

The Cloud Service Provider (CSP) gives the custody against the corruption of data [21]. The data stored at the cloud in encrypted form and if hackers or intruders get the data it would be useless, as difficult to decrypt. The data owner has no worries even if the server goes down as the replica of the server is always maintained, which provides data to the data owner. The question arises that if any data is lost, corrupted, damaged and leaked then how it could be acknowledged by the data owner? Could the Cloud recognize this? If recognized, would Cloud confess to the owner of the data? Even the data owner desire to understand regarding his data security, however, in reality, the data owner has limited potential and did not have any witness to raise any question to Cloud as incapable and helpless, compared to the cloud and has fewer privileges. Cloud posse immense computation power and storage compared to user who are restricted and unarmed to investigate against Cloud. Auditor would be only promising weapon and empowers the user to verify the Cloud.

V. NECESSITY OF AUDITING

The Auditor is an official, whose job is to carefully check the accuracy of business records and accounts. An auditor could be an autonomous body and self-directed, unaffiliated with the company being audited or captive auditors, as well as some are elected public officials. Whenever cloud is concerned, Auditor checks the accuracy of cloud and keeps his eye on the stored data [22]. The Cloud has numerous users and size of data could be extremely huge. Users don't have such technical knowledge and computation power to audit cloud, users cannot build informed decision if data has been lost in case. Imagine in the year of 2013, images of some occasion have been clicked and have stored it on the cloud. In 2030, if an attempt has been made to retrieve those clicks and surprisingly, only some pictures are available and many of the images are lost. These images could be deleted by service provider or got corrupted. Data owners need to be convinced about the safety of data in the cloud and users satisfy only when the data is correctly stored. This mainly occurs when Cloud believed that the user won't be aiming to open stored images. In such situations, Auditor plays a vital role, it reports to the data owner if cloud behaves illegally by providing witnesses to the user. The user gets the report for each single problem that has occurred and hence provides a warning to user. The Auditor can also be used to improve services of cloud by constantly monitoring it. Making profit through serving customers is the main objective of a Cloud Service Provider (CSP). It could delete data which were not accessed from last the few years to save space and cost of maintenance. It could sell the data to any company or someone like personal information and documents of data owner. The Auditor evaluates the standard of services through externally available interface and predicts cloud behavior. Auditing evaluates the extent to which a Cloud service provider follows best practices, evaluates structure and processes within a service. Auditing need specialized interfaces and have to establish extraordinary standard for comparison which helps understand what to audit. The auditing has simple process which consumes fewer resources and no new vulnerabilities were added. Auditor maintains the data privacy as no local copy of data is disclosed and auditing is carried through logical operations. It provides trustworthy results which were helpful to detect misbehavior of cloud. The Auditor is neither biased toward data owner nor toward the CSP, it checks for the corrupted data blocks and the result produced purely on technical facts. Auditor tells at what extent the corruption occurred, through the report of auditor, users prove that cloud has reliability or not, hence, inclines Cloud to use best practices for storage. Traditionally Audit was done by providing complete data and Auditor examines this stored data which might break privacy. New technologies and demand for audit without data download increased over the time and currently the audit is organized without actually providing data to Auditor and analyzes the data by checking the blocks of data, in that report, if any problems detected were notified to data owner.

These reports prevent Cloud

from doing banned practices like leaking, copying data, etc. and, also prevent the data owner to store data in such defective cloud. Auditor empowered the data owner to keep a watch on Cloud and reports can be used to question Cloud. An agreement is signed between the Cloud Service Provider and the Data Owner; if Cloud violates agreement, then cloud has to bear action according to agreement made [23].

VI. PROPOSED WORK

1) Tag Generation:

The data owner selects a file to check integrity of file. Key Generation generates the Secret key and Public key, in which Secret key pre-process a file that consists of a set of verification parameters and hash table which is stored in TPA.

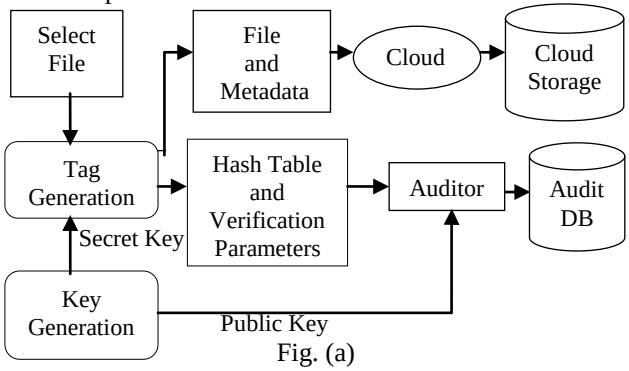

Fig. (a)

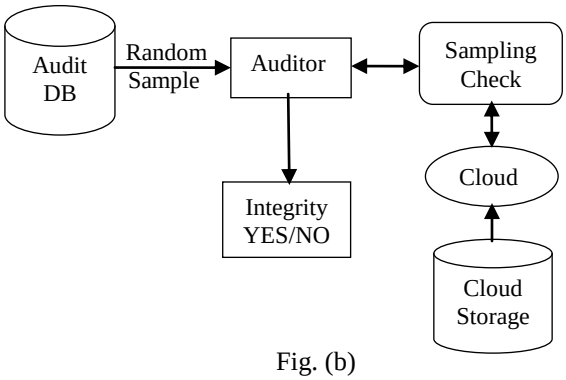

Fig. (b)

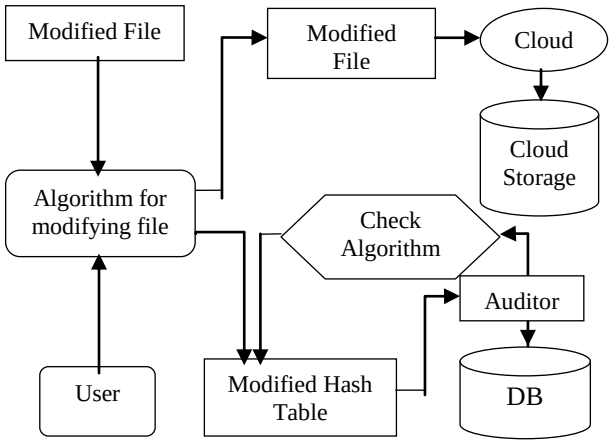

Fig. (c)

The file is transmitted in the form of collection of blocks and verification tags to CSP and data owner may delete its local copy, these verification Tags help carry out Sampling Auditing and Dynamic Auditing, which is in the next part. The Audit Database (DB) stores data which includes the hashes and verification parameters and helps carry out Audit without actual file which supports transparency (see Fig. a).

2) Sampling Auditing:

Auditor collects some random samples of data to perform audit which includes parameters like version, date, etc. Due to periodic sampling a large number of files of different owners can be checked in less time which increases reliability and availability of TPA, which reduces the load of cloud to provide all blocks of the file to TPA. TPA performs integrity check on some blocks of the file and checks for the block's corruptions.

Thus, no need to scan complete file, which reduces processing, saves time of TPA to process files, cost of transporting files and transmit blocks corresponding to a same file to the auditor which breaks privacy of data (see Fig. b).

3) Dynamic Auditing:

Simple auditing could not check the data integrity of data which is modifying simultaneously, due to continuous changes in tags and hashes, user may be attacked when uses an application or edit a file online. Whenever user uses a cloud application, the authorized application uses a user's secret key which may manipulate the outsourced data and update the associated Hash values stored at the TPA. With the assistance of Dynamic Auditing, the privacy of a secret key and the checking algorithm make sure that the storage server cannot cheat the users and forge the valid audit records. Thus, it prevents replay, reply and forge attacks that are widely utilized by professional hackers (see Fig. c).

VII. APPLICATION OF AUDITING:

1) Finding Non-trusted Storage:

During peak period processing at cloud might increase, above the estimated level which requires some time to extend Infrastructure. Cloud shifts some processes to some other servers, which could be trusted or non-trusted. The Cloud has to believe on the outside servers, on faith, but nothing would be understood if outside server follows any banned habit. Illegal activities of outside server were hard to notice by Cloud without the helping hand of the Auditor [24].

2) Accuracy of Cloud:

Many times data may be modified by different data owners and CSP modifies changes into respective data owner's stored data. If the data is modified by data owner, but allocated in an improper way, to the other data owner, may lead ambiguity. This would create confusion for both original as well as mistaken user whose data is accidentally modified; condition even gets worse if there are millions of users. In such a condition, users as well as CSP may get benefited from reports which could boost accuracy of Cloud.

3) Avoid Replace attack:

The server may choose another valid and uncorrupted pair of data block and data tag to replace the challenged pair of data block and data tag when already discarded. The Replace attack would be prevented by TPA because maintained a record of data blocks and tags in Audit Database (DB).

4) Avoid Forge attack:

This attack generally occurs during the dynamic operations are taking place in which, the server may forge the data tag of the data block and deceive the TPA, the owner's secret tag keys were reused for the different versions of data. The TPA keeps track of the tag which makes attack futile.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

5) Avoid Replay attack:

A replay attack, a form of network attack in which a valid data transmission would be maliciously or fraudulently repeated or delayed which would carry out either by the maker or by an adversary who intercepts the data and retransmit it like a masquerade attack by IP packet substitution like stream cipher attack. The server might produce the proof from the previous proof or replaced data, while not downloading the actual owner's data. The TPA always logically checks the generated proof and do not rely blindly on previous decisions taken.

6) Leaked content:

The Auditor report consists of the complete record of affected blocks, they facilitate to find the affected files. The corrupted blocks correspond to files which are modified, damaged or leaked by any unauthenticated one. The Auditor comes to know, when, block version is changed, which could be possibly done by the data owner. If the block version modified, without owner's consent, then data leakage may be possible in such scenario. Once the data leakage recognized, also, it can be possible to detect reasons and find the intruder.

6) Government Documentation:

Population explosion occurred already in countries like India and China, government were under huge pressure to maintain records. Some criminal minded people are making benefit out of this by creating duplicate certificates. They were using such fake certificates for getting various benefits and facilities from government, also used for accessing restricted areas like defense, industry, ministry, etc., which were hazardous for peace and integrity of the nation. The Government should have an Auditor on stored data that may stop such culprits who are against the humanity.

7) Immigration and Emigration:

Number of people move outside countries and everyone has unique passport. Sometimes a fake passport is created by using identity of such person who died or not existed. In such condition, identifying a fake one may be impossible. The systems of all nations if integrated with each other, such fake person could be tracked. The person carrying a fake passport could be caught when personnel at the counter refers to the global database, in such condition the system may provide actual status of passport and such culprits with the help of Auditor could control such illegal immigration as well as emigrations.

8) Hospitals:

In the medical field, patient's history plays a vital role throughout emergency situation, which consist of data like blood group, chronic diseases, allergies, blood pressure, etc., as they were critical while operating patient. In large hospitals, data stored on systems and there were numerous hospitals that have internet connectivity. Critical data should be shielded from hackers or native employees and also from accidental edit; which might not come to notice of database administrator easily unless some accident happens. Such technical defect may take someone's life just due to carelessness toward data security and integrity. In such condition, the auditor may warn the administrator and doctors to reexamine such patients, whose data is manipulated.

9) Corruption Calculation:

The Auditor checks for malicious activities of cloud and detects corrupted blocks, which were used to find affected files. The reports generated by TPA reflect all misbehavior of cloud and if, all reports were integrated with each other, percentage of corruption could be calculated from which calculation of corrupted files can be prepared.

10) Rating the Cloud Service Provider:

Reports are generated by Auditor, they were considered as mirrors of cloud, with the help of reports, rating system can be created. If the corruption occurred to the smaller extent, then higher rating is awarded to such Cloud Service Provider and if higher corruption is detected, then lower rating is delivered. These audit reports were purely based on logical operations and didn't have an external interference from Cloud service provider as well as data owner. These reports have technical proofs and logic which could be provided to opposing Cloud Service Provider, also the data owners would refer only to high rated cloud.

VIII. CONCLUSION

Storage and retrieval of documents on demand is becoming the real need of the life. Hence, the concept of cloud originates. This paper discusses first the need of cloud in the introduction along with its background. Then, discusses, it's characteristics followed by the previous work done by the different people in this field. Then, discusses different services provided by Cloud followed by the necessity of cloud in brief, and benefits. Then, the proposed approach of auditing is presented. Cloud has been enormous and due to huge data, one cannot guarantee data integrity. There need, some system regulating this fact of data integrity. Here, the system mentioned in the paper does the same work of regulation and is called an auditor. The Auditor does the dedicated work to perform cloud auditing. As discussed, Auditor uses two major approaches for cloud auditing. First, is simple auditing, in which the user has the privilege to select the files for auditing. The second one is the Dynamic Auditing, where the auditor chooses the files for auditing and performs the auditing, while users were modifying data. Both the approaches use Tag Generation technique for auditing. The Auditor has to do nothing with size of the Cloud. Due to Automation of Auditor; the efficiency is absolute. Whenever, the report generation involves no human intervention it will be quite accurate, transparent and confidential as well. Auditing system makes cloud storage more safe and reliable.

REFERENCES

1. Mirzaei, Nariman. "Cloud Computing." Pervasive Technology Institute Report, Community Grids Lab, Indiana University (2008): 1-12. Available:http://grids.ucs.indiana.edu/ptliupages/publications/ReportNarimanMirzaeiJan09.pdf.
2. Huth, Alexa, and James Cebula. "The Basics of Cloud Computing." United States Computer, Carnegie Mellon University, 2011. Available:https://www.us-cert.gov/sites/default/files/publications/CloudComputingHuthCebula.pdf.
3. Cloud Computing in education, UNESCO (United Nations Educational Scientific and Cultural Organization), UNESCO Institute for Information Technologies in Education, Policy brief, September 2010.Available:http://unesdoc.unesco.org/images/0019/001904/190432e.pdf
4. Cloud Computing: Benefits, risks and recommendations for information security, *ENISA, European Network and Information Security Agency*, November 2009.Available:http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at_download/fullReport

5. Hsiao, Hsu-Chun, Yue-Hsun Lin, Ahren Studer, Cassandra Studer, King-Hang Wang, Hiroaki Kikuchi, Adrian Perrig, Hung-Min Sun, and Bo-Yin Yang. "A study of user-friendly hash comparison schemes."In *Computer Security Applications Conference*, ACSAC'09. Annual, pp. 105-114. IEEE, 2009.
6. A.R. Yumerefendi and J.S. Chase, "Strong Accountability for Network Storage," *Proc. Sixth USENIX Conf. File and Storage Technologies (FAST)* , pp. 77-92, 2007.
7. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S. Yau, "Efficient
8. Provable Data Possession for Hybrid Clouds," *Proc. 17th ACM Conf. Computer and Comm. Security*, pp. 756-758, 2010.
9. G. Ateniese, R.C. Burns, R. Curtmola, J. Herring, L. Kissner, Z.N.J. Peterson, and D.X. Song, "Provable Data Possession at Untrusted Stores," *Proc. 14th ACM Conf. Computer and Comm. Security*, pp. 598-609, 2007.
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)*, pp. 1-10, 2008.
11. C.C. Erway, A. Kupcu , C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," *Proc. 16th ACM Conference, Computer and Communication Security*, pp. 213-222, 2009.
12. Hovav Shacham, and Brent Waters. "Compact proofs of retrievability." In *Advances in Cryptology-ASIACRYPT 2008*, pp. 90-107. Springer Berlin Heidelberg, 2008.
13. Min Xie, Haixun Wang, Jian Yin and Xiaofeng Meng. "Integrity auditing of outsourced data." In *Proceedings of the 33rd international conference on Very large data bases*, pp. 782-793. VLDB Endowment, 2007.
14. Philippe Golle, Stanislaw Jarecki, and Ilya Mironov. "Cryptographic primitives enforcing communication and storage complexity." In *Financial Cryptography*, pp. 120-135. Springer Berlin Heidelberg, 2003.
15. P. Syam Kumar and R. Subramanian. "An Efficient and Secure Protocol for Ensuring Data Storage Security in Cloud Computing." *IJCSI International Journal of Computer Science Issues 8, no. 6*, 2011.
16. Wang Qian, Cong Wang, Kui Ren, Wenjing Lou and Jin Li. "Enabling public auditability and data dynamics for storage security in cloud computing." *Parallel and Distributed Systems, IEEE Transactions on 22, no. 5*: 847-859. No 5, 2011.
17. Yan Zhu, Huaixi Wang, Zexing Hu, Gail-Joon Ahn, Hongxin Hu, and Stephen S. Yau. "Efficient provable data possession for hybrid clouds." In *Proceedings of the 17th ACM conference on Computer and communications security*, pp. 756-758. ACM, 2010.
18. "The Carbon emissions of server computing for small-to medium sized organizations", a performance studies of On-premise verses the Cloud, October 2012. Available:https://www.wspgroup.com/Globaln/USA/Environmental/Sustainability/Documents/NRDC-WSP_Cloud_Computing.pdf.
19. Available: http://www.storagemagazine.co.uk
20. Sudip Chahal, Krishnamurthy Anandarao, Chris Peters, Shane Healy, Nigel Wayman, Steve Owen, " Implementing Cloud Storage Metrics to Improve IT Efficiency and Capacity Management", *Intel IT, IT best practices, Cloud Computing and IT Efficiency*, July 2011. Available: http://www.intel.com/content/dam/doc/white-paper/intel-it-implementing-cloud-storage-metrics-paper.pdf.
21. Saurabh Kumar Garg and Rajkumar Buyya, "Green Cloud computing and Environmental Sustainability", *Cloud computing and Distributed Systems(CLOUDS)Laboratory.*Available:http://www.cloudbus.org/papers/Cloud-EnvSustainability2011.pdf
22. Available: http://www. Investopedia.com
23. Seny Kamara, Kristin Lauter, "Cryptographic Cloud Storage", *In Financial Cryptography and Data Security*, pp. 136-149, Springer Berlin Heidelberg, 2010, Available:http://www.informatik.uni-trier.de/~ley/pers/hd/k/Kamara:Seny
24. Mehul A. Shah, Mary Baker, Jeffrey C. Mogul, Ram Swaminathan ,"Auditing to Keep Online Storage Services Honest.", In *HotOS*, 2007. Available:http://www.hpl.hp.com/personal/Mary_Baker/publications/hotos2007.pdfpt
25. "Negotiating the cloud – legal issues in cloud computing agreements",
26. Better Practice Guide, Department of Finance and Deregulation, Australian Government, November2011. Available:http://agict.gov.au/files/2011/11/Cloud-Legal-Draft-Better-Practice-Guide-November-2011.pdf.

AUTHORS PROFILE

Sanket Sandesh Shahane is pursuing his M.E. in Computer Science and Engineering from Walchand Institute of Engineering, Solapur University, Maharashtra, India. He received his B.E. in Computer Science and Engineering from Rajaram Shinde College of Engineering, Pedhambe, Chiplun, Ratnagiri, Maharashtra, India from Mumbai University. His areas of interest include Cloud Computing and work in Cloud Auditing and maintaining Cloud's Data Integrity.

Dr. Raj B.Kulkarni, received Degree of Ph.d from Solapur University. He is a Associate Professor in Walchand Institute of Technology. His area of interest includes Restructuring, Data Mining and Web mining. He has published many papers in International journals, National journals and also in International Conference Proceedings. He is a CSI member and Akash Workshop coordinator held by IIT, Bombay. He attended many workshops of National and International level.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication