Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 2): pp. 837-843
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.02.94

PRIVACY PRESERVING AND PUBLIC AUDITING MECHANISM FOR SHARED DATA IN THE CLOUD INFRASTRUCTURE

Pavneet Kaur, M. Tech Scholar, SRM Institute of Science and Technology, SRM University, Kattankulathur, Chennai, India, Email: kaurpavneet1810@gmail.com
Dr. M. Pushpalatha, Professor, SRM Institute of Science and Technology, SRM University,Kattankulathur, Chennai, India, Email: pushpalm@srmist.edu.in

Abstract — More recently, the capability of computational power being dispersed in the cloud is seen as a massive leap forward that allows for improvements in network bandwidth, computational power capacity expansion, and the ability to store data on the cloud. The knowledge is updated and shared more readily on the cloud, which is beneficial for clients. To verify the consistency of shared data blocks, you must compute the authentication signatures on all of each one (client-side)node and each peer- (or all)nodes in the cluster. A distinct block of shared data was registered with a distinct user. The block must be re-signed by the user who still has access to it if it has been withdrawn from the previous signatory (a user previously permitted to sign blocks must still be present if blocks are being reissued to him). However, since files of this size affect cloud performance, there is the risk of bogging down the server while sending and accessing them. If the shared data that the user previously had downloaded has expired, the former client will automatically re-sign at the time of user revocation. Since the current system is less focused on protection and the data backup process, data is vulnerable to lose in a system failure event. It was the first paper to introduce the truth about shared data and the use of AES encryption to secure data against cloud server tampering and suggested the use of a new mechanism to locate it, such as the recovery procedure. Comparison of the two systems revealed far more flaws in the current system, but several additional problems with the new one and another critical aspect of the Group Policy Management feature are key revocation. When a user is kicked out of a group, their personal security credentials (i.e. his or her keystrokes) are revoked, new personal keys are distributed to current users in the group.

Keywords—Public Auditing, Cloud computing, encryption,surrogate server, decryption, signature,AES algorithm

I. INTRODUCTION

Like knowing readily available to the Internet is, users often gain the ability to work together on documents and other digital media projects like Dropbox, Google Drive, etc., by expanding cloud capacities like collaboration and storage. Shared data is made available to all those who are a part of the group and everyone else, and the group has access to all the data. The individuals can only use the data, but they may also share it and change it if given their explicit permission. Because of the extra data protection in the cloud, customers would not be concerned whether a breakdown in the environment causes the loss of information or human error and the fact that the environment itself is more secure, the information will not be altered.

Several different methods are provided to protect the integrity of the data inside the cloud. When we allocate a signature to all data, integrity depends on the number of signatures, allowing for the quick determination of which blocks of data are accurate. Generally, a public auditing approach is useful because it does not require entire data to be downloaded; it is useful and ordinary in this case because it allows an auditor to inspect the data integrity inside the cloud without having to receive the data. If a specific knowledge expander needs to use the cloud for searching, computing, analysis, or data mining, there is a public search paraphier that he or she can use. The third-party auditor will provide information about customers as well as act as support for providing authentication.

With universal, always-on, on-demand network and servers, plus networks, servers, storage, and utilities, provide networks, apps, and services to the client, which are always available and require no waiting on the part of the time. The fact that cloud computing offers storage options to clients and organizations places to store and process information in third-party data centres gives it value to all customers. In

addition to the idea of sharing resources to get an equitable inheritance, such as an electric power grid, the networks are shared to promote in market-oriented energy companies in networks, meaning that all users benefit equally from the use of resources in market economies, e.g. The architecture as a cloud computing model is predicated on widely available resources as well as many available services.

We delve into related work in Section II, where we see a model of the device, algorithms, algorithm instructions, and equipment for how the hardware should be designed, and a mathematical calculation, in Section III. In the final section (IV), we anticipate the possible outcomes and finally give a conclusion (summation) in the final section (V).

II. RELATED WORK

Bilin Shao et al. [1] discusses that since vehicles don't have much room for extra information, only the required information is stored on the cloud, which allows for an expansion of usable storage space. When data is stored locally, the vehicles are in charge of its availability; however, when data is held in the cloud, it's not in their possession and therefore vehicles have lost their ability to restrict the amount of data available. Storage owners are most concerned with keeping their customer's data secure; hence, cloud-based storage vehicles have privacy baked into their designs from the outset. Unfortunately, the majority of integrity auditing systems are ineffective because they do not keep track of their data, lack of data dynamics, and privacy protection, and are very expensive. Concerns regarding data integrity and data protection are allayed with the audit plan outlined in this document, which calls for both data expansion and data verification. When building the multiple regression parts, first create a multiple data structure to represent the multidimensional data An extra feature is to allow use of piecewise linearization (better handling of data dynamic changes) and Boneh-Lyn-Shach signature technology and is the development of a data auditing scheme that's independent of a log audit that's designed from several sources to determine changes across the log. In the evaluation, protection and performance evaluations, last but not least, risk analysis is conducted. This security audit proves that the system can protect the data's privacy, guard against improper use and tampering, and system changes, and deal with replay attacks.

According to Wang et al. [2] users can easily update and share data as a community using cloud data storage and sharing services. To ensure that the credibility of shared data can be publicly checked, users in the community must decide signatures on all blocks of shared data. Because of data changes performed by different users, different blocks in shared data are usually signed by different users. For security purposes, when a user is called from the party, any blocks previously signed by this called user must be re-signed by a current user. Because of the large size of shared data in the cloud, the direct process, which enables a current user to download the corresponding part of shared data and re-sign it during user revocation, is inefficient.

Armbrust et al. [3] states that in comparison to earlier models of computing, in which computing systems where computing resources were static, in a cloud computing model the tools are used as service components over the internet. The practice of keeping memberships in a multi-related cloud open and being open while still maintaining privacy and confidentiality are still still fascinating problems to deal with.

Ateniese et al. [4] presents an evidence to support provable data possession (PDP) that goes one step further and proposes that clients should present server that has possession of their data on untrusted computers with more convincing models to prove that the fact that it originated on the computer rather than proving the integrity of the data itself, because clients usually do not need the original data. Instead of iterating over all the elements of the server array to collect unique ownership proof one at a time, the model generates statistically-confirmed blocks from the server. The client must ensure that the data can be reviewed against a reasonable amount of metadata, ensuring that

H. Shacham and B. Waters, among others [5] Formal paraphrase Cloud computing refers to the different types of services and applications that are distributed via the internet cloud, and the devices used to access these services and applications do not need any special applications. Cloud computing has moved application applications and databases to centralized massive data centers, where data and service management might not be entirely trustworthy. This raises the issue of maintaining the confidentiality of data stored in the cloud.

Among those who have contributed to this work are Wang et al. [6]. Cloud computing is a type of internet-based computing that allows for the scalable acquisition of resources (hardware, software, platforms, and services) from the internet. Since a large number of users store their data in the cloud, data consistency, integrity, and protection are top priorities. This paper investigates the issue of maintaining the privacy and protection of data storage in cloud computing. To ensure the consistency of data, we consider the role

of having a Third Party Auditor (TPA) to validate the quality of data stored in the cloud on behalf of the cloud client. Batch auditing is performed using the bilinear aggregate signature methodology.
According to Wang et al. [7], the amount of sensitive data generated by various organizations has increased, and is beyond their storage capacity. More complex activities, including processing and analysis of vast quantities of data, such as this, are prohibitively costly because of the need for high-capacity and highly qualified personnel, not a lack of storage space. Data storage and/expansion through cloud service providers (DaaS) includes the transferable subscription models which permit businesses to share data storage across multiple remote servers. Because of that, SaaS saves the company from the cost and hassle of huge local storage, the workload that comes with it, it's free. Security should be paid for by someone who provides it, but the person providing it must be compensated if they are providing too much or if their level of protection is unacceptable.
Zhu et al. [8] provide an audit service to verify that source and substrate data in unregulated and outsourced storage are free of errors Outsourced data can be trusted in one way or the other: it is trustworthy in any case, as long as our audit service utilizes techniques such as fragmentary analysis, random sampling, and random hash functions. Additionally, we propose a probabilistic query enhancement using an audit and verification check with regularly scheduled database sample analysis, also known as periodic analysis auditing.
Cao and colleagues [9] the principles of cloud computing, the dynamic data management systems that are inspired to cloud storage from local systems have great versatility and economic savings on public cloud. However, the privacy of public cloud is very poor because data can be compromised. Taking these factors into account, we need a new approach MERS (Multi-keyword Encrypted Ranked Search). This technique helps us to handle privacy data as well as a large number of keyword searches that are ranked. Given the large number of data users and documents in the cloud, it is important to support several keywords in the search request and return documents in the order of their importance to these keywords.
Among those who have contributed to this work are Wang et al. [10]. Extending the cloud computing to apply the ideas of interactivity and interlacement to data storage, a cloud storage model gives users control of their data without it being limited to one location. on-demand cloud applications provide greater benefits because they enable users to be more readily able to enjoy cloud-based services regardless of whether they are in a mobile, branch, in the office, or in other locations that have restrictions Many people may want to use the data at the same time, but not everyone can have access, which means that access has to data must be handled so that everyone gets fair benefits. There are existing protection solutions in place to prevent unauthorized users from getting private data, but also avoid the issue that presents itself when someone pressures the cloud to provide the data that's access, which is that some security vulnerability is required when the next in line users are asked to verify.
According to Tate et al. [11], in clients may be putting their data on a (untrusted)server that has greater capacity and retention power than they know in outsourcing, allowing it to be available to everyone in the internet or to become accessible on a server that isn't well-secured. By incorporating security data accuracy, authenticity, and freshness into the model, this one at the client, while avoiding the need for computing resources, this model will result in considerable reductions in both client and server costs.
R. Patil et al. [12] discussed that it is possible to use cloud-based data clients to migrate data from their systems to cloud-based servers. As a result, the customer is relieved of the cost of maintenance while still receiving high-quality data storage facilities. Cloud storage raises many security concerns. Cloud service providers and data servers are not without flaws. The consumer is worried with whether or not the information stored on the cloud is in order. The public key hash algorithm is used in this article. Furthermore, for data dynamics, this facilitates dynamic operations such as insert, update, remove, and alter at the block stage. The Merkle Hash Tree is used to assist in determining the position of each complex operation. A third-party inspector verifies the correctness of the user's data and certifies the consistency of the data stored on the cloud server. The overhead of computational and communication is minimized. The defragmentation technique is used to determine whether or not the file that the user wishes to store in cloud storage already exists on the cloud server. This system is powerful and safe against malicious server-launched replace attacks.

III. EXISTING SYSTEM

The veracity of information is determined by any piece of data stored in this system, each of which is signed with a unique identification code. This is a common feature of these systems because it allows public verifiers to quickly audit data without uploading it: this involves ensuring that the public auditing functionality is successful while reducing storage costs. This public verifier (TP) may choose to use cloud data for use in the public domain, or a third-party auditor may be able to provide data integrity assurance for use in the private realm. Since users must contribute their personal data when a block is extended, a change necessitates yet another signature, this time from you. Everyone is distinct because no two blocks

are signed by the same user, they can each contain varying amounts of personal information. This user's membership in the community should be refused. To retain security, this user must be kicked out of the community or have their privileges revoked.
Because of the revocation of this revoked user, this user no longer has access to and cannot alter shared data. As a consequence, even after revocation, the contents of shared data will remain null. To retain data integrity, blocks signed by the previously revoked user must be re-signed by a member of the new user's team. The integrity of all data is also ensured by the fact that the public keys of existing users are notaries; as a result, new users cannot be introduced to the cloud.

Disadvantages of Existing System
1. For security reasons, when a user leaves the group or misbehaves, this user must be revoked from the group. As a result, this revoked user should no longer be able to access and modify shared data, and the signatures generated by this revoked user are no longer valid to the group.
2. Cost of communication & computation resources for downloading, verifying blocks, re-computing and uploading signature increase for the existing user.
3. It is difficult for existing user to maintain the correctness of shared data during efficient user revocation

IV. IMPLEMENTATIONMETHEDOLOGY

A. System Overview

The following architecturerepresents the proposed system:

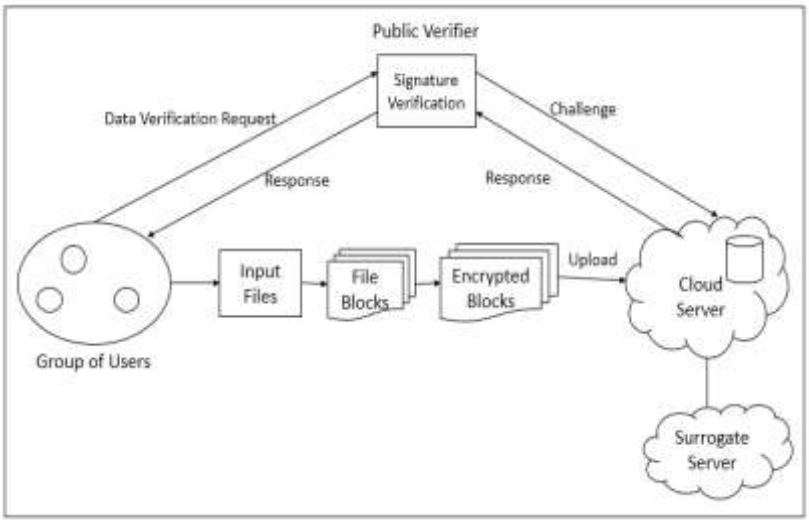

Figure 1. System Architecture

.

Initially user module can perform following operation"

- Registration: All clients have to register to view the files at the beginning. Clients that meet the minimum criteria have the option to log in in to the cloud. The registration module includes the name, email address, and password, plus your cell phone number, amongst other things.
- Login: To be authenticated, users can only use their registered usernames are allowed to log in to the system. There are username and password fields in it.
- Uploading file: With this tool user-friendly process, files can be uploaded to the cloud that are encrypted by using the user's personal key. This is done to ensure that files cannot be accessed by unauthorized users.
- File Encryption: After uploading the file, the file is encrypted by applying AES algorithm.
- Signature Generation:Signature is generated for the file.
- File Block Generation: Number of blocks is generated by providing the file
- User Revocation:This work shows the potential for data owners to revoke their data's permissions. As soon as any of the user leaves the organization, the data is erased and deleted. if users request an expansion of the organization's authority, which holds their private keys, the system must be reconfigured to make it more difficult for such users to withdraw information they will no longer be an

authority over the contents of that information This necessitates that the data owner change all the keys to maintain the safety of their users.

- Proxy Signature: Proxy signature is generated.

2. Auditor Module

- File Verification: the group member can use the public verifier to ensure that data hasn't been tampered with. Even if certain blocks of shared data have been re-signed, the verifier will audit the cloud's integrity.
- Request and Response from user and server:If an auditor is linked to the server, they can communicate with it and also communicate with the user and respond to a request from it.

3. Server Module

- Data Storage: In this module public auditor view the all details of upload, download, blocked user, re-upload.
- Store duplicate data at surrogate server:For the backup the data store in the surrogate server is done by the server.

B. Alogorithm

Algorithm 1: Proposed System Algorithm:

1. Initialization:

Let G_1 and G_2 be the two groups of order p,
g be the generator of G_1, $e: G_1 * G_1 \rightarrow G_2$ be a bilinear map,
w be a generator of G_1
$H: \{0,1\}^* \rightarrow G_1$ where H is a hash Function

2. KeyGen:

A user μA selects a random $\alpha\ \varepsilon\ Z_p^*$ and outputs public key $PK_A = g^a$ and private key $SK_A = a$

3. ReKey

i. A proxy generate a random $r \epsilon Z_p^*$ and sends it to userμA
ii. The μAcomputes and sends r/a to user μB, where $SK_A = a$
iii. User μB calculates and sends rb/a to proxy where $SK_B = b$
iv. The proxy recovers $rK_{A\rightarrow B} = \frac{b}{a} \epsilon Z_p^*$

4. Sign:

Given $SK_A = a$, block $m \epsilon Z_p^*$ and block identifier(id)
User μA output the signature on block m as:

$$\sigma = (H(id)w^m)^a \epsilon G_1$$

5. Resign

Given re-signing key $rK_{A\rightarrow B}$, public key PK_A, signature σ, block $m \epsilon Z_p^*$ and block identifier(id),
The proxy checks that $verify(pk_a, m, id, \sigma) = 1$, if the verification result is 0 the output $\perp$, otherwise its output,

$$\sigma' = \sigma^{rk} A \rightarrow B = (H(id)w^m)^{a.b/a} = (H(id)w^m)^b$$

Algorithm 2: AES Algorithm:

Step in AES Algorithm:

1. Key_expansion: - round keys are derived from cipher key inRijndael's key schedule Round keys which are derived from the cipher key.
2. If Dist_to_tree(u) >Dist_to_tree and First-Sending(u) then
3. Initial_round: - Add_Round_Keyin which every byte is shared with the round key utilizing bitwise XOR.
4. Rounds

Sub_Bytes :nonlinear substitution step.
Shift_Rows : transposition step.

Mix_Columns: mixing operation of each column. Add_Round_Key
5. Final_round: It contain SubBytes, ShiftRows and Ad-d_Round_Key

V. RESULTS ANALYSIS

C. Dataset

Files are taken as an input for our work, varying in sizes ranging from 1 KB to 100MB.

D. Result

Table I. Accuracy Comparison between Existing and Propose System

Fig 2 .Storage Availability Graph

Fig. 2 depicts the overall storage capacity differences between the current system and the system envisioned. Due to the existing system's lack of a provision of backups, which system expands the current server capacity, if the data in the cloud is lost or corrupted, no new or additional servers can be added to it will have to be installed on the existing system. When it comes to storage or file management, structures, the X-axis represents how much storage or file space a system has and the Y-axis reflects how much of that it is using.

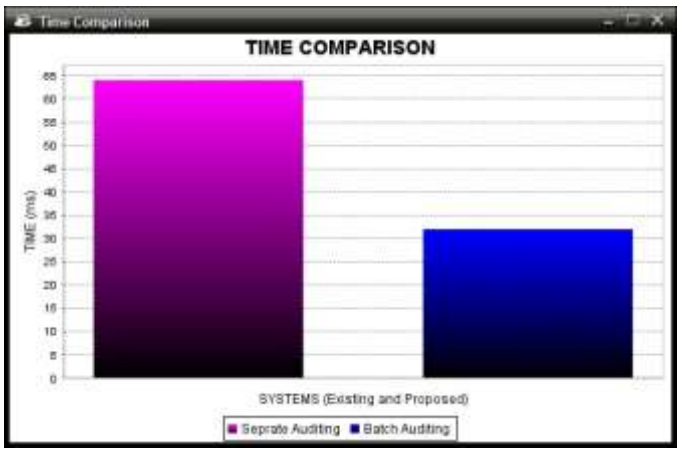

Fig 3. Time Comparison

Fig 3. Shows that the system time comparison shows how the audit procedures and auditing processes vary from each other, both a single-case system and a batch system. In order to accommodate the current scheme, the separate auditing period will be longer than the time used in the batch auditing

We've concluded with the table for the differences between the various device options. We demonstrates how the systems will compare to each other in our proposed solution, and other systems with different parameters.

sTable 1: Comparison Table

	Paper - Proposed System	Paper[1] – Base Paper (PANDA)	Paper[2] - ORUTA
Public Auditing	Yes	Yes	Yes
Proxy Re-signature	Yes	Yes	No
User Revocation	Yes	Yes	No
Encryption	Yes	No	Yes
Data Backup at Server	Yes	No	No

VI. CONCLUSION AND FUTURE SCOPE

The method for disseminating information in the cloud implemented to protect sensitive data is "resourceful dissemination," a practice that cloaks it by cloaking it. The confidence in the client's role is revoked, after which, framework re-expands the semi-trusted cloud signed blocks that were trusted on that user's behalf is invalidated. Has also implemented the AES (Advanced Encryption Standard) algorithm for the cloud for data protection. For backup purposes, a surrogate server is used on the cloud. Experimental studies demonstrate that the proposed system is more stable than the current system. Also, the overall system storage availability is greater than that of the current system.

REFERENCES

[1] B. Shao, G. Bian, Y. Wang, S. Su and C. Guo, "Dynamic Data Integrity Auditing Method Supporting Privacy Protection in Vehicular Cloud Environment," in IEEE Access, vol. 6, pp. 43785-43797, 2018, doi: 10.1109/ACCESS.2018.2863270.

[2] B. Wang, B. Li, and H. Li, "Public Auditing for Shared Data with Efficient User Revocation in the Cloud," Proc. IEEE INFOCOM, pp. 2904-2912, 2013.

[3] 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, "A View of Cloud Computing," Comm. ACM, vol. 53, no. 4, pp. 50-58, Apr. 2010.

[4] 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. 598-610, 2007.

[5] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT'08), pp. 90-107, 2008.

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

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

[8] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," Proc. IEEE INFOCOM, pp. 525-533, 2010.

[9] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S. Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds," Proc. ACM Symp. Applied Computing (SAC'11), pp. 1550-1557, 2011.

[10] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards Secure and Dependable Storage Services in Cloud Computing," IEEE Trans. Services Computing, vol. 5, no. 2, pp. 220-232, Jan. 2012.

[11] Y. Zhu, G.-J. Ahn, H. Hu, S.S. Yau, H.G. An, and C.-J. Hu, "Dynamic Audit Services for Outsourced Storages in Clouds," IEEE Trans. Services Computing, vol. 6, no. 2, pp. 227-238, Apr. - June 2013.

[12] N. Cao, S. Yu, Z. Yang, W. Lou, and Y.T. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.

[13] R. Patil Rashmi, Y. Gandhi, V. Sarmalkar, P. Pund and V. Khetani, "RDPC: Secure Cloud Storage with Deduplication Technique," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 1280-1283, doi: 10.1109/I-SMAC49090.2020.9243442.