A Survey of Zero Knowledge Proof with Batch Auditing Based Privacy Preserving Authentication for Secure Cloud Storage

Faizan Ahmed[1], Avinash Sharma[2], Sriram Yadav[3]

[1]M.Tech (CSE, Student), [2]Associate Professor (HOD), [3]Associate Professor

(Department of CSE), MITS, Bhopal

Abstract - Using cloud storage, users can remotely store their information and explore the on-demand high-quality applications and services from a common pool of configurable computing resources, without the burden of local information storage and maintenance. Cloud services support great conveniences for the users to enjoy the on-demand cloud applications without accepting the local infrastructure limitations. However, the concept that users no longer have physical possession of the outsourced information, that makes the data integrity protection in cloud computing a formidable task, especially for users with constrained computing resources. The existing cloud security solutions mainly focus on the authentication to realize that a user's privative data cannot be unauthorized accessed. Moreover, users should be able to just accept the cloud storage as if it is local, without worrying about the need to verify its cloud integrity. Thus, enabling batch audit ability for cloud storage is of critical importance so that users can resort to a zero knowledge proof (ZKP) for authenticate the data to check the integrity of outsourced data and be worry free. To securely introduce an effective ZKP, the batch auditing process should bring in no new vulnerabilities toward user data privacy, and explore no additional online burden to user. In this paper, we propose a secure cloud storage system supporting ZKP with batch auditing (ZKP-BA). We further extend our result to enable the ZKP to perform security for multiple users simultaneously and efficiently. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient. Our preliminary experiment conducted on Amazon EC2 instance further explores the fast performance of the design.

Keywords - Cloud Computing, Authentication, Zero Knowledge Proof, Batch Auditing.

I. INTRODUCTION

Cloud computing is a efficient information technology outline for both industry and individuals. It launches an engaging data storage and interactive paradigm with provable advantages, including on-demand self-services, present network access, and location independent resource pooling [1]. Towards the cloud computing, veritable service architecture is anything as a service (XaaS)**,** in which infrastructures, platform, software, and separate are applied for Present interconnections. Previous studies have been worked to enhanced the cloud computing specialize towards the internet of services [2], [3]. Afterward, security and privacy issues are seemly key concerns with the increasing popularity of cloud services. Formally security approaches primarily target on the strong authentication to realize that a user can remotely access its own data in on-demand mode**.** Along with the status of the application requirements, users may want to access data and share each other's authorized data Slots to achieve profitable benefits**,** which collects new authentication and privacy challenges for the cloud storage. An example is informs to recognize the main motivation. Inside the cloud storage based supply chain management, there are lots of interest groups (e.g., supplier, carrier, and retailer) in the structure**.** Every group owns its users which are allowing accessing authorized data fields, and unlike users owning relatively independent access authorities**.** It means that any two users from various groups should access unlike data fields of the same file. There into, a supplier advisedly may like to access a carrier's data fields, but it is not sure whether the carrier will allow its access request. If the carriers regret its request, the supplier's access desire will be revealed along with nothing obtained towards the desired data fields. But in real, the supplier may not send the access request or withdraw the unaccepted request in advance if it industry knows that its request will be decline by the carrier. It is illogical to exhaustively disclose the supplier's personal information without any privacy exploration. Figure 1 illustrates three revised cases to address above undetectable privacy issue.

• Case 1: The carrier also wants to access the supplier's data fields, and the cloud server should confirm each other and displace the shared access authority to the both users;

• Case 2: The carrier has no required on other users data fields, therefore its authorized data fields should be properly secured, meanwhile the supplier's access appeal will also be invisible;

• Case 3: The carrier may want to access the retailer's data fields, but it is not definite whether the retailer would accept

its appeal or not. The retailer's authorized data fields should not be public if the retailer has no interests in the carrier's data fields, and the carrier's request is also privately hidden.

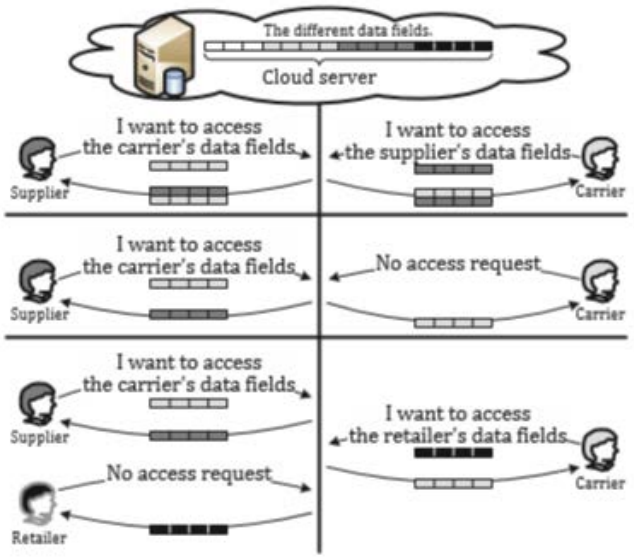

Figure 1: Three possible cases at the time of data accessing and data sharing in cloud applications.

II. SYSTEM MODEL

Following Figure 2 illustrates a structure model for the cloud storage architecture, which involve three most commonly network entities: users (U_x), a cloud server (S), and a trusted third party.

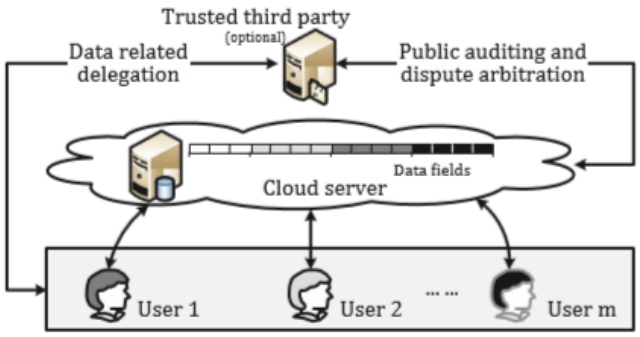

Fig 2: The cloud storage system model.

• User: A single or group entity, which owns its data stored in the cloud for online data storage and computing. Unlike users may be related to with a common industry, and are assigned with independent authorities on different data fields**.**

• Cloud server: an entity, which is managed by a particular cloud service provider or cloud application operator to provide data storage and computing services. The cloud server is consider as an entity with unrestricted storage and computational resources**.**

• Trusted third party: an optional and neutral entity, which has advanced capabilities on position of the user, to perform data public auditing and contestation arbitration.

In the cloud storage, a user remotely saved its data via online infrastructures, flat forms, or software for cloud service, that are run in the distributed, parallel, and cooperative modes. At the time of cloud data accessing, the user independently interacts with the cloud server without external interferences, and is assigned with the whole and independent access on its own data fields. It is required to guarantee that the users' outsourced data cannot be unauthorized accessed by the another users, and is of unfavorable importance to assure the private information during the users' data access challenges. In some scenarios, there are multiple users in a system (e.g., supply chain management), and the users could have unlike associated attributes from different interest groups. One of them a users may want to access other associate users' data fields to achieve bi-directional data sharing, but that is cares about two aspects: whether the goal user will like to share its data fields, and how cannot expose its access request if the aimed user declines or ignores its challenge. On the paper, we pay lots of attention on the operation of data access control and access authority sharing other than the proper file oriented cloud data transmission and its management.

In the system model, assume the point-to-point communication channels between the users and a cloud server are certain with the protection of secure shell protocol (SSH). The related authentication acknowledgment is not highlighted in the following protocol presentation. Towards the trust model, there is no trustable relationship between a cloud server S and a user Ux.

• S is semi-honest and curious: Being semi-honest means that S can be regarded as an entity that appropriately follows the protocol procedure. Being curious means that S may attempt to get Ux's private information (e.g., data content, and user preferences). It means that S is under the supervision of the cloud provider or operator, but it may be interested in accessing users' privacy. In the passive or honest- but-curious model, S cannot manipulate with the user data to maintain the system normal operation with undetected monitoring.

• Ux is rational and sensitive: To be rational means that Ux's behavior will be never depend on experience or emotion, and misbehavior may only occur for selfish interests. To be sensitive means the Ux is disinclined to disclosure its sensitive data, but it has strong interests in another users' privacy. Towards the threat model, it covers all the possible privacy threats and system vulnerabilities at the time of cloud data interactions. The communication channels are open in

public, and both internal and external attacks exist in the cloud applications [15]. The internal attacks mainly relate to the interactive entities (i.e., S, and Ux). There into, S may be self-centered and useful, and aims to get more user data contents and the associated user behaviors/habits for the maximization of commercial interests; Ux may attempt to capture other users' delicate data fields for certain purposes (e.g., curiosity, and malicious intent). The external attacks mainly consider the data CIA triad (i.e., privatives, integrity, and availability) threats from outside adversaries, which can compromise cloud data storage servers, and after modify (e.g., insert, or delete) the users' data fields**.**

III. PREVIOUS WORK

Dunning et al. [11] proposed an unknown ID assignment based data sharing algorithm (AIDA) for multiparty oriented cloud and distributed computing systems. In the AIDA, an integer data sharing algorithm is designed to the top of secure sum data mining operations, and takes a variable and limitless number of iterations for anonymous assignment. Specifically, Newton's identities and Sturm's theorem are used for the data mining, a distributed solution of certain polynomials over finite fields raise the algorithm scalability, and Markov chain representations are used to define statistics on the required number of iterations. Liu et al. [12] proposed a multi-owner data sharing secure scheme (Mona) for dynamic groups in the cloud applications. The Mona aims to realize that a user can securely share its data with other users via the untrusted cloud server, and can expeditiously support dynamic group interactions. In the scheme, a new granted user can directly decrypt data files without pre-contacting with data owners, and user abrogation is achieved by a revocation list without updating the secret keys of the remaining users. Access control is applied to ensure that any users in a group can unidentified utilize the cloud resources, and the data owners' real identities can only be revealed by the group manager for dissipation, arbitration. It indicates the storage overhead and encryption computation cost are independent with the amount of the users. Grzonkowski et al. [13] proposed a zero-knowledge proof (ZKP) based authorization schemes for sharing cloud services. Based on the social home networks, a user centric approach is applied to enable the sharing of personalized content and temporal network-based services via TCP/IP infrastructures, in which a trusted third party is introduced for decentralized interactions. Nabeel et al. [14] proposed a broadcast group key management (BGKM) to improve the weakness of symmetric key cryptosystem in public clouds, and the BGKM actualize that a user need not utilize public key cryptography, and can dynamically derive the symmetric keys during decryption. Accordingly, attribute based access control mechanism is designed to get that a user can decrypt the contents if and only if its identity attributes satisfy the content provider's policies. The fine grained algorithm applies access control vector (ACV) for assigning secrets to users based on the identity attributes, and allowing the users to derive actual symmetric keys based on their secrets and other public information. The BGKM has an obvious advantage during adding/revoking users and updating access control policies.

IV. PROPOSED METHOD

We consider a cloud data storage service involving three different entities, as illustrated in the cloud user, who has big amount of data files to stored inside cloud storage; the cloud server, which is owned by the cloud service provider to provide data storage service and has significant storage. To fully ensure the data integrity and save the cloud users computation resources as well as online concern, it is of critical importance to enable authentication and batch auditing service for cloud data storage, so that users may resort to an independent third party auditor (TPA) to audit the outsourced data when it is needed. The TPA, who has sophisticate and capabilities that users do not, can periodically check the integrity of all the data stored in the cloud on behalf of the users, which provides a much more easier and affordable way for the users to ensure their storage correctness in the cloud. Furthermore, in addition to help users to conceive the risk of their subscribed cloud data services, the audit result from TPA will also be beneficial for the cloud service providers to enhance their cloud based service platform, and even serve for independent arbitration purposes. In a one word, enabling public auditing services would play an important role for this nascent cloud economy to become fully established; where users will need ways to assess risk and gain trust in the cloud. The proposed architecture of ZKP-BA is as follows:

A. Role of TPA: To enable security for cloud data storage using ZKP (Zero Knowledge Proof), This protocol design should accomplish the following security and performance guarantee:

(1) Public Audit ability: allow TPA uses ZKP to verify the rightness of the cloud data on demand without retrieving a copy of the whole data or introducing additional on-line burden to the cloud users.

(2) Storage Correctness: guarantee that there exists no cheating cloud server that can pass the audit from TPA without indeed storing user data inviolate.

(3) Privacy-Preserving: guarantee that there exists no way for TPA to accelerate users data content from the information collected during the auditing process.

(4) Batch Auditing: enable TPA with batch auditing with secure and efficient auditing capability to cope with auditing delegations from probably large number of different users simultaneously.

(5) Lightweight: allow TPA to perform auditing with minimum communication and computation overhead. We are introducing an attacking module which will keep continuously track on the data alteration in the cloud if any, and will inform the user about the altered data. Attacking module will be in the form of small code to modify the database directly so that entry is sabotaged. This code will dwell on cloud server. Also the timer is going to be implemented where task may be schedule for one time execution, or for repeated execution at regular intervals and also we adapt some efficient servers for better performance and increase the speed of execution, such as glassfish server.

The general process of using zero-knowledge proof protocol is shown in following figure:

(1) The prover P sent promise random number r to the verifier V.

(2) V sent random challenge value e to P.

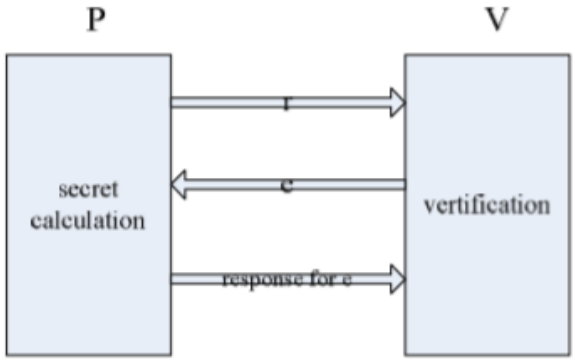

Figure 3: The general process of zero-knowledge proof protocol

(3) P calculates secretly and sent the result to V as the challenge-response for second step.

(4) V verifies the response. If the verification fails, the process of proof will end. Otherwise, the above steps will be repeated for N times. If every verification can be successful, V will receive P's proof in great probability.

B. Privacy preserving module: Homomorphism authenticators are unforgettable verification metadata generated from individual data blocks, which can be securely aggregated in such way to guarantee an auditor that a linear combination of data blocks is appropriately computed by verifying only the aggregated authenticator. Hence, to achieve privacy-preserving public auditing, we propose to unambiguously integrate the homomorphism authenticator with random mask technique. In our protocol, the linear collection of sampled blocks in the server response is masked with randomness generated by a pseudo random function (PRF) [9].

C. Batch auditing module: Through the organization of privacy-preserving batch auditing in Cloud Computing, TPA may concurrently handle multiple auditing delegations upon different user requests**.** The several, auditing of these tasks for TPA can be and very difficult and inefficient. Batch auditing not only allows TPA to perform the multiple auditing tasks at the same time, but also greatly reduces the computation cost on the TPA side This is because of aggregating K verification equations into helps to reduce the number of quite expensive paring operation from 2k, as required in individual auditing, to K+1, by which saves a considerable amount of auditing time [9].

V. CONCLUSIONS

Zero knowledge proof protocol has become a very important component in cryptographic algorithms and security protocols in cloud computing. In this paper, the main idea, nature, general proof process, mathematical theory and specific applications of zero-knowledge proof with batch auditing protocol are introduced. Zero-knowledge proof protocol has the advantage of zero leakage proof, so it can be applied to prove many key issues, like many classic mathematical problem the polynomial function roots, the graph isomorphism, as well as other NP problem, such as the Sudoku games. The user can prove that he has the method to solve some problem and he does not worry about the method revealed. Zero-knowledge proof protocol is very useful in the field of network and information security too, like authentication, digital signatures, etc. It is very important that proving to each other identify of the user without revealing the user information in authentication and digital signatures. In order to effectively prevent unauthorized users impersonating legitimate users, we can use zero- knowledge proof protocol to authenticate.

VI. FUTURE WORK

The future work is as follows:

1. We can use OTP, RES, AES, triple-DES, Digital Signature for authenticate cloud data.

2. We can use other auditing scheme like public auditing, data dynamics auditing with their authentication protocol.

REFERENCES

[1] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Storage Security in Cloud Computing", Proc. IEEE INFOCOM '10, (2010).

[2] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing", http://csrc.nist.gov/groups/SNS/cloud-computing/index.html, (2009).

[3] M. Armbrust, A. Fox, "Above the Clouds: A Berkeley View of Cloud Computing", Technical Report UCB-EECS-2009-28, Univ. of California, Berkeley, (2009).

[4] Cloud Security Alliance, "Top Threats to Cloud Computing", http://www.cloudsecurityalliance.org, 2010.

[5] M. Arrington, "Gmail Disaster: Reports of Mass Email Deletions", http://www.techcrunch.com/2006/12/28/gmail-disasterreports- of-mass-email-deletions/, (2006).

[6] J. Kincaid, "MediaMax/TheLinkup Closes Its Doors", http:// www.techcrunch.com/2008/07/10/mediamaxthelinkup-closes-its-doors/, (2008).

[7] Amazon.com, "Amazon s3 Availability Event: July 20, 2008", http://status.aws.amazon.com/s3-20080720.html, (2008).

[8] 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).

[9] 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, (2007).

[10] Lindell, Zarosim, "Adaptive Zero-knowledge Proofs and Adaptively Secure Oblivious Transfer", Journal of Cryptology.24(4),pp.761-799, (2011).

[11] Bayer, Stephanie, "zero-knowledge argument for correctness of a shuffle", 31st annual International Conference on the Theory and Applications of Cryptographic Techniques, EUROCRYPT (2012).

[12] 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 Symp. Research in Computer Security (ESORICS '09), pp. 355-370, (2009).

[13] 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), (2007).

[14] 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, (2007).

[15] 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).

AUTHOR'S PROFILE

Faizan Ahmed has received his Bachelor of Engineering degree in Computer Science & Engineering from All Saints College of Technology, Bhopal in the year 2012. At present he is pursuing M.Tech with the specialization of Computer Science Engineering in Millennium Institute of Technology and Science, Bhopal. His area of interest is Data Mining, Image Processing etc.

Avinash Sharma has received his M.Tech from BUIT, Bhopal in the year 2011. At present he is working as an Associate Professor and also Head of CSE Department at Millennium Institute of Technology and Sciences, Bhopal. His areas of interests are Data Mining, Image Processing, Clouds Computing etc.

Sriram Yadav has received his M.Tech from Berhampur University, Orrissa. Pursuing PhD in CSE from P.A.H.A.R University, Udaipur (Rajasthan). At present he is working as an Associate Professor in Millennium Institute of Technology and Sciences, Bhopal. His areas of interests are Data Mining, Image Processing, Cloud Computing, Grid Computing, Video Processing etc.