MJRET
Open Access
Multidisciplinary Journal of Research in Engineering and Technology
Journal homepage: www.mjret.in
ISSN:2348 - 6953

REVIEW ON AUDITING OF DYNAMIC BIGDATA ON CLOUD WITH FINED GRAINED UPGRADES

Shubham Burghate, Akash Chourange, Rahul Ghyar,
Akshay Dighe, Prof. Yogesh Thorat

Computer Engineering, Dr. D.Y.Patil, School of Engineering, Pune, India

Abstract: *Cloud is a new technology generally available as a service to anyone on the internet. Cloud computing is a computing model, where resources like computing power, storage, network and software are abstracted and provided as services on the internet in a remotely accessible fashion. In cloud computing, data while transferring/sharing as well as storage, users store their data in the cloud and get it back when it is needed. But there is no assurance about the security of that data stored and also it is not changed by the cloud or Third Party Auditor (TPA). Therefore Data Security is one of the major issues of using cloud. Process of verification is called data auditing and TPA is called an auditor. In existing research work, which already check for the data integrity, but having some drawbacks like the basic and mostly needed authorization/authentication process is not present in between Cloud Service provider and TPA. Some of the recent research which is based on Method like BLS signature can support dynamic updates but they only has support updates for fixed-size data blocks. In this paper, we provide support for authorized auditing and fine-grained update requests. The basic goal of our scheme, we propose an Service that can reduce communication overhead for verifying small updates, because now-a-days for big data applications having large number of frequent small changes, such as application in social media. In our Scheme, The CSS (Cloud Storage Server) , CSP (Cloud Server Provider) and T.P.A (Third Party Auditor) are the important things.CSS provide the facility of data virtualization by spreading it in various distributed data centre and data on demand service. The communication among CSS and cloud user is monitored and controlled by TPA (Third Party Auditor). The T.P.A. will perform overall the detailed operation to authenticate each user.*

Keywords*: Cloud computing, data security, BigData, provable data control, authorized auditing, fine-grained dynamic data update, TPA (Third Party Authenticator)*

1. INTRODUCTION

Cloud is a new technology generally available as a service to anyone on the internet. Cloud computing is a computing model, where resources like computing power, storage, network and software are abstracted and provided as services on the internet in a remotely accessible fashion. In cloud computing, data while transferring/sharing as well as storage, users store their data in the cloud and get it back when it is needed. Cloud service providers provide an enterprise-class infrastructure that offers a scalable, secure and also reliable good environment for users, at a much less cost due to the sharing nature of resources. It is good way for users to use cloud storage services to divide data with others in a team, as data sharing becomes a standard feature in most cloud storage scenarios. As Compared to traditional systems, scalability and elasticity are key advantages of cloud. As such, efficiency in supporting dynamic data is of great importance.

The integrity of data in cloud storage is subject to uncertainty and analysis, while data stored in an untrusted cloud can simply be lost or corrupted, due to hardware failures or human errors. To keep the integrity of cloud data, it is best to carry out public auditing by introducing a third party auditor (TPA), who offers its auditing service with more dominant computation and communication abilities than normal users. Even though existing data auditing schemes previously have different properties, potential risks and inefficiency such as security risks in an unauthorized auditing requests and inefficiency in processing small updates still exist. Here, we will focus on enhanced support for small dynamic updates, which benefits the scalability and efficiency of a cloud storage server.

2. PROBLEM DESCRIPTION

The identity of the signer on every block in shared data is private and confidential for the user. During this procedure of auditing, a semi-trusted TPA, who is merely responsible for auditing the integrity of shared data, may try to disclose the identity of the signer on every block in shared data based on verification information. Once the TPA discloses the individuality of the signer on every block, it can easily distinguish a high-value target.

It is especially recommended that data auditing is to be conducted on a regular basis for the users who have high-level security demands over their data. A necessary authorization/authentication process is not there in between the auditor and cloud service provider anyone can challenge the service provider for a proof of integrity of certain file.

In our purposed scheme for the first time, we formally analyze different types of fine-grained dynamic data update requests on variable-sized file blocks in a single dataset. To the best of our knowledge, we are the first to propose a public auditing scheme based on BLS signature and Merkle hash tree (MHT) that can support fine-grained update requests. Compared to existing schemes, our scheme supports updates with a size that is not restricted by the size of file blocks, thereby offers extra flexibility and scalability compared to existing schemes.
For better security, our scheme incorporates an additional authorization process with the aim of eliminating threats of unauthorized audit challenges from malicious or pretended third-party auditors, which we term as ‘authorized auditing’.

We investigate how to improve the efficiency in verifying frequent small updates which exist in many popular cloud and big data contexts such as social media. Accordingly, we propose a further enhancement for our scheme to make it more suitable for this situation than existing schemes. Compared to existing schemes, both theoretical analysis and experimental results demonstrate that our modified scheme can significantly lower communication overheads.

3. SYSTEM ARCHITECTURE

To enable the TPA efficiently and securely verify shared data of users, Our system should be designed to achieve following properties:

(1) **Public Auditing:** The third party auditor is able to verify the integrity of shared data without retrieving the entire data.

(2) **Correctness:** The third party auditor is able to correctly detect whether there is any corrupted block in shared data.

(3) **Enforceability:** Only a user which is valid can generate valid verification information on shared data.

(4) **Identity Privacy:** During auditing, the TPA cannot distinguish the identity of the signer on each block in shared data.

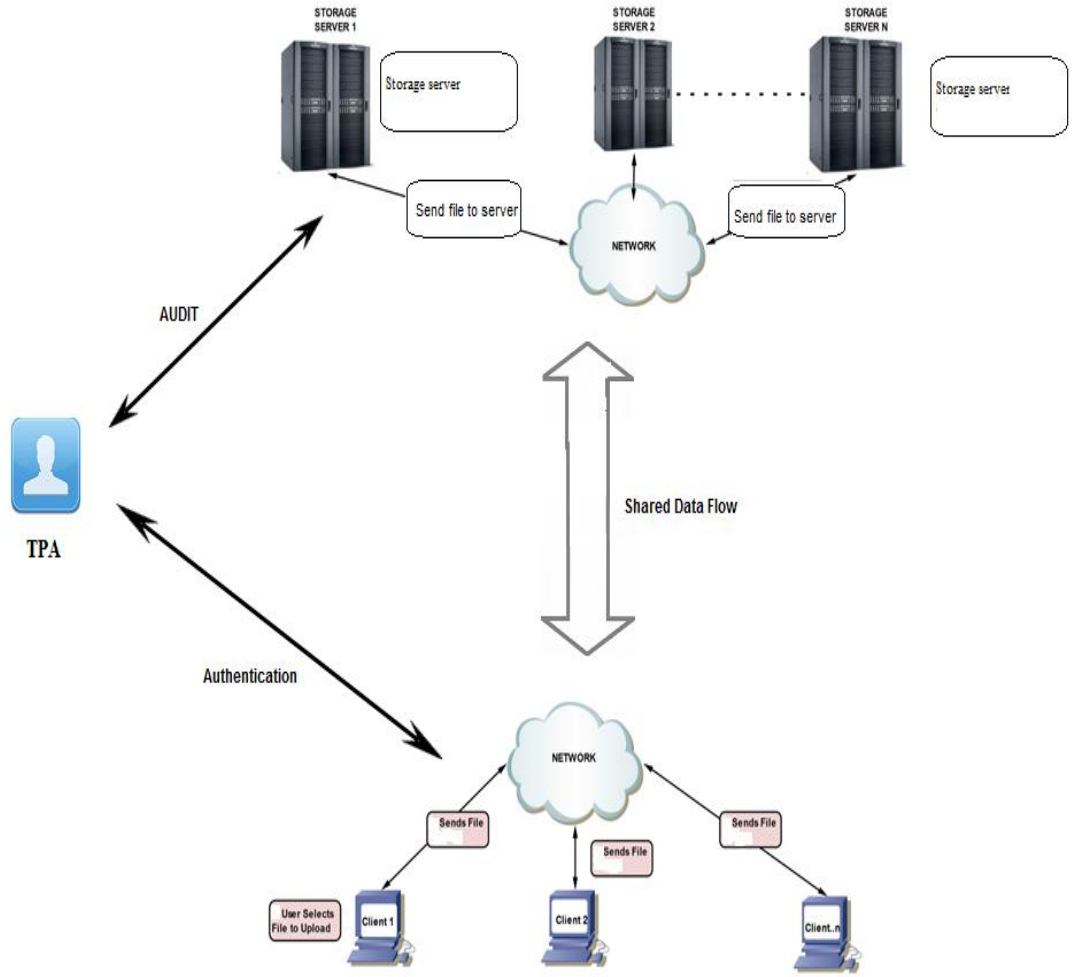

Fig.1: Architecture of proposed system

Architecture of our proposed system main goal is to support variable sized data block, authorized TPA (third party auditing), and fined grained updates.

4. RELATED WORK

Ralph C. Merkle[1] proposed new scheme "Digital signature based on a conventional encryption function such as **DES** is described which is more secure as the basic function. Existing Research like "New direction in cryptography" in 1976 and "Making the digital signature legal and safeguard" by S.M.Lipton, S.M.Matyas in Feb. 1978. Previous work of Ralph C.Merkle in 1982 "Secrecy, Authentication, and public key systems these all rely on conventional encryption functions **like one-way-function.** But not single among them are much succeed in providing the convenience of system based on more complex mathematical problem. **Ralph C. Merkle[1]** provides advantage to reduce computational cost as compared with system which require modular arithmetic. The data encryption standard software implementation which runs faster than exponentiation modulo N, because of this digital signature system which is based on use of DES would get benefit from it.DES chips are already available at low cost of different manufacturer. New digital signature system is very fast when Retro fitted to a system that already has a DES chip. In this paper, they describe how new one time signature system can be used in a new way to provide a digital signature system that overcomes the limitation of previous. The general idea in this new system is to use an infinite tree of one time signature.

Suthan and Kesavaraja[2] proposed scheme " Granule based File Storage System with Secure Transparent Availability". In this system, they focus on various security algorithms to provide security to the data that we store. The file is to be splitted into n number of different particles and this particles be distributed within the system providing transparency to the user. The GFMS makes sure that the file is splitted and spread inside the system in such a way that, even if some part of the file is retrieved no information can be recognized. **Three major issues: Spyware, Encryption, File splitting so for Security purpose- The** file is Encrypted with AES algorithm. and the secure key is generated for AES File security by using merkle damagard hash construction. This hash key is used for AES Encryption process.

Cong Wang, Qian Wang, Kui Ren and Wenjing Lou[3] proposed system in which they focus on cloud data storage security, which has always been an very much important aspect of quality of service. To make sure the correctness of users data in the cloud. And also an effective and flexible distributed scheme with two salient features, opposing to its predecessors. By utilizing the homomorphic token with distributed verification of erasure-coded data, our scheme achieves the integration of storage correctness insurance and data error localization. Algorithms used- token Pre-computation, correctness verification & error localization, error recovery. Advantages are providing dynamic operations support and security strength against weak advisory.

Suganya .S, Mrs. Sumathi[8] they proposed system in the recent years, the Network-Attached Storage (NAS) and the Network File System (NFS) provide storage devices over the network so that user can access the storage devices through network connection.

Drawbacks: Encryption schemes supports confidentiality of the data, but the functionality of the storage system is limited because only certain operations are supported. Data robustness is the important need for storage systems. Participating parties in the auditing scheme Rank-based Merkle hash tree.

Ari Juels and Burton S. Kaliski Jr Proofs of **retrievability [4]** for Large Files some facility provided to user like archive and back up service provide to the user. User can retrieve the target file into the storage/server.POR can specially design to handle large file cryptographic techniques help users ensure the privacy and integrity of files they retrieve. In this system for users those want to verify that archives do not delete or modify files prior to retrieval. The main advantages of a POR are to accomplish these checks *without users having to download the target files themselves.* POR can be efficient enough to provide regular checks of file **retrievability**. They introduce a POR protocol in which the verifier stores only a single cryptographic key—irrespective of the size and number of the files whose retrievability it seeks to verify—as well as a small amount of dynamic state (some tens of bits) for each file. it is worth considering a straightforward design involving a keyed hash function. Data-integrity protection is one of the fundamental goals of cryptography. Primitives such as digital signatures and message-authentication codes (MACs), are used in POR(Proofs of Retrievability).POR protocol encrypts file and randomly embeds a set of randomly-valued check blocks called *sentinels*. The verifier challenges the prove by specifying the positions of a collection of sentinels and ask the prove to return the associated sentinel values. Some challenges to accept in this system First, they offer a formal, concrete security definition of PORs that we believe to be of general interest and applicability in practical settings. Second, they introduce a sentinel-based POR scheme with several interesting properties, such as its uses function key generates secret key and encoded the file which is store in server, extraction, response and verify function used in this system to check the file will secure or not when we retrieve the file.

5. COMPARISON

The idea of proofs of retrievability (POR) and its first model was proposed by Jules et al. [4]. but their system can only be applied to static data storage such as records or files. As such, good organization in behind dynamic data is of great importance. Protection and privacy protection on dynamic data has been considered extensively in the past [3], [7]. "Granule based File Storage System with Secure Transparent Availability" by Suthan and Kesavaraja[2] suffering from drawbacks such as it only provide an security until the intruders get hold of the file. "Digital signature based on conventional encryption function"[1] by Ralph C. Merkle which is symmetric encryption technique."Dynamic Big Data Storage on Cloud with Efficient Verifiable Fine-grained Updates using Secure Erasure Code-Based Cloud Storage System"[8] which uses concept of proxy servers but it's limited for certain file types. "Proofs of retrievability for Large Files"[4] by Ari Juels provide checks for regular retrievability but it can only protect the static archived files. In any of this application public auditability and variable sized data block are not supported by default. Compared to conventional systems, scalability and flexibility are key advantages of cloud [5], [6], Cloud users may also want to split big datasets into smaller datasets and store them in different physical servers for reliability, privacy-preserving or efficient processing purposes. Among the most vital problems linked to cloud is data security/privacy [1], [8]. It has been one of the most recurrently raised concerns. There is a lot of work trying to enhance cloud data

security/privacy with technological approaches on CSP side.

6. CONCLUSION

We have provided a formal analysis on possible types of fine-grained data updates and proposed a scheme that can fully support authorized auditing and fine-grained update requests. Based on our scheme, we have also proposed a modification that can dramatically reduce communication overheads for verifications of small updates.

Based on the contributions of this paper on better data auditing, we plan to additional investigate the next step on how to develop other server-side protection methods for efficient data security with effective data confidentiality and availability. Also, we also plan to investigate audit ability-aware data scheduling in cloud computing**.** For providing more security we are using TPA(Third party authenticator). Which is able to verify our data from cloud and check our data's integrity. We are providing authenticity to the TPA using MD5 hashing algorithm which is going to perform main function in our system. It will allow to achieve us the security of our data from TPA also. MD5 hashing algorithm gives 128 bit hash key which is allocate to every TPA which should be given at the time of verifying data at cloud.

ACKNOWLEGEMENT

We would like to thank Dr.D.Y.Patil School of Engineering for providing us with all the required amenities. We would thank our guide Prof. Yogesh Thorat sir for giving us all the help and guidance we needed. We are also grateful to Prof. S. S. Das, Head of Computer Engineering Department, DYPSOE, Lohegaon, Pune for their indispensable support, suggestions and motivation during the entire course of the project.

REFERENCES

[1] Digital signature based on conventionl encrypion function" by Ralph C. Merkle in 1998.
[2] Granule based File Storage System with Secure Transparent Availability " by Suthan and Kesavaraja in 2011.
[3] Ensuring Data Storage Security in Cloud Computing" by Cong Wang, Qian Wang, Kui Ren and Wenjing Lou in 2012
[4] Proofs of retrievability for Large Files" by Ari Juels and Burton S. Kaliski Jr in 2007.
[5] Cloud Computing and Emerging IT Platforms: Vision, Hype, Reality for Delivering Computing as the 5th Utility,'' Future Generation Computation System," by . R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic in June 2009.
[6] A View of Cloud Computing,'' by M.Armbrust, A.Fox,R.Griffith, A.D.Joseph, R.Katz,A.Konwinski, G.Lee, D.Patterson, A.Rabkin, I.Stoica, and M.Zaharia Commun in April 2010.
[7] Scalable and Efficient Provable Data Possession," by G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, in 2008.
[8] Dynamic Big Data Storage on Cloud with Efficient Verifiable Fine-grained Updates using Secure Erasure Code-Based Cloud Storage System" by Suganya .S, Mrs.Sumathi in 2014.