A Review on Towards Audit of Secure Storage Correctness Verification and Dependable Storage Error Recovery in Cloud Computing

Gadekar Anant[#], Dr. S. Lomte[*], Prof. Neha Walmik

[#]*Department of Computer Science & Engineering, Dr. B. A.. M. University, Aurangabad*

Everest College of Engineering & Technology, Aurangabad

Abstract: Cloud Computing provide reliable, cost effective and scalable method for delivery of computing and delivery of data. To Store the data in cloud system without burden hardware and software management on remote site and enjoying the on demand cloud application.Entering the benefits are clear, such a service is also 'metaphorically surrender users' physical control of their outsourced data, which avoid position new security risks in direction of correctness of the data in cloud. In order to overcome the this new problem and further enhancement a secure and dependable cloud storage service, we propose in this paper a flexible distributed storage integrity third party auditing mechanism, utilizing the homomorphism token and distributed ensure-coded data with error recovery.The proposed design allows users to audit the cloud storage with less computation cost and lightweight communication. The auditing result shows not only ensures strong cloud storage correctness verification guarantee, but also simultaneously achieves data error localization, i.e., the identification of misbehaving server. Considering the cloud data are dynamic in nature, with consideration block modification, deletion and append the proposed system design further supports secure and efficient dynamic operations on outsourced data. Analysis shows the proposed scheme is highly efficient and flexible against Byzantine failure, malicious data modification attack, and even server colluding attacks. This article also focuses on parallel data processing for dynamic resource allocation for both task scheduling and execution.

Index Terms—Data integrity, dependable distributed storage, error localization, data dynamics, Cloud Computing.

I. INTRODUCTION

Cloud Computing has began to receive mass attract in corporate organizations as it makes the data center be able to work like the Internet to share and access resources in safe and secure manner. CLOUD computing has been envisioned as the next- generation information technology (IT) architecture for enterprises, due to its long list of unprecedented advantages in the IT history: on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing and transference of risk [2]. To provide data storage service, Cloud Computing utilizes network of enormous amount of servers generally running lower cost customer PC technology with peculiar connections to disperse data processing tasks across end users.As a disruptive technology with profound implications, cloud computing is transforming the very nature of how businesses use information technology. One fundamental aspect of this paradigm shifting is that data are being centralized or outsourced to the cloud. From users' perspective, including both individuals and IT enterprises, storing data remotely to the cloud in a flexible on-demand manner brings appealing benefits: relief of the burden for storage management, universal data access with location independence, and avoidance of capital expenditure onhardware, software, and personnel maintenances, etc., [3].

Moving data into the cloud offers great convenience to users since they don't have to care about the complexities of direct hardware management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud EC2) [2] are both well known examples. While these internet-based online services do provide huge amounts of storage space and customizable computing resources, his computing platform shift, however, is eliminating the responsibility of local machines for data maintenance at the same time. As a result, users are at the mercy of their cloud service providers for the availability and integrity of their data [3]. On the one hand, although he cloud infrastructures are much more powerful and reliable than personal computing devices, broad range of both internal and external threats for data integrity still exist. Examples of outages and data loss incidentsof noteworthy cloud storage services appear from time to time [4]–[8]. On the other hand, since users maynot retain a local copy of outsourced data, there exist various incentives for cloud service providers (CSP) to behave unfaithfully towards the cloud users regarding the status of their outsourced data. For example, to increase the profit margin by reducing cost, it is possible for CSP to discard rarely accessed data without being detected in a timely fashion [9]. Similarly, CSP may even attempt to hide data loss incidents so as to maintain a reputation [10]–[12]. Therefore, although outsourcing data into the cloud is economically attractive for the cost and complexity of long-term large-scale data storage, it's lacking of offering strong assurance of data integrity and availability may impede its wide adoption by bothenterprise and individual cloud users. Reason for moving into Cloud is simply because of Cloud allows users to access applications from anywhere at any time through internet. But in past, consumers run their programs and applications from software which downloaded on physical server in their home or building. Cloud provides benefits such as flexibility, disaster recovery, software updates automatically, pay-per-use model and cost reduction. However Cloud also includes major risks such as security, data integrity, network dependency and centralization. When storing customer'sdata into cloud data storage, security plays a vital role. Sometimes customers store some sensitiveinformation in cloud storage environment. This causes some serious security issues. So providing security to such sensitive information is one of the difficult problems in Cloud computing. In preceding works, several methods areproposed for securely storing data into Cloud. This paper discussed those methodologies and various techniques to effectively store data. Also analyzed the advantages, drawbacks of those techniques and provides some directions for future research work.

We also provide the extension of the proposed main scheme to support third-party auditing, where users can safely delegate the integrity checking tasks to third-party auditors and are worry-free to use the cloud storage services. Our work is among the first few ones in this field to consider distributed data storage security in Cloud Computing. Our contribution can be summarized as the following three aspects:

1) Compared to many of its predecessors, which only provide binary results about the storage status across the distributed servers, the proposed scheme achieves the integration of storage correctness insurance and data error localization, i.e., the identification of misbehaving server(s).

2) Unlike most prior works for ensuring remote data integrity, the new scheme further supports secure and efficient dynamic operations on data blocks, including: update, delete and append.

3) The experiment results demonstrate the proposed scheme is highly efficient. Extensive security analysis shows our scheme is resilient against Byzantine failure, malicious data modification attack, and even server colluding attacks.

II. PROBLEM STATEMENTS

A. System Model

The System and Threat Model

We consider a cloud data storage service involving three different entities, as illustrated in Fig. 1: the cloud user, who has large amount of data files to be stored in the cloud; the cloud server, which is managed by the cloud service provider to provide data storage service and has significant storagespace and computation resources (we will not differentiate CS and CSP hereafter); the third-party auditor, who has expertise and capabilities that cloud users do not have and is trusted to assess the cloud storage service reliability on behalf of the user upon request. Users rely on the CS for cloud data storage and maintenance. They may also dynamically interact with the CS to access and update their stored data for various application purposes. As users no longer possess their data locally, it is of critical importance for users to ensure that their data are being correctly stored and maintained. To save the computation resource as well as the online burden potentially brought by the periodic storage correctness verification, cloud users may resort to TPA for ensuring the storage integrity of their outsourced data, while hoping to keep their data private from TPA.

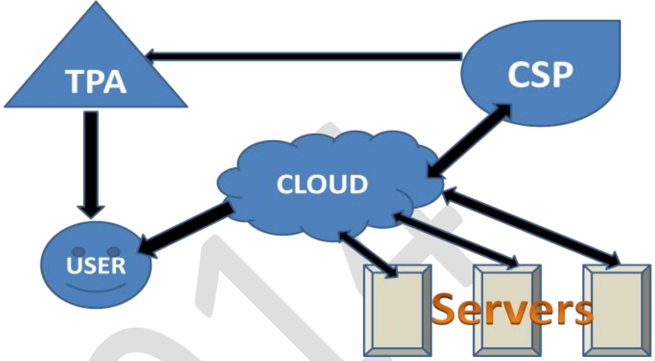

Fig. 1: Cloud storage service architecture

We assume the data integrity threats toward users' datacan come from both internaland external attacks at CS. These may include: software bugs, hardware failures, bugs in the network path, economically motivated hackers, malicious or accidentalmanagement errors, etc. Besides, CS can be self-interested. For their own benefits, such as to maintain reputation, CS might even decide to hide these data corruption incidents to users. Using third-party auditing service provides a cost-effective method for users to gain trust in cloud. We assume the TPA, who is in the business of auditing, is reliable and independent. However, it may harm theuser if the TPA could learn the outsourced design should achieve the following security and performance guarantees:

1. Public auditability: to allow TPA to verify the correctness of the cloud data on demand without retrieving a copy of the whole data or introducing additional online burden to the cloud users.
2. Storage correctness: to ensure that there exists no cheating cloud server that can pass the TPA's audit without indeed storing users' data intact.
3. Privacy preserving: to ensurethat the TPA cannot derive users' data content from the information collected during the auditing process.
4. Batch auditing: to enable TPA with secure and efficient auditing capability to cope with multiple auditing delegations from possibly large number of different users simultaneously.
5. Lightweight: to allow TPA to perform auditingwith minimum communication and computation overhead.

III. THE PROPOSED SCHEMES

This section presents our public auditing scheme which provides a complete outsourcing solution of data—not only the data itself, but also its integrity checking. After introducing notations and brief preliminaries, we start from an overview of our public auditing system and discuss two straightforward schemes and their demerits. Then, we present our main scheme and show how to extent our main scheme to support batch auditing for the TPA upon delegations from multiple users. Finally, we discuss how to generalize our privacy-preserving public auditing scheme and its support of data dynamics.

STORAGE TECHNIQUES IN CLOUD COMPUTING

In this section, various existing techniques have been dies-cussed. Cloud storage is regarded as a system of disseminated data centers that generally utilizes virtualization technology and supplies interfaces for data storage.

A. Implicit Storage Security to Data in Online

Providing implicit Storage Security to data in Online is more beneficial in a cloud environment. Presented implicit storage security architecture for storing data where security is disseminated among many entities [8] and also look at some common partitioning methods. So data partitioning scheme is proposed for online data storage that involves the finite field polynomial root. This strategy comprises of two partitioning scheme. Partitioned data are saved on cloud servers that are chosen in a random manner on network and these partitions are regained in order to renovate the master copy of data. Data pieces are accessible to one who has knowledge of passwords and storage locations of partitioned pieces.

B. Identity-Based Authentication

In Cloud Computing, resources and services are distributed across numerous consumers. So there is a chance of various security risks. Therefore authentication of users as well as services is an important requirement for cloud security and trust. When SSL Authentication Protocol (SAP) was employed to cloud, it becomes very complex. As an alternative to SAP, proposed a new authentication protocol based on identity which is based on hierarchical model with corresponding signature and encryption schemes [2]. Signature and encryption schemes are proposed to achieve security in cloud communication. When comparing performance, authentication protocol based on identity is very weightless and more efficient and also weightless protocol for client side.

C. Public Auditing with Complete Data Dynamics Support

Verification of data integrity at unreliable servers is the major concern in cloud storage. Proposed scheme first focused to discover the potential security threats and difficulties of preceding works and build a refined verification scheme Public auditing system with protocol that supports complete dynamic data operations is presented [3]. To accomplish dynamic data support, the existent proofread of PDP or PoR scheme is improved by spoofing the basic Markle Hash Tree (MHT). Proposed system extended in the direction of allowing TPA to perform many auditing jobs by examining the bilinear aggregate signature technique.

D. Efficient Third Party Auditing (TPA)

Cloud consumers save data in cloud server so that security as well as data storage correctness is primary concern. A novel and homogeneous structure is introduced [4] to pro- vide security to different cloud types. To achieve data storage security, BLS (Boneh–Lynn–Shacham) algorithm is used to signing the data blocks before outsourcing data into cloud. BLS (Boneh–Lynn–Shacham) algorithm is efficient and safer than the former algorithms. Batch auditing is achieved by using bilinear aggregate signature technique simultaneously. Reed-Solomon technique is used for error correction and to ensure data storage correctness. Multiple batch auditing is an important feature of this proposed work. It allows TPA to perform multiple auditing tasks for different users at the same.

E. Way of Dynamically Storing Data in Cloud

Securely preserving all data in cloud is not an easy job when there is demand in numerous applications for clients in cloud. Data storage in cloud may not be completely trustable because the clients did not have local copy of data stored in cloud. To address these issues, proposed a new protocol system using the data reading protocol algorithm to check the data integrity [5]. Service providers help the clients to check the data security by using the proposed effective automatic data reading algorithm. To recover data in future, also presented a multi server data comparison algorithm with overall data calculation in each update before outsourcing it to server's remote access point.

F. Effective and Secure Storage Protocol

Current trend is users outsourcing data into service provider who have enough area for storage with lower storage cost. A secure and efficient storage protocol is proposed that guarantees the data storage confidentiality and integrity [6]. This protocol is invented by using the construction of Elliptic curve cryptography and Sobol Sequence is used to confirm the data integrity arbitrarily. Cloud Server challenges a random set of blocks that generates probabilistic proof of integrity. Challenge-Response protocol is credential so that it will not exposes the contents of data to outsiders. Data dynamic operations are also used to keep the same security assurance and also provide relief to users from the difficulty of data leakage and corruptions problems.

G. Storage Security of Data

Resources are being shared across internet in public surroundings that creates severe troubles to data security in cloud. Transmitting data over internet is dangerous due to the intruder attack. So data encryption plays an important role in Cloud environment. Introduced a consistent and novel structure for providing security to cloud types and implemented a secure cross platform [7]. The proposed method includes some essential security services that are supplied to cloud system. A network framework is created which consists of three data backups for data recovery. These backups located in remote location from main server. This method used SHA Hash algorithm for encryption, GZIP algorithm for compression and SFSPL algorithm for splitting files. Thus, a secure cross platform is proposed for cloud computing.

H. Secure and Dependable Storage Services

Storage service of cloud permits consumers to place data in cloud as well as allowed to utilize the available well qualified applications with no worry about data storage maintenance. Although cloud provides benefits, such a service gives up the self-control of user's data that introduced fresh vulnerability hazards to cloud data correctness. To handle the novel security issue, accomplish the cloud data integrity and availability assurances, a pliable mechanism is proposed for auditing integrity in a dispersed manner [8]. Proposed mechanism allows users to auditing the cloud data storage and this auditing result utilized Homomorphic token with Reed-Solomon erasure correcting code technique that guarantee the correctness insurance and also identifying misconduct servers rapidly. The proposed design is extended to support block-level data dynamic operations. If cloud consumer is not able to possess information, time and utility then the users can assign their job to an evaluator i.e. TPA for auditing process in safe manner.

I. Optimal Cloud Storage Systems

Cloud data storage which requires no effort is acquiring more popularity for individual, enterprise and institutions data backup and synchronization. A taxonomic approach to attain storage service optimality with resource provider, consumer's lifecycle is presented [9]. Proposed scheme contributes storage system definition, storage optimality, ontology for storage service and controller architecture for storage which is conscious of optimality. When compared with existing work, more general architecture is created that works as a pattern for storage controller.

J. Process of access and Store Small Files with Storage

To support internet services extensively, Hadoop distributed file system (HDFS) is acquired. Several reasons are examined for small file trouble of native Hadoop distributed file system: Burden on Name Node of Hadoop distributed file system is enforced by large amount of small files, for data placement correlations are not considered, perfecting mechanism is not also presented. In order to overcome these small size problems, proposed an approach that improves the small files efficiency on Hadoop distributed file system [10]. Hadoop distributed file system is an Internet file system representative, which functioning on clusters. The cut-off point is measured in Hadoop distributed file system's circumstance in an experimental way, which helps to improve I/O performance. From taxonomic way, files are categorized as independent files, structurally and logically-related files. Finally perfecting technique is used to make better access efficiency and considering correlations when files are stored.

K. File Storage Security Maintenance

To assure the security of stored data in cloud, presented a system which utilizes distributed scheme [11]. Proposed system consists of a master server and a set of slave servers.

There is no direct communication link between clients and slave servers in the proposed model. Master server is responsible to process the client's requests and at slave server chunking operation is carried out to store copies of files in order to provide data backup for file recovery in future. Users can also perform effective and dynamic data operations. Clients file is stored in the form of tokens on main server and files were chunked on slave servers for file recovery. Thus proposed scheme achieved storage correctness insurance and data availability by using Token generation algorithm with homomorphic token and merging algorithm were used.

L. File Assured Deletion (FADE) for Secure Storage

Proposed a file assured deletion scheme based on policy to dependably efface files of cancelled file access policies [14]. Working prototype of FADE is implemented at the top of Amazon S3. Performance overhead is also evaluated on Amazon S3.

a. . File Assured Deletion Based on Policy

Data file is logically connected with file access policy and a data key. Each file access policy should be attached with control key. Maintenance of control key is the responsibility of key manager. When a policy is cancelled, control key of that policy will be dispatched from the key man- ager. The main idea is as follows: each file with data key is saved and control key is used to protect data key. Here key manager is responsible for retaining keys. The control key is deleted when a policy is cancelled. So that the encrypted file and data key could not be regained. In case the file is removed still a copy exists, that file is encrypted and unavailable to everyone. Multiple policies such as conjunctive and disjunctive policies are also presented. Conjunctive policies are used to recover file by satisfying all policies whereas disjunctive policies satisfying only one policy. Conclusion is FADE is executable in practice and this approach includes all dynamic data operations. Cryptographic operations are less and meta-data over- head is small.

M. Accessing Outsourced Data Efficiently

An approach is proposed to attain flexible access control and dynamic large-scale data in a safe and effective way. An Owner-write-user-read scenario is presented for accessing data [12]. Original data owner be only able to update/modify their data. Cloud users will be able to read information with corresponding access rights. Proposed approach deals with key generation, dynamics handling and overhead analysis. In key generation part, a key derivation hierarchy is generated and Storage over- head is moderated. Dynamics handling part consists of dynamic data operations and access rights of user. Eavesdropping can be overcome by over-encryption and lazy revocation.

IV. CONCLUSION

Cloud Computing is an emerging computing paradigm, allows users to share resources and information from a pool of distributed computing as a service over Internet. Even though Cloud provides benefits to users, security and privacy of stored data in cloud are still major issues in cloud storage. Cloud storage is much more beneficial and advantageous than the earlier traditional storage systems especially in scalability, cost reduction, portability and functionality requirements. This paper presented a survey on secure storage techniques in Cloud Computing. First several storage techniques that provide security to data in cloud have been discussed in detail and also highlighted the necessity for future research on storage methods to provider much better security and accountability. Finally, presented a comparative analysis on storage techniques, that includes the proposed approach, advantages and limitations of those storage techniques.

REFERENCES

[1] C. Wang, Q. Wang, K. Ren, and W. Lou, “Privacy-Preserving Public AuditingforStorage Security inCloud Computing,”Proc. IEEEINFOCOM ’10,Mar.2010.

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

[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,“Above theClouds:ABerkeley ViewofCloud Comput- ing,” Technical Report UCB-EECS-2009-28,Univ. of California, Berkeley, Feb.2009.

[4] P. Mell and T.Grance, “Draft NIST Working Definition of Cloud Computing,” http://csrc.nist.gov /groups/SNS/cloud-computing/index.html,June 2009.

[5] M. Armbrust, A.Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski,G.Lee,D.A. Patterson,A.Rabkin, I.Stoica, and M. Zaharia,“Above theClouds:ABerkeley ViewofCloud Computing,” Technical Report UCB-EECS-2009-28,Univ. OfCalifornia, Berkeley, Feb.2009.

[6] DongB,ZhengQetal.(2012). Anoptimized approach forstoring and accessingsm all files oncloudstorage, JournalofNetworkand Computer Applications,35(6),1847–1862.

[7] DeshmukhPM,GughaneASetal.(2012).Maintaining File Storage Security in Cloud Computing, International JournalofEmergingTechnologyand Advanced Engineering, vol 2(10), 2250–2459.

[8] TangY, Lee P PC et al (2010). FADE: a secure overlay cloud storage system with File Assured Deletion,6thInternationalICST Conference,Secure Comm.

[9] WangW, Li Z et al. (2009). Secure and Efficient AccesstoOutsourcedData,CCSW ‘09Proceedingsof the2009ACMworkshoponCloudcomputingsecurity, 55–66.

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

[11] M. A. Shah, R. Swaminathan, and M. Baker, “Privacy-preserving audit and extraction of digital contents,” Cryptology ePrint Archive, Report 2008/186, 2008, http://eprint.iacr.org/.

[12] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, “Scalable and efficient provable data possession,” in Proc. of SecureComm’08, 2008, pp. 1–10.

[13] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, “Enabling public verifiability and data dynamics for storage security in cloud computing,” in Proc. of ESORICS’09, volume 5789 of LNCS. SpringerVerlag, Sep. 2009, pp. 355–370.

[14] C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, “Dynamic provable data possession,” in Proc. of CCS’09, 2009, pp. 213–222.