Active Inspection Services with Log Maintenance Scheme for Cloud Storage

Karthikkeyan. S[1]
[1]Department of Computer Science & Engineering
[1]Saveetha School of Engineering College, Chennai

Abstract— In present days, the up-and-coming cloud-computing model is rapidly in advance force as an unconventional to traditional information technology. Cloud computing make available a scalability environment for emergent amounts of data and processes that work on a variety of services and applications by means of on-demand self-services. One fundamental aspect of this model shifting is that data are being centralized and outsourced into clouds. This category of outsourced storage space services in clouds have turn out to be a new profit growth point by given that a comparably low-cost, scalable, location-independent policy for managing clients' data. The cloud storage service (CSS) mitigates the load of maintenance and storage management. However, if such a significant service is weak to attacks or failures, it would take permanent losses to users since their data or records are stored into an unsure storage space pool outside the enterprises. These security risks move about in the direction of from the following reasons: the cloud infrastructures are much more authoritative and reliable than personal computing devices. If they are still susceptible to security threats both from inside and outside the cloud for the benefits of their control, there exist various motivations for cloud service providers (CSP) to behave falsely toward the cloud users in addition, the dispute infrequently suffers from the lack of trust on cloud service provider. As a result, their behaviors may not be known by the cloud users. Therefore, it is necessary for cloud service providers to offer a scalable audit service to check the integrity and accessibility of the stored data. While Cloud Computing makes these advantages more appealing than ever, it also brings new challenging security threats towards users' outsourced data. Since cloud service provider is separate administrative units, data outsourcing is actually resigning user's control over the destiny of their data. The correctness of the data in the cloud is being put at risk due to the subsequent reasons. First of all, although the infrastructures beneath the cloud are much more powerful and reliable than private computing devices, they are silent facing the broad range of both internal and external threats for data integrity.

Key words: Storage security, provable data possession, audit service, cloud storage

I. INTRODUCTION

Cloud computing provides a scalable environment for growing amounts of data and processes that work on various applications and services by means of on-demand self-services. Especially, the outsourced storage in cloud shas become a new profit growth point by providing a comparably low-cost, scalable, location-independent platform for managing clients' data. The cloud storage service (CSS) relieves the burden for storage management and maintenance. However, if such an important service is vulnerable to attacks or failures, it would bring irretrievable losses to the clients because their data or archives are stored in an uncertain storage pool outside the enterprises. These security risks come from the following reasons: First, the cloud infrastructures are much more powerful and reliable than personal computing devices, but they are still susceptible to internal threats (e.g., via virtual machine)and external threats (e.g., via system holes) that can damage data integrity [1]; second, for the benefits of possession, there exist various motivations for cloud service providers(CSP) to behave unfaithfully toward the cloud users [2];furthermore, disputes occasionally suffer from the lack of trust on CSP because the data change may not be timely known by the cloud users, even if these disputes may result from the users' own improper operations [3]. Therefore, it is necessary for CSP to offer an efficient audit service to check the integrity and availability of stored data [4].Security audit is an important solution enabling tracebackand analysis of any activities including dataaccesses, security breaches, application activities, and soon. Data security tracking is crucial for all organizations that should comply with a wide range of federal regulations including the Sarbanes-Oxley Act, Basel II,HIPAA, and so on.1 Furthermore, compared to the common audit, the audit services for cloud storages should provide clients with a more efficient proof for verifying the integrity of stored data. Unfortunately, the traditional cryptographic technologies, based on hash functions and signature schemes, cannot support for data integrity verification without a local copy of data. In addition, it is evidently impractical for audit services to download the whole data for checking data validation due to the communication cost, especially for large-size files. Therefore, following security and performance objectives should be addressed to achieve an efficient audit for outsourced storage in clouds:

- Public auditability: To allow a third party auditor(TPA) or clients with the help of TPA to verify the correctness of cloud data on demand without retrieving a copy of the whole data or introducing additional online burden to cloud services;
- Dynamic operations. To ensure there is no attack to compromise the security of verification protocol or cryptosystem by using dynamic data operations;
- Timely detection. To detect data errors or losses in outsourced storage, as well as anomalous behaviors of data operations in a timely manner;
- Effective forensic. To allow TPA to exercise strict audit and supervision for outsourced data, andoffer efficient evidences for anomalies; and
- Lightweight. To allow TPA to perform audit tasks with the minimum storage, lower communication cost, and less computation overhead. *Usage* data represent a Web site's usage, such as visitors' IP address, time and date of access, complete path (files or directories) accessed, referrers' address,

and other attributes that can be included in a Web access log.

In this paper, we introduce a dynamic audit service for integrity verification of untrusted and outsourced storages. Constructed on interactive proof system (IPS) with the zero knowledge property, our audit service can provide public audit ability without downloading raw data and protect privacy of the data. Also, our audit system can support dynamic data operations and timely anomaly detection with the help of several effective techniques, such asfragment structure, random sampling, and index-hash table(IHT). We also propose an efficient approach based on probabilistic query and periodic verification for improving the performance of audit services. A proof-of-concept prototype is also implemented to evaluate the feasibility and viability of our proposed approaches. Our experimental results not only validate the effectiveness of our approaches, but also show that our system does not create any significant computation and require less extra storage for integrity verification.

We list the features of our scheme in Table 1. We also make a comparison of related techniques, involving provable data possession (PDP) [5], scalable PDP (SPDP) [6], dynamic PDP (DPDP) [7], and compact proofs of irretrievability (CPOR) [8]. It clearly shows that our scheme not only supports complete privacy protection and dynamic data operations, but also enables significant savings in computation and communication costs, as well as a high detection probability of disrupted blocks.

The audit service has been done to trace out the data for security with a log based maintenance strategy. A maintenance log is a document that records who did what, when, and why. Maintenance of logs are extremely useful for troubleshooting recurring or obscure problems, as they provide a record of all work performed on the system and may shed light on hard-to-spot interactions between seemingly unrelated symptoms. The concept of tag creation for outsourced file, periodic sampling audit and dynamic data operation and audit has been done by generating efficient algorithm. The problem occurred in the existing system has been overcome in this system by auditing the cloud storage with an accurate log maintenance. The methodologies such as tag creation for outsourced file, periodic sampling audit and dynamic data operation and audit has been done to trace out the data for security. Tag is a visual representation of the text data that depicts keyword tags on websites. A secret key has been generated with the tag so as to represent the file to be stored in the cloud. The third party auditor is the one who verifies the genialness of the data stored by the privileged person and it also identifies the illegal changes occurred using the technique called random sampling method. A hash function is any algorithm that maps data of variable length to data of a fixed length. Each page constitutes of a hash function and a block of pages is called as a file. Dynamic data operation audit fastens the operation of update/delete/insert. Modification made in a single page does not resave the entire file and saves the modified page alone. But this system does not maintain any log so as to identify the users who accessed the data/modified the data/time period of accessing and so on.

The remainder of this paper is organized as follows:Active inspection services with log maintenance scheme for cloud storage. Section 2, reviews the literature in cloud storage. In section 3, the proposal's model is presented. Section 4,discusses the evaluation of proposed model Section 5 draws proposal's conclusion and future work.

II. Related Work

This section briefly delivers literature on Cloud Storage, andLog Maintenance.

Traditional cryptographic technologies for data integrityand availability, based on hash functions and signature schemes,cannot work on the outsourced datawithout a local copy of data. In addition, it is not a pacticalsolution for data validation by downloading them due tothe expensive communications, especially for large-size files. Moreover, the ability to audit the correctness of data in a cloud environment can be formidable and expensive for cloud users. Therefore, it is crucial to realize public audit ability for CSS, so that data owners (DOs) may resortto a TPA, who has expertise and capabilities that a common user does not have, for periodically auditing the out sourced data. This audit service is significantly important for digital forensics and data assurance in clouds

A. Cloud Storage

Cloud computing is an emerging, on-demand and internet-based technology. It provides variety of services over internet such as, software, hardware, data storage and infrastructure. This technology has been used by worldwide customers to improve their business performance. However , to utilize these services by authorized customer, it is necessary to have strict authentication check. At present, authentication is done in several ways: such as , textual, graphical, bio-metric, 3D password and third party authentication.

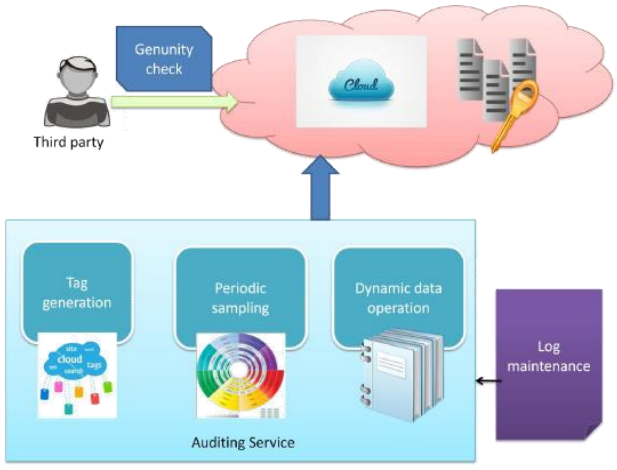

Fig. 1: Architecture of log maintenance scheme

B. Log Maintanance

Traditional cryptographic technologies for data integrity and accessibility, based hash functions and on signature schemes cannot work on the outsourced data lacking a local copy of data. In accumulation, it is not a realistic solution for data validation by downloading them due to the exclusive transaction, especially for large-size files. Moreover, the solutions to audit the correctness of the data in a cloud environment can be formidable and expensive for the cloud

users. Therefore, it is critical to recognize public audit ability for Cloud Storage Service, so that data owners may remedy to a third party auditor (TPA), who has proficiency and capabilities that a common user does not have, for from time to time auditing the outsourced data. This audit service is extensively significant for digital forensics and data assurance in clouds. who has proficiency and capabilities that a common user does not have, for from time to time auditing the outsourced data. This audit service is extensively significant for digital forensics and data assurance in clouds.

III. ARCHITECTURE

A. Tag Generation

Tag is a visual representation of the text data that depicts keyword tags on websites. The data's are stored in the form of blocks and each blocks constitute of set of files. Each file in the blocks are represented by an individual specification by tagging it. In this module, the tags for every files in blocks has been generated with a tag generation algorithmic technique.The client (DO) uses a secret key to preprocess a file, which consists of a collection of nblocks, generates a set of public verification parameters (PVPs) and IHT that are stored in TPA, transmits the file and some verification tags to CSP, and may delete its local copy.Fig.2.a. To maximize the storage efficiency and audit performance, our audit system introduces general fragment structure for outsourced storages. An instance for this framework, which is used in our approach is shown in Anoutsourced file F is split into n blocks , and each block mi is split into s sectors . The fragment framework consists of n block-tag pair where is a signature tag of a block mi generated by some secrets. We can use such tags and corresponding data to construct a response in terms of the TPA's challenges in the verification protocol, such that this response can be verified without raw data. If a tag is unforgeable by anyone except the original signer, we call it a secure tag. These schemes, built from collision resistance hash functions (see Section 5) and a random oracle model, support the features of scalability, performance, and security.

B. Periodic Sampling

By using an interactive proof protocol of retrievability, TPA (or other applications) issues a "random sampling" challenge to audit the integrity and availability of the outsourced data in terms of verification information (involving PVP and IHT) stored in TPA Fig. 2.b; and Maintaining security is an ongoing process that must be reviewed and revisited periodically. Maintaining a secure system requires vigilance, because the default security configuration for any system tends to become increasingly open over time. In this module, the audits are randomly sampled so as to monitor the data to preserve security.In contrast with "whole" checking, random "sampling" checking greatly reduces the workload of audit services, while still achieves an effective detection of mis behaviours. Thus, a probabilistic audit on sampling checking is preferable to realize the anomaly detection in a timely manner, as well as to rationally allocate resources. The fragment structure shown in Fig. 3 provides probabilistic audit as well: Given a randomly chosen challenge (or query)Q ¼ fði; viÞgi2I , where I is a subset of the block indices and vi is a random coefficient, an efficient algorithm is used to produce a constant-size response ð_1; _2; _ _ _ ; _s; _0Þ, where_i comes from all fmk;i; vkgk2I and _0 is from all f_k; vkgk2I .Generally, this algorithm relies on homomorphic propertiesto aggregate data and tags into a constant-size response, which minimizes network communication costs. Since the single sampling checking may overlook a small number of data abnormality, we propose a periodic sampling approach to audit outsourced data, which is named as Periodic Sampling Audit. With this approach, the audit activities are efficiently scheduled in an audit period, and a TPA merely needs to access small portions of files to perform audit in each activity. Therefore, this method can detect exceptions periodically, and reduce the sampling numbers in each audit.

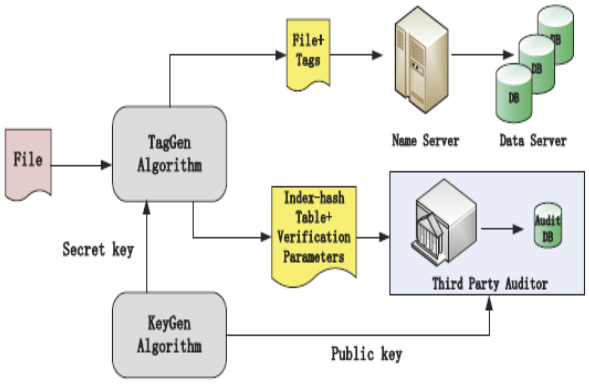

Fig. 2: (a) Tag Generation for Outsourced File

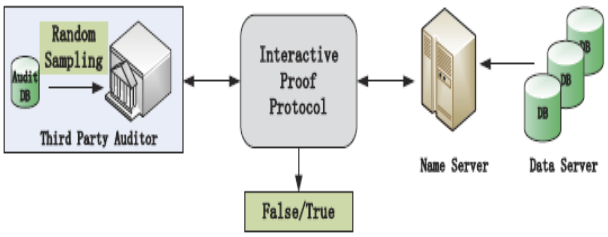

Fig. 2: (b) Periodic Sampling Audit

C. Dynamic Data Operations

An AA, who holds a DO's secret key, can manipulate the out sourced data and update the associated IHT stored in TPA. The privacy of secret key and the checking algorithm ensure that the storage server cannot cheat the AAs and forge the valid audit records.dynamic usually means capable of action and/or change. Fig 3.c The occurrence of updating/modification in any one page of a file does not consume the processing time by saving the entire file in this system whereas it saves only the updated page of the file alone.To support dynamic data operations, we introduce a simple IHT to record the changes of file blocks, as wellas generate the hash value of each block in the verification process. The structure of our IHT is similar to that of file lock allocation table in file systems. Generally, the IHT _consists of serial number, block number, version number, and random integer (see Table 2 in Section 5). Note that we must assure all records in the IHT differ from one another to prevent the forgery of data blocks and tags. In addition to recording data changes, each record _i in the table issued to generate a unique hash value, which in turn issued for the construction of a signature tag _i by the secret key sk. The relationship between _i and _i must be cryptographically

secure, and we make use of it to design our verification protocol. Although the IHT may increase the complexity of an audit system, it provides a higher assurance to monitor the behavior of an untrusted CSP, as well as valuable evidence for computer forensics, due to the reason that anyone cannot forge the valid _i (in TPA) and _i (in CSP) without the secret key sk. In practical applications, this architecture can be constructed into a virtualization infrastructure of cloud-based storage service [14]. In Fig. 4, we show an example of Hadoop distributed file system (HDFS),2 which is a distributed, scalable, and portable file system [15]. HDFS'

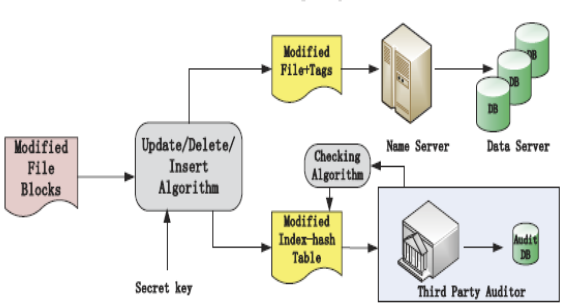

Fig. 2: (c) Dynamic Data Operations and Audit

IV. Experimental Evaluation

In market research, log maintenancecustomers with different preference value different featuresof the product or service.

In this section, we describe the construction of algorithms inour audit architecture. First, we present the definitions forthe tag generation process as follows:

KeyGen (1_) takes a security parameter _ as an input,and returns a public/secret keypair ðpk; skÞ; and

T agGen (sk; F) takes a secret key sk and a file F,

and returns a triple ð_; ; _Þ, where _ denotes thesecret used to generate verification tags, is a set ofPVPs u and IHT _, i.e., ¼ ðu; _Þ, and _ denotes aset of tags.

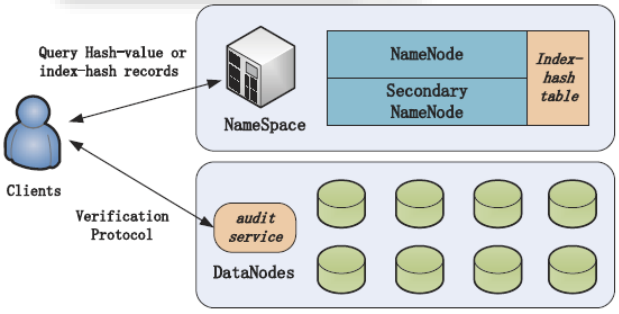

Fig. 3: Index Table Verification

A. Supporting Periodic Sampling Audit

At any time, TPA can check the integrity of a file F as follows: TPA first queries database to obtain the verification information and initializes an interactive protocolProofðCSP;ClientÞ; then, it performs a 3-move proofprotocol: Commitment, Challenge, and Response; and it finally verifies the interactive data to get the results. In fact, because our scheme is a publicly verifiable protocol, anyone can run this protocol, but s/he is unable to get any advantage to break the cryptosystem, even if TPA and CSP cooperate for an attack. Let PðxÞ denotes the subject Pholds the secret x and hP;V iðxÞ denotes both parties P and V share a common data x in a protocol. This process cans be defined as follows:

B. Supporting Dynamic Data Operations

To meet the requirements from dynamic scenarios, we introduce following definitions for dynamic data operations:

Update(sk; _i;m0i) is an algorithm run by AA to update the block of a file m0i at the index i by using sk, and it returns a new verification metadata ð_0i; _0i;m0iÞ;

Delete(sk; _i;mi) is an algorithm run by AA to delete the blockmi of a filemi at the index i by using sk, and it returns a new verification metadata ð_0 i; _i; _0iÞ; and

Insert(sk; _i;mi) is an algorithm run by AA to insert the block of a file mi at the index i by using sk, and it returns a new verification metadata ð_0i; _0i;m0iÞ.

To ensure the security, dynamic data operations are available only to DOs or AAs, who hold the secret key sk. Here, all operations are based on data blocks. Moreover, to implement audit services, applications need to update the IHTs. It is necessary for TPA and CSP to check the validity of updated data. In Fig. 5c, we describe the process of dynamic data operations and audit. First, an AA obtains the Public verification information from TPA. Second, the application invokes the Update, Delete, and Insert algorithms, and then sends the new 0 and _0 to TPA and CSP, respectively. Next, the CSP makes use of an algorithm Check to verify the validity of updated data. Note that theCheck algorithm is important to ensure the effectiveness ofthe audit. Finally, TPA modifies audit records after the confirmation message from CSP is received and thecompleteness of records is checked.

C. Work Flow Of Audit System

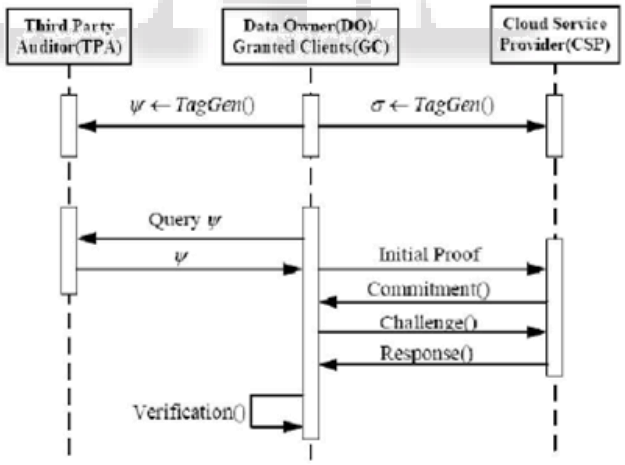

Fig 4.1: Tag Generation and User Verification

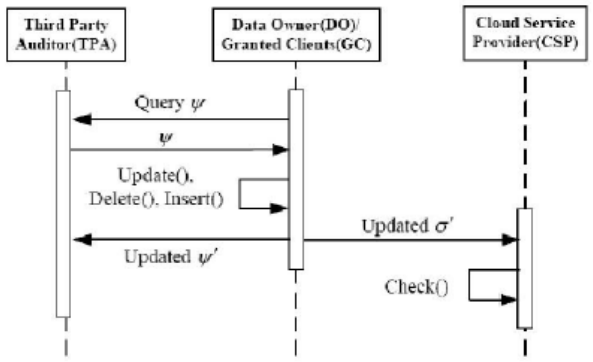

Fig. 4.3: Dynamic Data Operations And Dynamic

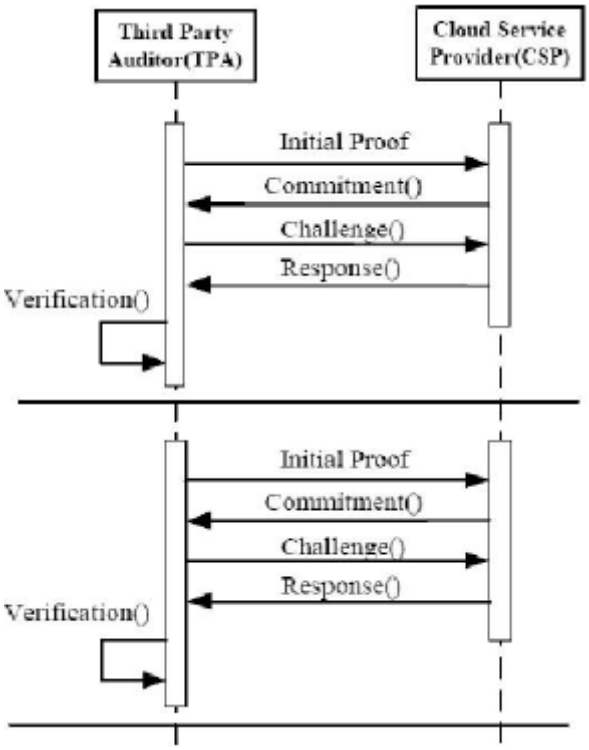

Fig. 4.2: Periodic Sampling Audit

V. Conclusion And Future Work

In this paper, Cloud Computing releases the world of computing to a wider range of uses and increases the ease of usage by giving access through any kind of internet connection. Though with these increased ease of usage also come drawbacks. Privacy security is a key issue for cloud storage and is to be considered very important. To ensure that the risks of privacy have been mitigated a variety of techniques that may be used in order to achieve privacy. This paper has addressed some privacy approaches for overcoming the issues in privacy on un trusted data stores in cloud computing using the audit services strategies with log maintenance. Categories the methodologies in the literature as encryption based methods, access control based mechanisms, query integrity/keyword search schemes, and audit ability schemes. The work is giving an efficient privacy preserving storage compared to other works. Even though there are many approaches in the literature for mitigating the concerns in privacy no approach is fully sophisticated to give a privacy-preserving storage that overcomes all the other privacy concerns. Thus to deal with the concerns of privacy, we need to develop privacy–preserving framework that overcomes the worries in privacy security and encourage users to adopt cloud storage services more confidently. Cloud is a vast area to provide security and to generate traceability by means of audit services. In future, the privacy leakage audit in terms of entity relational datasets access can also be considered to monitor.

References

[1] P. Ammann and S. Jajodia, "Distributed Timestamp Generation inPlanar Lattice Networks," ACM Trans. Computer Systems, vol. 11,pp. 205-225, Aug. 1993.

[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," Proc. ACM Conf. Computer and Comm. Security, pp. 598- 609, 2007.

[3] E. Barka and A. Lakas, "Integrating Usage Control with SIP-Based Communications," J. Computer Systems, Networks, and Comm., vol. 2008, pp. 1-8, 2008.

[4] D. Boneh and M.K. Franklin, "Identity-Based Encryption from the Weil Pairing," Proc. Int'l Cryptology Conf. Advances in Cryptology, pp. 213-229, 2001.

[5] R. Bose and J. Frew, "Lineage Retrieval for Scientific Data Processing: A Survey," ACM Computing Surveys, vol. 37, pp. 1- 28, Mar. 2005.

[6] P. Buneman, A. Chapman, and J. Cheney, "Provenance Management in Curated Databases," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD '06), pp. 539-550, 2006.

[7] B. Chun and A.C. Bavier, "Decentralized Trust Management and Accountability in Federated Systems," Proc. Ann. Hawaii Int'l Conf. System Sciences (HICSS), 2008.

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

[9] H.-C. Hsiao, Y.-H. Lin, A. Studer, C. Studer, K.-H. Wang, H.Kikuchi, A. Perrig, H.-M. Sun, and B.-Y. Yang, "A Study of User-Friendly Hash Comparison Schemes," Proc. Ann. ComputerSecurity Applications Conf. (ACSAC), pp. 105-114, 2009.

[10] A.R. Yumerefendi and J.S. Chase, "Strong Accountability forNetwork Storage," Proc. Sixth USENIX Conf. File and StorageTechnologies (FAST), pp. 77-92, 2007.

[11] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S. Yau, "EfficientProvable Data Possession for Hybrid Clouds," Proc. 17th ACMConf. Computer and Comm. Security, pp. 756-758, 2010.

[12] M. Xie, H. Wang, J. Yin, and X. Meng, "Integrity Auditing ofOutsourced Data," Proc. 33rd Int'l Conf. Very Large Databases(VLDB), pp. 782-793, 2007.

[13] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-PreservingPublic Auditing for Data Storage Security in Cloud Computing,"Proc. IEEE INFOCOM, pp. 1-9, 2010.

[14] B. Sotomayor, R.S. Montero, I.M. Llorente, and I.T. Foster, VirtualInfrastructure Management in private and Hybrid Clouds," IEEEInternet omputing, vol. 13, no. 5, pp. 14-22, Sept./Oct. 2009

