Indian Journal of Science and Technology

A Survey on Secure Storage in Cloud Computing

A. Rajathi[1*] and N. Saravanan[2]

[1] PG Student, School of Computing , SASTRA University, 613401, Thanjavur, Tamilnadu, India; rajichakkaravarthii@gmail.com

[2] Assistant Prof., School of Computing , SASTRA University, 613401, Thanjavur, Tamilnadu, India; saranindia@gmail.com

Abstract

Cloud Computing is an environment for providing information and resources that are delivered as a service to end-users over the Internet on demand. Thus cloud enables users to access their data from any geographical locations at any time and also has brought benefits in the form of online storage services. Cloud storage service avoids the cost expensive on software, personnel maintenance and provides better performance, less storage cost and scalability. But the maintenance of stored data in a secure manner is not an easy task in cloud environment and especially that stored data may not be completely trustworthy. Cloud delivers services through internet which increases their exposure to storage security vulnerabilities. However security is one of the major drawbacks that preventing several large organizations to enter into cloud computing environment. This work surveyed on several existing cloud storage frameworks, techniques and their advantages, drawbacks and also discusses the challenges that are required to implement secure cloud data storage. This survey results help to identify the future research areas and methods for improving the existing drawbacks.

Keywords: Cloud Computing, Data, Security, Storage Techniques and Survey.

1. Introduction

Cloud Computing is a kind of computing whereby shared resources and IT-related capabilities are provided as a service to outer customers using Internet techniques. Cloud Computing depends on sharing information and computing resources instead of using local servers or personal devices for to manage supplications. 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. 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.

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's data into cloud data storage, security plays a vital role. Sometimes customers store some sensitive information 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.

* *Corresponding author:*
A. Rajathi (rajichakkaravarthii@gmail.com)

In preceding works, several methods are proposed 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.

2. Storage Techniques in Cloud Computing

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

2.1 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 [1] 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.

2.2 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.

2.3 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.

2.4 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 provide 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.

2.5 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.

2.6 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.

2.7 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.

2.8 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.

2.9 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. A new prototype NubiSave is also proposed which is available freely and it implements almost all of RAOC concepts.

2.10 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 NameNode of Hadoop distributed file system is enforced by large amount of small files, for data placement correlations are not considered, prefetching 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 prefetching technique is used to make better access efficiency and considering correlations when files are stored.

2.11 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.

Table 1. Comparative analysis on advantage and limitations of existing storage techniques

Storage Scheme	Proposed Approach	Advantages	Restrictions
1. Implicit Storage Security to Online data	Data partitioning scheme for online data storage.	Partitioned data pieces cannot bring out any user information.	In case user forgot where the data stored, it will become difficult for users.
2. Identity-Based Authentication	New authentication protocol based on identity which is based on hierarchical model	Weightless and more expeditious.	Only certificate communication is taken into account.
3. Public Auditing with Complete Data Dynamics support	PKC-based homomorphic authenticator is used to outfit the verification protocol.	Basic Markle Hash Tree (MHT) is manipulated for block tag authentication.	Computation cost of BLS scheme is prominent.
4. Efficient Third Party Auditing (TPA)	Novel and uniform security structure. Storage security is accomplished by utilizing BLS algorithm.	Auditor performs auditing jobs for different users at the same.	Unable to support both public verification and dynamic data correctness.
5. Dynamic Storage way in Cloud Computing	New protocol system using the data reading protocol algorithm. Multi server data comparison algorithm to recover data.	Integrity can be verified before and after data insertion.	TPA is not considered for integrity checking process.
6. Effective and Secure Storage Protocol	Efficient and secure storage protocol is implemented by utilizing Elliptic curve cryptography and Sobol Sequence	Block level data dynamic operations are also used to maintain the same security assurance.	Elliptic Curve Cryptography scheme is only suitable for devices with restricted low power.
7. Storage Security of data	Uniform and modern structure of security for different cloud types. SHA Hash, GZIP algorithm and SFSPL algorithm.	Provided data backups for data recovery. Includes essential security services such as authentication, encryption and decryption and compression.	Data back ups are available at multiple servers. So there is a chance for servers to behave unreliably.
8. Secure and Dependable Storage Services	Homomorphic token with Reed-Solomon erasure correcting code.	Guaranteed the correctness insurance and also identified the immoral server behavior.	Gross overhead approximately stays equal with other.
9. Optimal Cloud Storage Systems	Taxonomic approach for achieving cloud storage service optimality. Proposed a new NubiSave prototype	Proposed generic architecture served as blueprint for optimal storage controller. NubiSave is available freely.	NubiSave is needs to integrate with frontends for future research.
10. Process of access and store small files with storage	Prefetching technique should be used to make better access efficiency.	Improves the access ability of small files. Cut-off point is measured to improve I/O performance	Formula for cut-off point not available. It will be identified in future.
11. File Storage Security Maintenance	Distributed scheme contains master server and a set of slave servers. Token generation algorithm and merging algorithm are used.	File chunking operation is carried out to provide data backup in case of server failure.	Data chunks are stored in slave server will lead to an opportunity of corrupting data by servers.
12. File Assured Deletion (FADE) for Secure Storage	Conjunctive and disjunctive policies are used for file recovering process.	Support for dynamic data operations and meta data overhead is less.	Time and Space are the major overhead of this scheme.
13. Accessing outsourced data efficiently	An Owner-write-user-read Scenario for accessing data.	Original data owner be only able to update/ modify their data.	Combination of multiple policies is not supported.

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.

2.12 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 [12]. Working prototype of FADE is implemented at the top of Amazon S3. Performance overhead is also evaluated on Amazon S3.

2.12.1. 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 manager. 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 overhead is small.

2.13 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 [13]. 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 overhead 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.

3. 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.

4. References

1. Parakh A, and Kak S (2009). Online data storage using implicit security, Information Sciences, vol 179(19), 3323–3331.
2. Li H, Dai Y et al. (2009), Identity-Based Authentication for Cloud Computing, M. G. Jaatun, G. Zhao, and C. Rong (Eds.): Cloud Computing, Lecture Notes in Computer Science, vol 5931, 157–166.
3. Wang Q, Wang C et al. (2011). Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing, IEEE Transactions on Parallel and Distributed Systems, vol 22(5), 847–859.
4. Balakrishnan S, Saranya G, et al. (2011). Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud, International Journal of Computer Science and Technology, vol 2(2), 397–400.

5. Dinesh C (2011). Data Integrity and Dynamic Storage Way in Cloud Computing, Distributed, Parallel, and Cluster Computing.
6. Kumar S P, Subramanian R (2011). An efficient and secure protocol for ensuring data storage security in Cloud Computing, International Journal of Computer Science Issues, vol 8(6), No 1, 261–274.
7. Sajithabanu S, Raj E G P (2011). Data Storage Security in Cloud, International Journal of Computer Science and Technology, vol 2(4), 436–440.
8. Wang C, Wang Q et al. (2012), Toward Secure and Dependable Storage Services in Cloud Computing, IEEE Transactions on Services Computing, vol 5(2), 220–232.
9. Spillner J, Müller J et al. (2012), Creating optimal cloud storage systems, Future Generation Computer Systems, vol 29(4), 1062–1072.
10. Dong B, Zheng Q et al. (2012). An optimized approach for storing and accessing small files on cloud storage, Journal of Network and Computer Applications, 35 (6), 1847–1862.
11. Deshmukh P M, Gughane A S et al. (2012). Maintaining File Storage Security in Cloud Computing, International Journal of Emerging Technology and Advanced Engineering, vol 2(10), 2250–2459.
12. Tang Y, Lee P P C et al (2010). FADE: a secure overlay cloud storage system with File Assured Deletion, 6th International ICST Conference, Secure Comm.
13. Wang W, Li Z et al. (2009). Secure and Efficient Access to Outsourced Data, CCSW '09 Proceedings of the 2009 ACM workshop on Cloud computing security, 55–66.