An Approach for Data Storage Security in Cloud Computing

Deepanchakaravarthi Purushothaman[1]and Dr.Sunitha Abburu[2]

[1] Master of Computer Application ,Adhiyamaan College of Engineering Hosur, Tamilnadu-635109, India.

[2] Professor and Director, Master of Computer Application,Adhiyamaan College of Engineering Hosur, Tamilnadu-635109, India.

Abstract

Cloud computing is now days emerging field because of its performance, high availability, low cost. In the cloud many services are provided to the client by cloud. Data store is main future that cloud service provides to the companies to store huge amount of storage capacity. But still many companies are not ready to implement cloud computing technology due to lack of proper security control policy and weakness in protection which lead to many challenge in cloud computing.. The main objectives of this paper are, 1) To prevent Data access from unauthorized access, it propose a distributed scheme to provide security of the data in cloud .This could be achieved by using homomorphism token with distributed verification of erasure-coded data. 2) Proposed scheme perfectly stores the data and identifies the any tamper at the cloud server.3) And also performs some of the tasks like data updating, deleting, appending. This paper also provides a process to avoid Collusion attacks of server modification by unauthorized users. The proposed techniques is been implementation and results are shown at the below.

Keyword: *cloud computing, Authentication, homomorphism token, Collusion attacks.*

1. Introduction

Cloud computing is the most demanding and emerging technology throughout the world. Cloud computing is an Internet based computer technology. Some of the major firms like Amazon, Microsoft and google have implemented the "CLOUD" and have been using it to speed up their business. Cloud computing has given a new dimension to the complete outsourcing arena (SaaS, PaaS and IaaS) and they provide ever cheaper powerful processor with these computing architecture The major thing that a computer does is to store in the available space and retrieve information whenever requested by the authenticated user. The pioneer of Cloud Computing vendor,(example) Amazon S3 is storage for the Internet. Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web. It also allows developer to access the highly scalable, reliable, secure, fast, inexpensive infrastructure that Amazon uses to run its own global network of web sites. From the viewpoint of data security, which has always been an important aspect of quality of service, Cloud Computing unavoidably poses new challenging security threats for number of reasons.

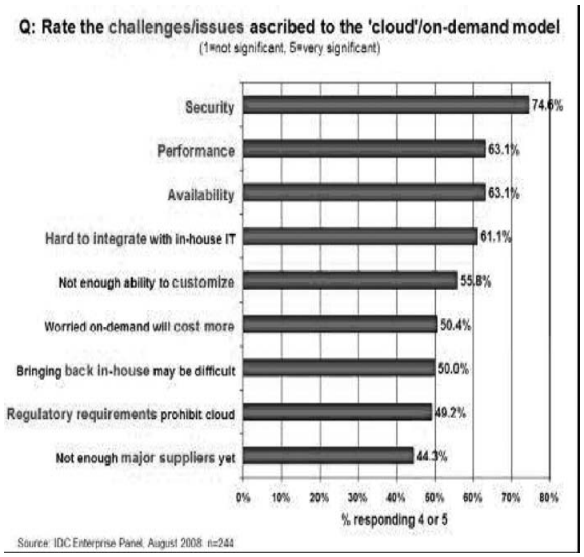

Fig.1. Results of IDC survey ranking security challenge[1]

- Unauthenticated person don't attack the authorized file
- Avoids Collusion attacks
- Malicious data modification attack

- Dynamic data operations
- Identification tamper server

2. Related Work

The Internet began to grow quickly in the 1990s and the increasingly sophisticated network infrastructure and increased bandwidth developed in recent years has dramatically enhanced the stability of various application services available to users through the Internet, thus marking the beginning of cloud computing network services. Many organizations tried to enhance for their security constraints, for their secure database, for their web application but they have not achieved a high-level security for their organizations. Data integrity quality of correctness, completeness, wholeness, soundness and compliance with the intention of the creator of the data. It is achieved by preventing accidental or deliberate but unauthorized insertion, modification or destruction of data in a database. Ensuring the integrity of the data really means that it changes only in response to authorized transactions.(see Fig.1) given stats confirm that the "Security" is the main Challenge in Cloud Computing For example IDC recently conducted a survey of 244 IT executives/CIOs and their line-of-business (LOB) colleagues to gauge their opinions and understand their companies' use of IT cloud services. Security ranked first as the greatest challenge or issue of cloud computing.

Shacham et al. [2] In a Compact proof-of-retrievability system, a data storage center must prove to a verifier that he is actually storing all of a client's data safely. The central challenge is to build systems that are both efficient and provably secure that is, it should be possible to extract the client's data from any prover that passes a verification.Juels and Kaliski [3] proposed a scheme called "provable data possession" (PDP) model for ensuring possession of file on untrusted storages. Their scheme utilized public key based homomorphic tags for auditing the data file, thus providing public verifiability. However, this scheme requires sufficient computation overhead that can be expensive for an entire file. Check.M.A Shah, M.Baker, J.C.Mogul;[4]proposed a scheme called "Auditing to Keep online Storage Services Honest" The need for auditing to support an online service-oriented economy. They highlight issues around both internal and external auditing. This paper [2][4][9],allows TPA to audit the cloud data storage without demanding user's time feasibity (or) resources. The proposed method provides public key verification for secured storage and investigate the problem of fine-grained data error Localization in the cloud.

3. Problem Statement

From the perspective of data security, which has always been an important aspect of quality of service, Cloud Computing inevitably poses new challenging security threats for number of reasons.

- Data stored on cloud servers is not completely secure from infection**.** While popular cloud services such as Google Docs are equipped with virus scanning software, there is still the possibility of an internal or external attack affecting your data.
- The data stored in the cloud may be frequently updated by the users, including insertion, deletion, modification, appending, reordering, etc. To ensure storage correctness under dynamic data update, distributed protocol is used.

A descriptive architecture for secure data storage is illustrated in(see Fig .2) Data storage in a cloud is a process where the owner stores his data, files and applications through a Cloud Storage Provider (CSP) into a set of cloud servers. At the time of file storage, security key is used to secure the file from unauthorized access and then safely stored in the cloud. User who likes to access the file from cloud needs the security key to retrieve the file. User sends a key request to the owner and retrieves the file from the cloud after security key sent by the owner. File can't be accessed by any Unauthorized person or person who entering unmatching security key. For additional security, blocking IP address of the system those who illegally trying to access the file.

.

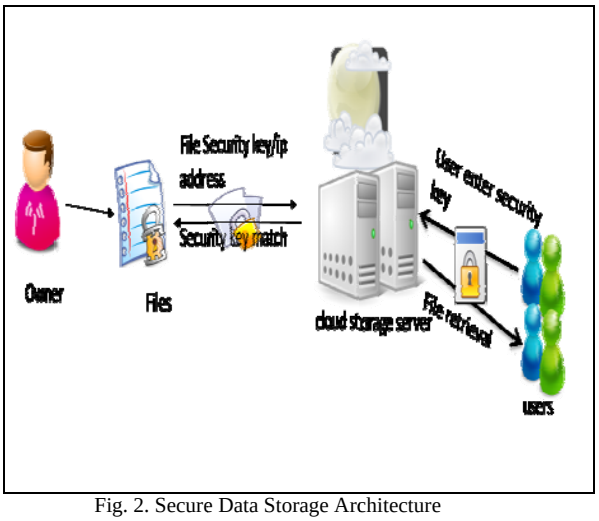

Fig. 2. Secure Data Storage Architecture

Data storage in a cloud is a process where the owner stores his data

This paper proposes using homomorphic token & verification of erasure-coded the current research provides cloud data security along with minimizes the redundancy.

- The distributed protocol in our work future provides the localization of data error. Which only provides binary results about the storage state across the distributed service in predecessors.
- Operations like Update, delete and integrity are also provided in the proposal methods.
- Extensive security and performance analysis shows that the proposed scheme is highly efficient and resilient against Byzantine failure, malicious data modification attack, and even server Collusion attacks.

4. Secure Data Storage in Cloud

In cloud storage system, companies stores their data in the remotely located data server. Accordingly, correctness of the data is assured. Even though sometimes unauthorized person may modify or delete the data which leads to server compromise and/or random Byzantine failures. Because it can be the first step for fast recovery of the storage errors. The cloud storage systems propose an effective and flexible distributed scheme with explicit dynamic data support for file distribution across cloud servers. By computing homomorphic token using universal hash function [7] which can be perfectly integrated with the verification of erasure-coded data. As well as it identifies misbehaving servers. Finally, the procedure for file retrieval and error recovery based on erasure-correcting code is outlined.

4.1Token correctness

It achieves assurance for data storage correctness and data error localization, using pre-computed token. Before sharing file distribution using pre-computes a certain number of shortest verification token are generated that will ensure security for a block of data in a file in cloud storage. When the user wants to make sure the storage correctness for the data in the cloud, he challenges the cloud servers with a set of randomly generated block indices. After getting assurance of the user it again asks for authentication by which the user is confirmed to be the authenticated user. Upon receiving assurance, each cloud server computes a short “signature” over the specified blocks and returns them to the user. The values of these signatures should match the corresponding tokens pre-computed by the user. All servers operate over the same subset of the indices, the requested response values for integrity check must also be a valid codeword determined by a secret matrix. Suppose the user wants to challenge the cloud server’s t times to make sure the correctness of data storage. Then, he must pre-compute t verification tokens for each function, a challenge key and a master key are used. To generate the ith token for server j, the user acts as follows the details of token Generations are shown in Algorithm 1.

- Derive an arbitrary value i and a permutation key based on master permutation key.
- Calculate the set of randomly-chosen index.
- Calculate the token using encoded file and the arbitrary value derived.

Algorithm 1 Token Pre-computation

Block of data is represented as l;

No. of .blocks is denoted as n;

Let f be the function and t be the token ;

Index per proof is denoted as r;

Generate M_k and C_k;

For point G (j); j->1, n execute

/*j server position*/

For round i->1, t execute

/*i block index*/

Derive i = f (i) and k (i) from master key.Compute v(j)

End for

End for

Store all the vis locally.

End procedures

4.2. Correctness Verification and Error Localization

Error localization is a key requirement for eradicate errors in storage systems. However, many previous schemes do not explicitly consider the problem of data error localization.

The challenges response protocol in our work future provides the localization of data error. Which only provides binary results about the storage state across the distributed service in predecessors. The response values from servers for each challenge not only determine the correctness of the distributed storage, but also contain information to locate potential data error(s).
Specifically, the procedure of the ith challenge-response for a cross-check over the n servers is described as follows:

- The client reveals the i as well as the ith key k (i) to each servers
- The server storing vector G aggregates those r rows
- Specified by index k(i) into a linear combination R
- Upon receiving R is from all the servers, the user takes away values in R.
- Then the user verifies whether the received values remain a valid codeword determined by secret matrix.

Because all the servers operate over the same subset of indices, the linear aggregation of these r specified rows (R (1)i , . . . ,R(n)i) has to be a codeword in the encoded file matrix. If the above equation holds, the challenge is passed. Otherwise, it indicates that among those specified rows, there exist file block corruptions. Once the inconsistency among the storage has been successfully detected, we can rely on the pre-computed verification tokens to further determine where the potential data error(s) lies in. Note that each response R(j) i is computed exactly in the same way as token v(j) i , thus the user can simply find which server is misbehaving by verifying.

5. Implementation

5.1 Secure Software Development Life Cycle

The Security Development Lifecycle (SDL) is a software development security assurance process consisting of security practices grouped by seven phases Investigation, Analysis, Logical design, Physical design, Implementation, Maintenance.

Phase1.Investigation**:**Define project processes and goals, and document them in the program security policy.
Phase2.Analysis: Analyze existing security policies and programs, analyze current threats and controls, examine legal issues, and perform risk analysis.

Phase3.Logical design: Develop a security blueprint, plan incident response actions, plan business responses to disaster, and determine the feasibility of continuing and/or outsourcing the project.

Phase4.Physical design: Select technologies to support the security blueprint, develop a definition of a successful solution, design physical security measures to support technological solutions, and review and approve plans.

Phase5.Implementation: Buy or develop security solutions. At the end of this phase, present a tested package to management for approval.

Phase6.Maintenance: Constantly monitor, test, modify, update, and repair to respond to changing threats.

5.2 Main Modules

5.2.1 Client Module

The client sends the query to the server. Based on the query the server sends the corresponding file to the client. Before this process, the client authorization step is involved. In the server side, it checks the client name and its password for security process. If it is satisfied and then received the queries form the client and search the corresponding files in the database. Finally, find that file and send to the client. If the server finds the intruder means, it set the alternative Path to that intruder. Using screen shown in fig.3.

5.2.2 System Module

- User

Users, who have data to be stored in the cloud and rely on the cloud for data computation, consist of both individual consumers and organizations.

- Cloud Service Provider (CSP)

A CSP, who has significant resources and expertise in building and managing distributed cloud storage servers, owns and operates live Cloud Computing systems.

- Third Party Auditor (TPA)

An optional TPA, who has expertise and capabilities that users may not have, is Trusted to assess and expose risk of cloud storage services on behalf of the users upon request.

5.2.3 Cloud Data Storage Module

Cloud data storage, a user stores his data through a CSP into a set of cloud servers, which are running in a simultaneous, the user interacts with the cloud servers via CSP to access or retrieve his data. In some cases, the user may need to perform block level operations on his data.users should be equipped with security means so that they can make continuous correctness assurance of their stored data even without the existence of local copies. In case that users do not necessarily have the time, feasibility or resources to monitor their data, they can delegate the tasks to an optional trusted TPA of their respective choices. In our model, we assume that the point-to-point communication channels between each cloud server and the user is authenticated and reliable, which can be achieved in practice with little overhead. Using screen shown in Fig.4.

5.2.4 Cloud Authentication Server

The Authentication Server (AS) functions as any AS would with a few additional behaviors added to the typical client-authentication protocol. The first addition is the sending of the client authentication information to the masquerading router. The AS in this model also functions as a ticketing authority, controlling permissions on the application network. The other optional function that should be supported by the AS is the updating of client lists, causing a reduction in authentication time or even the removal of the client as a valid client depending upon the request. Using screen shown in Fig.5.

5.2.5 Misbehaving server model

When the user enters into cloud server and the user will start to access the file, but at the same time an unauthorized user enters into the cloud server without the proper authentication to the cloud server the particular IP address will be noticed and it makes some attention to the cloud owner. Using screen shown in fig.6.

6. Conclusion

This paper briefly explained the problems of data security in cloud data storage. And also provided a way out to ensure user correctness. We propose a

7. Result

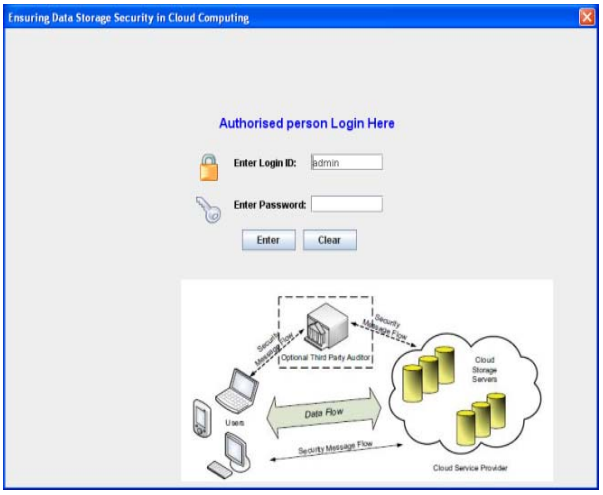

Fig-3: Authorized Person Login

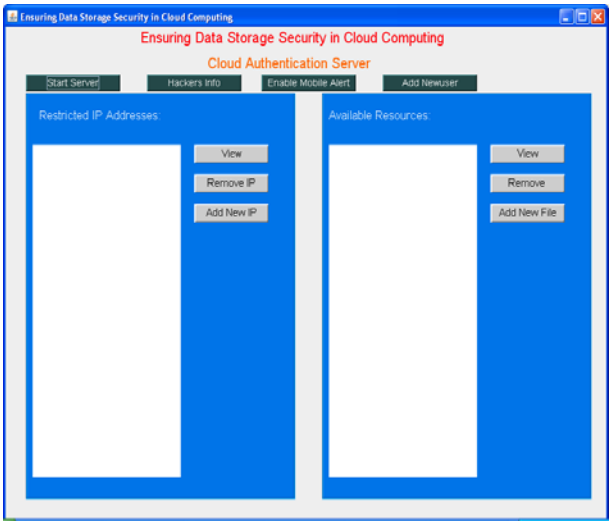

Fig-4: Cloud Data Storage

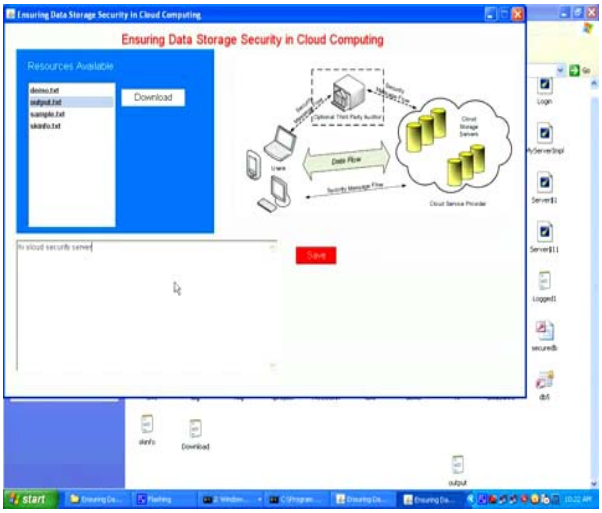

Fig-5: User Side Login

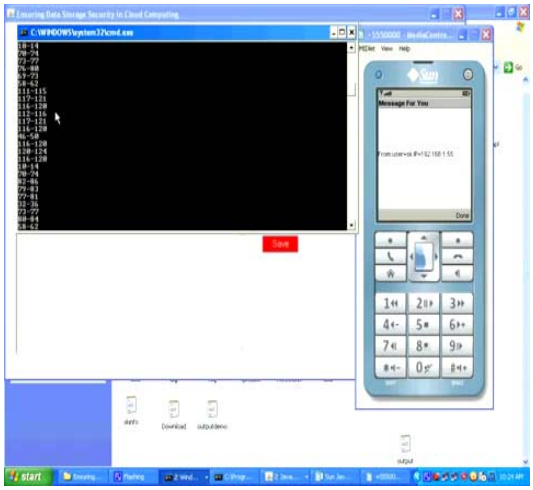

Fig-6: Misbehaving Server Model

distributed scheme through homomorphism token with distributed verification of erasure-coded data. Additionally, the technique provides a process to avoid colluding attacks of server modification by unauthorized users. We believe that data storage security in Cloud Computing, an area of challenges and of dominant significance, is still in its infancy to be identified. We envision several possible directions for future research on this area. It allows Third Parity Auditor to audit the cloud data storage without demanding users' time, probability.

References

[1] John W. Rittinghouse, James F. Ransome, "Cloud Computing Implementation, Management, Security", CRC Press 2009 by Taylor and Francis Group, LLC.

[2] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc.of Asiacrypt '08, Dec. 2008.

[3] G.Ateniese. R. D. Pietro ,L.V.Mancini and G. Tsudik, "Scalable and Efficient Provable Data Possession,"Proc of Securecomm'08.

[4] M.A Shah, M.Baker, J.C.Mogul, "Auditing to Keep Online storage services Honest,"Proc 11th USENIX Workshop on Hot Topics in Operating Systems(HOTOS'07). pp. 1–6, 2007.

[5] T.S.J.Schwarz and E.L.Millerds, "Store, Forget, and check: Using Algebraic Signatures to Check Remotely administered Storage,"proc.of ICDCS'06, pp. 12–12, 2006.

[6] Case study: http://eyeos.org/ cloud desktop

[7] L. Carter and M. Wegman, "Universal Hash Functions," *Journal of Computer and System Sciences*,

[8] http://searchcloudcomputing.techtarget.com/resources#parentTopic4

[9] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L.Kissner, Z. Peterson,and D. Song, "Provable Data Possession at Untrusted Stores," Proc. Of CCS '07, pp. 598–609, 2007.

[10] Cloud security Alliance"Top threats to Cloud *http://www.cloudSecurityalliance.org/topthreats/casthreats.v1.0.pdf*.