ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

Minimum Replication of User Data Integrating Anti Collusion Scheme in Cloud Groups

M.Mahendran[1], M.Mulk Raj[2], M.Gabriel[2], V.Siranjeevi[2]

Asst Professor, Department of Computer Engineering, Panimalar Engineering College, Poonamallee, Chennai, India[1]

U.G. Student, Department of Computer Engineering, Panimalar Engineering College, Poonamallee, Chennai, India[2]

ABSTRACT: In cloud computing, users can share data among group members with the characters of less maintenance and little management cost. Sharing data must have security guarantees, if they are out sourced. Sharing data while providing privacy preserving is still a challenging problem, when change of the membership. It might cause to the collusion attack for an unsecured cloud. For existing technique, security of key distribution is based on the secure communication channel, however, to have such channel is a strong assumption and is difficult for practice. We propose a secure data sharing scheme for dynamic users. Key distribution done without any secure communication channels and the user can get the individual key from group manager.
Data deduplication is one of the techniques which used to solve the repetition of data. The deduplication techniques are generally used in the cloud server for reducing the space of the server. To prevent the unauthorized use of data accessing and create duplicate data on cloud the encryption technique to encrypt the data before stored on cloud server. CloudMe is proposed for cloud storage. All files of data owners are encrypted using AES algorithm and stored in real cloud.

I. INTRODUCTION

Cloud Computing is an innovative technology that is revolutionizing the way we do computing. The key concept of cloud computing is that you don't buy the hardware, or even the software, you need anymore, rather you rent some computational power, storage, databases, and any other resource you need by a provider according to a pay-as-you-go model, making your investment smaller and oriented to operations rather than to assets acquisition.
Cloud computing can be defined as a model for enabling ubiquitous, convenient and on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort from the user side and minimal service provider interaction. Cloud computing is considered the evolution of a variety of technologies that have come together to change an organizations' approach for building their IT infrastructure. Actually, there is nothing new in any of the technologies that are used in the cloud computing where most of these technologies have been known for ages. It is all about making them all accessible to the masses under the name of cloud computing. Cloud is not simply the latest term for the Internet, though the Internet is a necessary foundation for the cloud, the cloud is something more than the Internet. The cloud is where you go to use technology when you need it, for as long as you need it. You do not install anything on your desktop, and you do not pay for the technology when you are not using it. The cloud can be both software and infrastructure. It can be an application you access through the Web or a server like Gmail and it can be also an IT infrastructure that can be used as per user's request. Whether a service is software or hardware, the following is a simple test to determine whether that service is a cloud service.
In the simplest terms, cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive. The cloud is just a metaphor for the Internet. It goes back to the days of flowcharts and presentations that would represent the gigantic server-farm infrastructure of the Internet as nothing but a puffy, white cumulus cloud, accepting connections and doling out information as it floats.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

Hybrid services like Box, Dropbox, and SugarSync all say they work in the cloud because they store a synced version of your files online, but they also sync those files with local storage. Synchronization is a cornerstone of the cloud computing experience, even if you do access the file locally.

The size of Cloud storage is expanding at a dramatic speed. It is estimated that by 2015 the data stored in the Cloud will reach 0.8 ZB while even more data is "touched" by the Cloud within the data lifecycle. Meanwhile, with the development of the Cloud computing paradigm, Cloudbased applications have put forward a higher demand for Cloud storage. While the requirement of data reliability should be met in the first place, data in the Cloud needs to be stored in a highly cost-effective manner. In this paper, our research focuses on minimizing the Cloud storage consumption by minimizing data replication while meeting the data reliability requirement.

Data owner: This is a client who owns data, and wishes to upload it into the cloud storage to save costs. A data owner encrypts the data and outsources it to the cloud storage with its index information, that is, a tag. If a data owner uploads data that do not already exist in the cloud storage, he is called an initial uploader; if the data already exist, called a subsequent uploader since this implies that other owners may have uploaded the same data previously, he is called a subsequent uploader. Hereafter, we refer to a set of data owners who share the same data in the cloud storage as an ownership group.

Cloud service provider: This is an entity that provides cloud storage services. It consists of a cloud server and cloud storage. The cloud server deduplicates the outsourced data from users if necessary and stores the deduplicated data in the cloud storage. The cloud server maintains ownership lists for stored data, which are composed of a tag for the stored data and the identities of its owners. The cloud server controls access to the stored data based on the ownership lists and manages (e.g., issues, revokes, and updates) group keys for each ownership group as a group key authority. The cloud server is assumed to be honest-but-curious. That is, it will honestly execute the assigned tasks in the system; however, it would like to learn as much information about the encrypted contents as possible. Thus, it should be deterred from accessing the plaintext of the encrypted data even if it is honest.

II. RELATED WORKS

Hierarchical identity-based encryption:

The concept of Identity Based Encryption (IBE) was pro-posed by Shamir [11] first in 1984, differing from traditional symmetrical encryption system, IBE took arbitrary character strings that can represent the identities of users, such as ID numbers, e-mail addresses, as public keys to encrypt data. One advantage of IBE is that the sender didnt have to search the public keys information on certificate authority (CA) online, which solved the problem of poor CA performance. The shortage of IBE system was that all users keys were generated by the private key generation (PKG), which would become the bottleneck in the system.

Horwitz [23] proposed the idea of hierarchical IBE (HIBE) in 2002, a user in the higher hierarchical position of the system could create private keys for lower position users with his/her private keys. Which mean that only the first level users private keys need be created by PKG, while lower-level users private keys could be generated and managed by their ancestors. This improved system relieved PKG of great burden and enhanced the system efficiency by authenticating identities and transporting keys within locality area instead of global area.

The public key of a user is described by a set of IDs composed of the public key of father node and the users own ID in the method of G-HIBE [9], the most important feature of the proposal is that the users public key could reflect precise position of the user in the hierarchical structure.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

III. PROPOSED METHODOLOGY AND DISCUSSION

- The users can securely obtain their private keys from group manager.
- User send request to group manager for access the wanted group, at that time our system provide individual secure key to user without activation.
- Then group manager see the requests and activate the keys after confirm them.
- After user's private key gets activation, then only user can access the group.
- Our scheme have fine-grained access control, any user in the group can use the source in the cloud and revoked users cannot access the cloud again after they are revoked.
- In our proposed system the group manager performs the below tasks when an new user joins the group or a user has left the particular group,
  - Update the whole user name list.
  - Generate a secure key and encrypt the key without activation and send to the updated user list.
  - Update the rights in the cloud server.
- We proposed public cloud named Dropbox for data storage.
- Group manager makes sure that the revoked users cannot access the file if they conspire with untrusted cloud.
- In using advanced de-duplication system supporting authorized duplicate check. In this new de-duplication system, a hybrid cloud architecture is introduced to solve the problem.

ADVANTAGE:

- By integrating algorithms / techniques we can implement deduplication concepts and reduce the storage cost in a cloud for the data owners.
- Secure protocol for anti-collusion attack.
- Faster recovery and processing of data.
- This would significantly decrease the processing time of load balancer.
- Effective and Efficient usage of cloud Storage Space

IV. PSEUDO CODE

Step 1:**Authority User Verification**
At first Initial stage all users must create own username and password
Step 2:**Detect Deduplication**
Deduplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.
Step 3:**Image Based effective authentication**
Step 4:**Key distribution &Access control**
Step 5:**Privacy-preserving**

V. SYSTEM DESCRIPTION

- For user authentication Image based password system to decrypt and encrypted the file based authentication. When the Admin uploads the file in the cloud, the admin will split the image into 4 parts. The admin will hold 2 parts and the user of that respective group can view the other 2 parts. The images are spilt randomly using pseudo random generator technique. When the user tries to download a file, the user can send the requisition to the respective admin along with the user side available 2 parts. The admin will verify both the parts and if the authentication is passed, the file will be sent to the user in an encrypted way
- In our proposed project, we propose a secure architecture for handling file access in a dynamic cloud group. The user belonging to an particular group is analysed and identified. After that a private key is sent to the user

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 2, February 2018

by the group manager in a encrypted format using RC4 encryption algorithm. The group manager performs the below tasks when an new user joins the group or a user has left the particular group,

- Update the whole user name list.
- Generate a secure key and encrypt the key without activation and send to the updated user list.
- Update the rights in the cloud server.

- Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage. The technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. Deduplication can take place at either the file level or the block level. For file level deduplication, it eliminates duplicate copies of the same file. Deduplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.

VI. CONCLUSION

- Our proposed scheme provides a possible way to fight against immoral interference with the right of privacy.
- We hope more schemes can be created to protect cloud user privacy.

REFERENCES

1. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “A view of cloud computing,” Commun. ACM, vol. 53, no. 4, pp. 50–58, Apr. 2010.
2. S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Proc. Int. Conf. Financial Cryptography Data Security, Jan. 2010, pp. 136–149.
3. M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, “Plutus: Scalable secure file sharing on untrusted storage,” in Proc. USENIX Conf. File Storage Technol., 2003, pp. 29–42.
4. E. Goh, H. Shacham, N. Modadugu, and D. Boneh, “Sirius: Securing remote untrusted storage,” in Proc. Netw. Distrib. Syst. Security Symp., 2003, pp. 131–145.
5. G. Ateniese, K. Fu, M. Green, and S. Hohenberger, “Improved proxy re-encryption schemes with applications to secure distributed storage,” in Proc. Netw. Distrib. Syst. Security Symp., 2005, pp. 29–43.