Cloud Data Storage for Group Collaborations

Jyh-Shyan Lin

***Abstract*—Cloud computing has been an important development trend in information technology. By moving data and application software from traditional local hosts to network servers, cloud computing provides more flexible and convenient access to data and services, with cheaper software obtainment and hardware maintenance costs. Cloud computing also may provide some value-added services, such as automatic data backup and group collaboration support. Many researches about cloud computing have been proposed in the literature. However, how these methods could be used for group collaborations is still unclear. In this paper we develop a secure cloud data storage scheme which can be used for group collaborations.**

***Index Terms*—Cloud computing, Cloud data storage, Security**

I. Introduction

In recent years, mobile communication devices and mobile computing devices become increasingly popular. The computing power of lightweight netbooks and smart phones is growing stronger. Many electronic devices, such as medical diagnosis and healthcare instruments, electrical facilities, automobile equipments, and home appliances, are gradually toward the development of network interconnection. At the same time, network communication environment is also well constructed. Broadband networks and wireless communication networks become more and more speedy and popular. These facts, in addition to the demand that people desire flexible, convenient, and geographical location independent access to data and services, bring forth the era of cloud computing. In cloud computing, data and application software are moved from traditional local hosts to remote data centers and application servers, providing on-demand service, heterogeneous and ubiquitous network access, location independent resource pooling, rapid resource elasticity, and usage-based pricing [4]. Cloud computing also may provide some value-added services, such as automatic data backup and group collaboration support. Users can use various client devices, no matter desktop PCs or lightweight thin client devices such as netbooks and smart phones, to subscribe services from cloud server providers with relatively cheaper software and hardware costs, and relief from the complexity of direct hardware maintenance.

Although cloud computing possesses so many advantages as stated above, security is a seriously concerned issue. There must be some ways to convince users that the data stored in the cloud are intact, confidential, and retrievable. There are many researches that address the issues of integrity and retrievability of cloud computing, for example [1], [2], [3], [5] and [6]. However, how these methods could be used for group collaboration is still unclear. We claim that group collaboration is an important application in cloud computing. On one hand, the client-server essence of cloud computing makes it suitable to support group collaboration. On the other hand, the requirement for group collaboration may be the main reason for an organization or a corporation deciding to subscribe a cloud service. In this paper, we propose a secure and public verifiable cloud data storage scheme which can be used for group collaboration. This is the contribution of this paper.

II. Architecture

A representative cloud data storage architecture for group collaboration is illustrated in Figure 1. In the architecture, messages between entities are transmitted by secured channels. The entities in the architecture are described as follows:

- **User**: A user is an individual who stores data in the cloud storage system. Users may be organized as a group, in which a group manager is responsible for adding and removing group members. Members of a group may work on a set of data on behalf of the group.
- **Cloud Service Provider (CSP)**: A CSP is an organization that has abundance of resources and expertise in order to construct and maintain a cloud data storage system.
- **Third Party Auditor (TPA)**: A TPA is an agency authorized by users or groups to verify the integrity and the retrievability of their data. In the cloud data storage architecture, TPA is an optional entity.

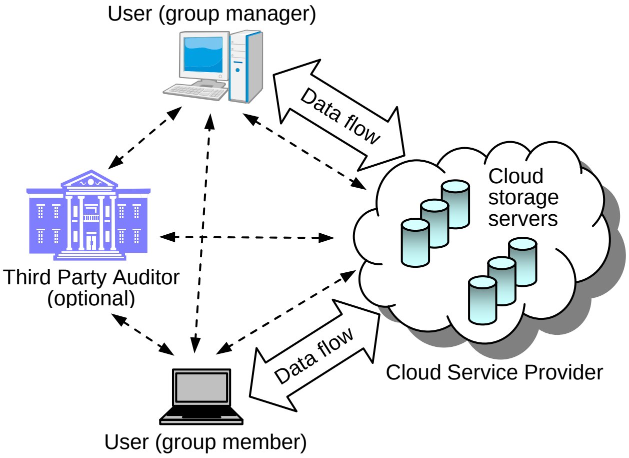

Figure 1. Cloud data storage architecture for group collaboration

Manuscript received March 1, 2010. This work was supported in part by the National Science Council of the Republic of China under contracts NSC 98-2221-E-264-009-.

J.-S. Lin is with the Department of Information Management, Yuanpei University, Hsinchu, Taiwan, ROC (phone: :+886-3-5381183 ext. 8294; fax: :+886-3-5381183 ext. 2362; e-mail: jslin@ mail.ypu.edu.tw).

A CSP provides a vast amount of storage space that shared by all users. The storage space is usually constructed by multiple storage servers in a distributed manner. All storage servers work simultaneously and collaboratively. In order to ensure the accuracy of stored data, appropriate redundancies may be stored in the storage servers and use error correction codes or erasure correction codes to prevent data loss due to accident or deliberate destruction. Users can access their private data through the interfaces provided by the CSP and manipulate the data with appending, insertion, modification, and deletion operations according to their well. The redundancies in the storage servers must be adjusted immediately corresponding to the changes made by these manipulations. In order to let users fell relieved to store their data on the cloud storage servers, there must be some ways to convince users that the data stored in the cloud are intact, confidential, and retrievable. For data confidentiality, traditional cryptographic primitives can be applied. For data accuracy and retrievability, traditional cryptographic primitives are not sufficient since users are unable to directly control the data stored in the cloud, and downloading the whole data from the storage servers for verification is impracticable when the size of the data is huge. We need an efficient method by which users can verify their data, and the verification will cause little computational load to the storage servers. Furthermore, the amount of the message transmitted between the users and the CSP for the verification is as small as possible. When the data are shared by a group of users, i.e. under group collaborations, every member of the group can verify the data independently. This is a challenge when data are updated dynamically, and is even more challenging under group collaborations.

III. Proposed Scheme

As described in the previous section, users and TPAs must have an efficient way to verify specific data stored in the cloud. In the proposed scheme, verifications are carried out by a challenge-response interaction. A user or a TPA can submit a request to the CSP as a challenge. The CSP then computes a value corresponding to the challenge and sends it back to the user or the TPA as a response. If the response coincides with the knowledge about the data, then it has proved that the data stored in the storage servers are intact and retrievable. The complete scheme contains the following procedures:

- **Setup**: This procedure generates parameters used in the scheme.
- **GrpSetup**: On input a group ID, this procedure generates the secret key for the group.
- **Join**: This procedure is executed by a group manager and, on input a user ID, adds the user into the group and generates the secret key for the user.
- **SigGen**: This procedure executed by a user generates the verification metadata for a data block.
- **GenProof**: This procedure executed by the CSP generates the proof for the verification of a data block.
- **CheekProof**: Users or TPAs use this procedure to validate the proof generated by the CSP.

We adapt the methods as in [7] to accomplish the Setup, GrpSetup, and Join procedures, and the method as in [8] to accomplish the SigGen, Genproof, and CheckProof procedures. The identity based approach of this scheme makes it possible for public verification. That is, a TPA can be authorized by users or groups to verity the accuracy and retrievability of their data.

IV. Conclusion

We have proposed a secure cloud data storage scheme for group collaborations. A group of users can operate on a set of data collaboratively with appending, insertion, modification, and deletion operations. Every member of the group can verify the data independently. The verification can also be authorized to a TPA for convenience.

References

[1] 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.

[2] K. D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," *Cryptology ePrint Archive*, available at http://eprint.iacr.org/2008/489.

[3] A. Juels and J. Burton S. Kaliski, "PORs: Proofs of Retrievability for Large Files," *Proc. of CCS '07*, pp. 584–597, 2007.

[4] P. Mell and T. Grance, "Draft nist working definition of cloud computing," Online at http://csrc.nist.gov/groups/SNS/cloud-computing/index.html, 2009.

[5] J. S. Plank and Y. Ding, "Note: Correction to the 1997 Tutorial on Reed-Solomon Coding," *University of Tennessee, Tech. Rep. CS-03-504, 2003.*

[6] A. Shamir, "Identity-based Cryptosystems and Signature Schemes", *Advances in Cryptology– CRYPTO'84*, pp. 47-53.

[7] N. P. Smart and B. Warinschi, "Identity Based Group Signatures from Hierarchical Identity-Based Encryption," *Pairing 2009*, pp.150-170.

[8] 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, Saint Malo, France, Sep. 2009.*