Chapter 21
An Authentication Technique for Accessing De-Duplicated Data From Private Cloud Using One Time Password

Prakash Mohan
Karpagam College of Engineering (Autonomous), India

Saravanakumar Chelliah
St. Joseph's Institute of Technology, India

ABSTRACT

Objective: The main aim is to de-duplicate the redundant files in the cloud and also to improve the security of files in public cloud service by assigning privileges to the documents when it is uploaded by confidential user. Methods: To achieve the objective the authors have used the AES algorithm to encrypt the file stored after de-duplication in the cloud. De-duplication is done based on comparison of contents, file type and size. For an authorized user to access the file from the cloud, generation of OTP using SSL protocol is adopted. Findings: Files uploaded in the cloud are encrypted using traditional encryption algorithms which don't provide high levels of security. Files can be accessed by anyone who is authorized. Privileges are not considered. During de-duplication, only the name and size of the files are considered. Application: Files within the public cloud can't be viewed by everyone who has registered with the cloud. Those who have the respective privileges can only view the file. Proof of Ownership is assured. Since de-duplication is done based on the content redundancy within the cloud storage is avoided. Usage of OTP ensures that the content is viewed by the individuals who have the respective privileges related to the file. These concepts provide additional security to the files stored in the public environment.

INTRODUCTION

Cloud computing technology is used to store enormous amount of data and appear to be a virtual resource to the users. It is dynamic and can be easily accessed from anywhere provided with internet. It encapsulates the platform and execution details from the user. Instead of using costly hardware components, cloud

DOI: 10.4018/978-1-5225-8176-5.ch021

service is comparatively cheap. It is extensible, scalable and updated with ease. Ex: If the user currently has 2GB of space and is in need of further storage space (Li, 2013; Itani, 2009), he can expand it easily. Private cloud provides more security (Mohan, 2013; Popović, 2010; Prakash, 2012) with less storage space. It can be accessed easily. It is suitable to use within the organisation. Data (Annamalai, 2015) can be accessed based on privileges. The keys for the files stored in public cloud are usually stored in private cloud. On contrary, public cloud provides data storage with less security. It is mostly concerned with the private cloud. To secure the data from losing its confidentiality, privileges (Prakash, 2015; Annamalai, 2015) are given to the files, so that only specific people can access the file. Privileges can be given both the types of cloud service (Saravanakumar, 2012). Authorization to the cloud is provided based on the credentials stored in the database during registration with the cloud. De-duplication is a data compression technique used to eliminate the redundant copies in the cloud enhancing (Mell, 2009; Khan, 2016) the storage capacity. It is done at both file level and block level. In file level, it eliminates the duplicated files and at block level redundant blocks in the file are eliminated in non-identical files. The file attributes like size, content and type are checked. Privacy concerns are present due to insider and outsider attacks. Data are encrypted (Corena, 2012; Ryan, 2011) for security (Pearson, 2013; Subashini, 2011) reasons. In traditional encryption, when the same file is uploaded by different people different cipher texts are created for each individual. This makes de-duplication difficult. In convergent encryption, a convergent key is generated by calculating the cryptographic hash value of the file. This key is used to encrypt or decrypt the file. Keys are present with the user and the cipher texts thus generated are stored in the cloud. Here the cipher text produced for identical copies of file will be same and helps in de-duplication. Proof of ownership is provided to the files to ensure the user holds the file in spite of duplicate copies. When convergent encryption is used de-duplication of cipher text is possible and proof of ownership helps to enhance confidentiality. The de-duplication systems based on this fail to provide duplicate check with privileges. Issues arise when de-duplication with privileges are tried to be implemented at same time.

RELATED PREVIOUS WORK

Venkatesh, Sharma, Desai et al. (2014), aimed to minimize the data duplication along with data. Security is to protect the confidentiality of the data. The security is provided by using many encryption techniques to encrypt the data before outsourcing. The users are checked whether they are authorized or not. Encryption is symmetric. SCSP is used to reduce de-duplication. The algorithm used here is novel encryption key generation algorithm. During the process of uploading the file, a tag is generated. It helps to identify the duplicates. These tags are stored in a separate table. The project has the advantage that the system is suitable for backup storage by using authorized de-duplication. In data duplication the encrypted keys are generated by private key cloud server. But the decryption of cipher text cannot be done by private cloud server and S-CSP based on security of symmetric encryption with traditional data duplication in cloud computing is semantically secure. The proposed system is storing the authorised privileges.

Kumaresan, and Visuwasam (2015) had tried to overcome the disadvantage of traditional encryption techniques using convergent encryption. Hash calculation is done at block level. If the target device finds a duplicate, then it doesn't store a duplicate block. Instead it references to the existing block. This takes quite a longer time. Data Duplication is an important technology used in many companies to save a lot of money on storage cost and bandwidth by avoiding the replication. This paper proposed that the server will pop-up the duplication message if a file duplication is found. The security is provided by