International Journal of Emerging Technology and Advanced Engineering

Authorized Hybrid Cloud Storage Service Provider with Normalization

Dinesh Shinde¹, Amit Dangi²
M.Tech¹, Asst. Professor², CSE, IITM Bhopal, M.P., India

Abstract—Data deduplication is one of important data compression techniques for eliminating duplicate copies of repeating data, and has been widely used in cloud storage to reduce the amount of storage space and save bandwidth. To provide security to achieve basic goals of computer security introduced convergent key to convert plaintext data to ciphertext format. To better protect data security, in this to overcome the problem of data duplication we used proposed system, which is different from recent applications, the differential priorities of users are used for to check duplicate data content itself. To check duplicate data content in hybrid cloud we introduce new deduplication techniques in proposed system. This method is secured and provide confidentiality as per data analysis. Conduct testbed experiments to check duplicate content using our method. In this paper we implemented that this system eliminate number of overheads.

Keywords-- Deduplication, confidentiality, hybrid cloud, ciphertext, convergent key.

I. INTRODUCTION

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. Although data deduplication brings a lot of benefits, security and privacy concerns arise as users’ sensitive data are susceptible to both insider and outsider attacks. Traditional encryption, while providing data confidentiality, is incompatible with data deduplication. Specifically, traditional encryption requires different users to encrypt their data with their own keys. Thus, identical data copies of different users will lead to different cipher texts, making deduplication impossible. Convergent encryption has been proposed to enforce data confidentiality while making deduplication feasible. It encrypts/decrypts a data copy with a convergent key, which is obtained by computing the cryptographic hash value of the content of the data copy. After key generation and data encryption, users retain the keys and send the ciphertext to the cloud.
II. MODULES

User Description:
In this users having valid authentication, access permission and security to access the detail which is presented in the ontology system. Before accessing or searching the details user should have the authority in that otherwise they should registered first.

DeDuplication System:
To support authorized deduplication, the tag of a file $F$ will be determined by the file $F$ and the privilege. To show the difference with traditional notation of tag, we call it file token instead.

To support authorized access, a secret key $kp$ will be bounded with a privilege $p$ to generate a file token. Let $\phi' F;p = \text{TagGen}(F, kp)$ denote the token of $F$ that is only allowed to access by user with privilege $p$. In another word, the token $\phi' F;p$ could only be computed by the users with privilege $p$. As a result, if a file has been uploaded by a user with a duplicate token $\phi'F;p$, then a duplicate check sent from another user will be successful if and only if he also has the file $F$ and privilege $p$. Such a token generation function could be easily implemented as $H(F, kp)$, where $H(\_)$ denotes a cryptographic hash function.[11]
Security Of Duplicate Check Token:

We consider several types of privacy we need protect, that is, i) unforgeability of duplicate-check token; There are two types of adversaries, that is, external adversary and internal adversary. As shown below, the external adversary can be viewed as an internal adversary without any privilege. If a user has privilege p, it requires that the adversary cannot forge and output a valid duplicate token with any other privilege p' on any file F, where p does not match p'. Furthermore, it also requires that if the adversary does not make a request of token with its own privilege from private cloud server, it cannot forge and output a valid duplicate token with p on any F that has been queried.

Send Key:

Once the key request was received, the sender can send the key or he can decline it. With this key and request id which was generated at the time of sending key request the receiver can decrypt the message.

III. METHODOLOGY

In cloud computing we uses various techniques or new deduplication method to check identical data to eliminate and reduce storage space on hybrid cloud. following are the main function to check duplicate data.

• Differential Authorization. Every legitimate user is able to get individual token of his file to perform duplicate check based on his priority to access particular file. Under this assumption, any user cannot generate a token for duplicate check out of his priority or without the access from the private cloud server.

• Legitimate Duplicate Check. Legitimate user is able to use individual private keys to generate query for certain file and the priority owned with the help of private cloud, while the public cloud performs duplicate check directly and tells the user if there is any duplicate.

The security requirements considered in this paper lie in two folds, including the security of file token and security of data files. For the security of file token, two aspects are defined as duplicate token check and unauthorized user authentication of file token. The details are given below.

• Verify Duplicate token. Unauthorized users without any authority to access particular file or file should be prevented from getting or generating the file tokens for duplicate check of any file stored at the S-CSP. The users are not allowed to collude with the public cloud server to break the unforgeability of file tokens. In our system, the S-CSP will honestly perform the duplicate check upon receiving the duplicate request from users. The duplicate check token of users should be given from the private cloud server in our scheme.

• Indistinguishability of file token. It requires that any user without querying the private cloud server for some file token, he cannot get any useful information from the token, which includes the file information or the privilege information.

• Confidential Information. Unauthorized users without any priority access particular file or files, including the S-CSP and the private cloud server, should be prevented from access to the underlying plaintext stored at S-CSP. In another word, the goal of the adversary is to retrieve and recover the files that do not belong to them. In our system, compared to the previous definition of data confidentiality based on convergent encryption, a higher level confidentiality is defined

IV. PROPOSED SYSTEM

Prototype of the proposed authorized deduplication system, in which we model three entities as separate C++ programs. A Client program is used to model the data users to carry out the file upload process.
A PrivateServer program is used to model the private cloud which manages the private keys and handles the file token computation. A Storage Server program is used to model the SCSP which stores and deduplicated file implementation of the Client provides the following function calls to support token generation and deduplication along the file upload process.

In this proposed system it check duplicate content on private cloud using hash value if it matches with another data which is already had same hash value then it is rejected that data before the data uploaded and if data is unique then data is accepted and uploaded then it is downloaded from public cloud. the data uploaded on public cloud it is in encrypted form for data confidentiality to protect data from unauthorized user.

1. FileTag Pointer (File) – Hash algorithm SHA is used to hash of the File as File Tag;
2. Token(Tag, UserID) It requests the Private Server for File Token generation with the File User ID;
3. It requests the Storage Server for Duplicate Check of the File by sending the file token received from private server using token duplicate checker;
4. ShareToken(Tag, [Priv.]) It requests the Private Server to generate the Share Token with the File Tag and Target Sharing Privilege Set;

V. CONCLUSION

In this system authorized data deduplication was introduced to authorized data security by including different authorities given to users according to designation in duplicate check data in hybrid cloud architecture. In which duplicate of check of token is generated by private cloud server with private keys. In term of insider and outsider attack is eliminated and only authorized unique content is uploaded which is proved in security analysis model. We showed to secure data and provide confidentiality we compressed data by using lossless compression method which assured no data loss at the time of compression.

REFERENCES