International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue:08 | Aug 2019 www.irjet.net p-ISSN: 2395-0072

DATA DEDUPLICATION SECURITY WITH DYNAMIC OWNERSHIP MANAGEMENT

Tanvi Rudrashetty[1]

[1]*B. Tech Student, Department of CSE, Gurunanak Institutions, Ibrahimpatnam, Hyderabad, India*

Abstract - *To provide security of deduplication system with high authentication during which the information is being distributed across multiple cloud servers. Deduplication technique eliminates redundant data and stores only a single copy of that data. By the use of hash table formula, the requirement of information confidentiality and reducing space storage are achieved. It converts a range of key values into a range of indexes of an array. It performs a basic operations of a research, insert, and deletion. As a result, security analysis demonstrate that our deduplication systems are secure in the proposed system. To protect the confidentiality of sensitive data while supporting deduplication, the hash table algorithm technique has been proposed to encrypt the data before outsourcing. To safeguard information security which has been different from conventional deduplication systems, the differential rights of users are further considered induplicate check besides the data itself. Data deduplication is one of the principal information compression approach for terminating duplicate copies of redundant information, and has been widely incorporated in cloud storage to decrease the quantity of storage space and save bandwidth. To protect the privacy of sensitive information while supporting deduplication, the convergent coding technique has been planned to encode the information before outsourcing.*

Key Words: Security, Data De-duplication, Convergent Encryption, Privacy, Authentication.

1. INTRODUCTION

Because of the importance in serving users to form real time decisions, information dissemination has become important in several large-scale emergency applications, like earthquake observation, disaster weather warning, and status update in social networks. Recently, information dissemination in these emergency applications presents a number of contemporary trends. One is the rapid growth of live content. For example, Facebook users publish over 600,000 pieces of content and Twitter users send over 100,000 tweets on average per minute.

The other is the highly dynamic network environment. For instance, the measurement studies indicate that most users' sessions in social networks only last several minutes. In emergency eventualities, the fulminant disasters like earthquake or weather condition might cause the failure of a large range of users instantly. These characteristics require the data dissemination system to be scalable and reliable. Firstly, the system should be scalable to support the massive quantity of live content. The key is to supply an ascendable event matching service to separate extraneous users. Otherwise, the content may have to traverse a large number of uninterested users before they reach interested users. Secondly, with the dynamic network environment, it's quite necessary to provide reliable schemes to keep continuous data dissemination capacity. Otherwise, the system interruption might cause the live content becomes obsolete content. Driven by these necessities, publish/subscribe (pub/ sub) pattern is widely used to disseminate data due to its flexibility, scalability, and efficient support of complex event processing. In pub/sub systems (pub/subs), a receiver (subscriber) registers its interest in the form of a subscription. Events are published by senders to the pub/ sub system. In conventional information dissemination applications, the live content is generated by publishers at an occasional speed, that makes several pubs/subs adopt the multi-hop routing techniques to distribute events. A large body of broker-based pub/subs forward events and subscriptions through organizing nodes into diverse distributed overlays, such as tree based design cluster-based design and DHT-based design. However, the multichip routing techniques in these broker-based systems cause an occasional matching turnout, that is insufficient to apply to current high arrival rate of live content.

2. EXSISTING SYSTEM

In the existing deduplication system, each user is issued a set of privileges during system initialization.

Each file uploaded to the cloud is additionally finite by a group of privileges to specify which type of users is allowed to perform the duplicate check and access the files. Before presenting his duplicate check request for a file, the user needs to take that particular file and his very own benefits as inputs. The user is in a position to seek out a replica for this file if and provided that there's a duplicate of this file and a matched privilege stored in cloud.

3. PROPOSED SYSTEM

The proposing system, we terminating redundant copies of repeating information has been widely used in cloud storage to minimize the amount of storage capacity and save bandwidth. To protect the privacy of sensitive information while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing. To better protect information security, this

paper makes the first attempt to formally address the problem of authorized data deduplication.

4. SYSTEM ARCHITECTURE

Architecture diagram shows the connection between completely different elements of system.

This diagram is extremely necessary to know the general idea of system. Architecture representation might be a plan of a system, within which the essential components or roles and tasks are described by blocks connected by lines that show the relationships of the blocks. They are heavily employed in the engineering world in hardware model, electronic plan, package representation, and process flow diagrams.

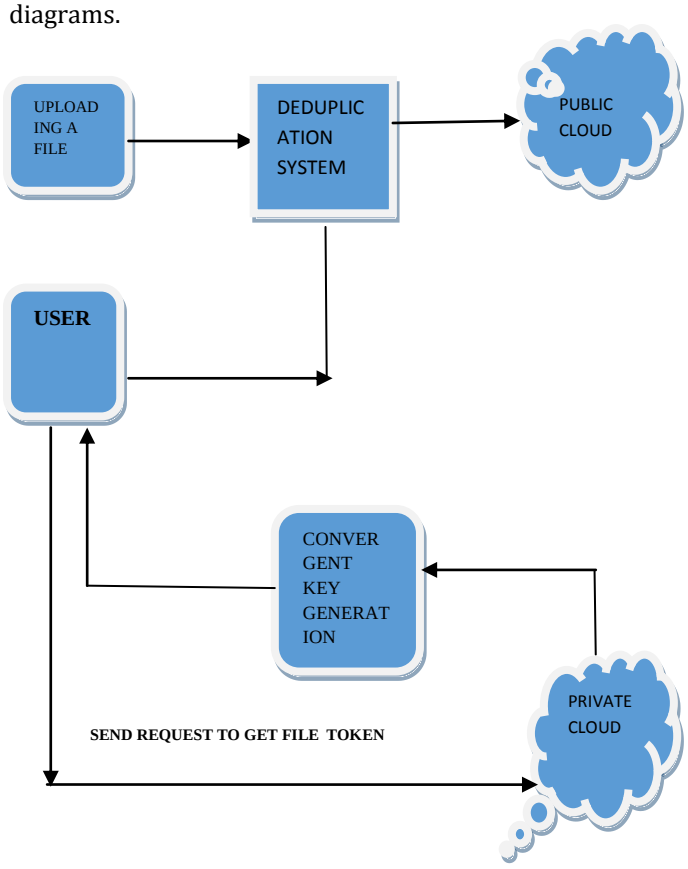

Fig-1: System Architecture

5. IMPLEMENTATION

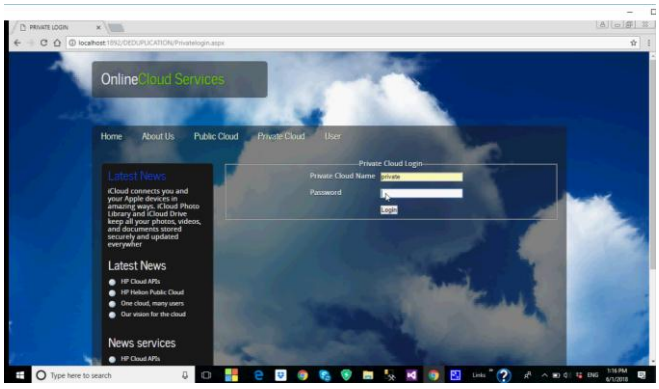

Fig-2.1: Cloud Login

Description: Page when we run the project where the authentication server provides authorized logins.

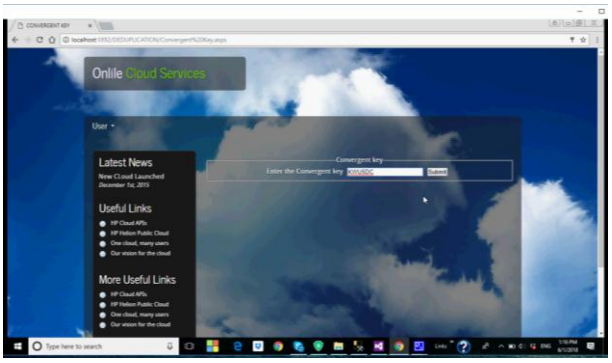

Fig-2.2 Entering Convergent key

Description: creating convergent key for securing data deduplication where we get access to user Id details.

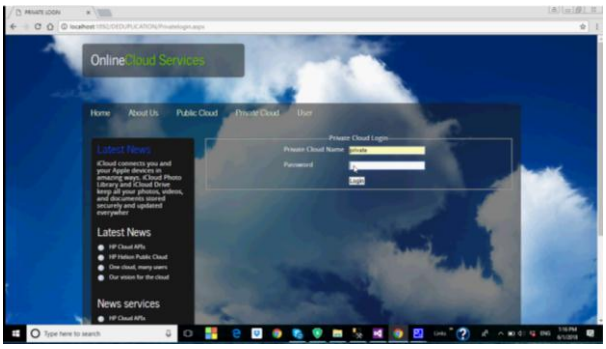

Fig-2.3 Private cloud Login

Description: Private cloud connects to public cloud infrastructure, allowing an organization to orchestrate workloads across the two environments.

Fig 2.4 Document creation and data representation

Description: Document creation for security purpose after user access.

6. CONCLUSIONS

The notion of approved information deduplication was planned to safeguard the information security by including divergent rights of users within the duplicate check.

In which the replicated -check tokens of files are generated by the private cloud server with private(special) keys.

7. FUTURE ENHANCEMENT

In future there are still some possible extensions of our current work remaining. If user wants to modify a file in already stored in cloud, user can modify.

REFERENCES

1. Paul Anderson, Le Zhang[1], Fast and Secure Laptop Backups with Encrypted De-duplication,2011

2. Pasquale puzio secludit and eurecom, refik molva eurecom, melek o nen eurecom[2], Clouded up: secure deduplication with encrypted data for cloud storage,2012

3. N.O.agrawal,Prof Mr. s.s. kulkarni[3], Secure deduplication and data security with efficient and reliable CEKM,2014

4. Pierre-Louis Cayrel1, Philippe Gaborit1 and Marc Girault2[4], Identity-based signature schemes using correcting codes,2011.

5. B.C. Tea, M.R.K. Ariffin and J.J. Chin[5], An Efficient Identification Scheme in Standard Model Based on the Diophantine Equation Hard Problem,2013.