Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 6, Issue 06, June -2019

"Efficient Client-Side Deduplication of Encrypted Data in Cloud Storage"

Tejaswini Joshi[1], Vikas Singh[2], Raj Waykos[3], Swati Jadhav[4], Tejaswini Kailuke[5]

[1]*Department of Information Technology, PCCOE Pune, Maharashtra, India*
[2]*Department of Information Technology, PCCOE, Pune, Maharashtra, India*
[3]*Department of Information Technology, PCCOE, Pune, Maharashtra, India*
[4]*Department of InformationTechnology, PCCOE, Pune, Maharashtra, India*
[5]*Department of InformationTechnology, PCCOE, Pune, Maharashtra, India*

Abstract — *Attribute-based secret writing (ABE) has been wide used in cloud computing where knowledge|a knowledge|an information} provider outsources his/her encrypted information to a cloud service provider, and will share the information with users possessing specific credentials (or attributes). However, the standard ABE system does not support secure DE duplication, that's crucial for eliminating duplicate copies of identical data thus on save a lot of house for storing and network metric. throughout this paper, we've a bent to gift associate attribute-based storage system with secure deduplication throughout a hybrid cloud setting, where a private cloud is answerable for duplicate detection and a public cloud manages the storage. Compared with the previous data deduplication systems, our system has two blessings. Firstly, it square measure usually accustomed confidentially share data with users by specifying access policies rather than sharing secret writing keys. Secondly, it come back throughs the standard notion of linguistics security for data confidentiality whereas existing systems only bring home the bacon it by shaping a weaker security notion. to boot, we've a bent to position forth a method to change a ciphertext over one access policy into ciphertexts of identical plaintext but at a lower place completely different access policies whereas not revealing the underlying plaintext.*

Keywords *- Attribute-based encryption, access control, audit logs, broadcast encryption, delegation, hierarchical identitybased encryption.*

INTRODUCTION

Cloud computing greatly facilitates info suppliers World Health Organization need to provide their info to the cloud whereas not revealing their sensitive info to external parties and would love users with certain credentials to be able to access the information. this needs info to be hold on in encrypted forms with access management policies such nobody except users with attributes (or credentials) of specific forms will decipher the encrypted info. degree cryptography technique that meets this demand is termed attribute-based cryptography (ABE), wherever a user's personal secret's related to degree attribute set, a message is encrypted below degree access policy (or access structure) over a bunch of attributes, and a user will decipher a ciphertext with his/her personal key if his/her set of attributes satisfies the access policy related to this ciphertext. However, the quality ABE system fails to understand secure deduplication, which can be a way to avoid wasting space for storing and network system of measure by eliminating redundant copies of the encrypted info hold on among the cloud. On the choice hand, to the foremost effective of our knowledge, existing constructions for secure deduplication aren't designed on attribute-based cryptography. however, since ABE and secure deduplication unit wide applied in cloud computing, it'd be fascinating to vogue a cloud storage system possessing each properties.

LITERATURE REVIEW

1.**Paper name:**Cloud Cryptography: Theory, apply and Future analysis Directions
Author: Elsevier B.V.
Cloud computing, a convenient manner of accessing services,resources and applications over the net, shifts the main target of industries ANd organizations faraway from the preparation and day-after-day running of their IT facilities by providing an on-demand,self-service, and pay-as-you-go business model. It is, therefore,unsurprising that cloud computing has continuing to extend in quality in recent times.

2.Paper name:Cloud primarily based information sharing with fine-grained proxy re-encryption
Author: Yanjiang rule a , Haiyan Zhu , Haibing Lu , Jian Weng , Youcheng Zhang ,
Kim-Kwang Raymond Choo.
Conditional proxy re-encryption (CPRE) allows fine-grained delegation of secret writing rights, and has several real-world applications. during this paper, we tend to gift a ciphertext-policy attributebased CPRE scheme,together with a formalization of the primitive an dits security analysis. we tend to demonstrate the utility of the theme during a cloud

preparation, that achieves fine-grained information sharing. This application implements cloud server-enabled user revocation, providing another nevertheless additional economical answer to the user revocation downside in the context of fine-grained encoding of cloud information. High user-side potency is another prominent feature of the application,which makes it possible for users to use resource con-strained devices, e.g., mobile phones, to access cloud information. Our evaluations show promising results on the performance of the planned theme.

3. **Paper name:** Google Drive: rhetorical analysis of information remnants
Author: Darren fast n , Kim-Kwang Raymond Choo
Cloud storage is AN rising challenge to digital rhetorical examiners. The services square measure progressively employed by shoppers, business, and government, and may doubtless store giant amounts of information. The retrieval of digital proof from cloud storage services (particularly from offshore providers) is a challenge during a digital rhetorical investigation, thanks to virtualisation, lack of data on location of digital proof,privacy problems, and legal or territorial boundaries. Google Drive may be a common service, providing users an economical, and in some cases free, ability to access, store, collaborate, and air information. victimization Google Drive as a case study, artefacts were known that square measure probably to stay once the utilization of cloud storage, within the context of the experiments, on a laptop disc drive and Apple iPhone3G, and therefore the potential access point(s) for digital forensics examiners to secure proof

4.Paper name: Fuzzy Identity-Based encoding
Author: Mihir Bellare, Sriram Keelveedhi ,Thomas Ristenpart
Cloud storage service suppliers like Dropbox, Mozy,and others perform deduplication to avoid wasting house by solely storing one copy of every file uploaded. ought to shoppers conventionally cypher their files, however, savings square measure lost. Message-locked encoding (the most distinguished manifestation of that is focused encryption) re-solves this tension. but it's inherently subject to brute-force attacks that may recover files falling into a best-known set. we tend to propose AN design that gives secure deduplicated storage resisting brute-force attacks, and are aware of it during a system known as DupLESS. In DupLESS, shoppers cypher underneath message-based keys obtained from a key-server via AN oblivious PRF protocol.It allows shoppers to store encrypted information with AN existing service, have the service perform deduplication on their behalf, and nevertheless achieves sturdy confidentiality guarantees. we tend to show that encoding for deduplicated storage can do performance and house savings on the brink of that can deliver the goods performance and house savings on the brink of that can do performance and house savings on the brink of that of victimization the storage service with plaintext information.

5.Paper name: Attribute-Based encoding for Fine-Grained Access management of Encrypted information
Author: Vipul Goyal Omkant Pandey Amit Sahai brent goose Waters
As additional sensitive information is shared and hold on by third-party sites on the net, there'll be a requirement to cypher information hold on at these sites. One disadvantage of encrypting information, is that it is by selection shared solely at a coarse-grained level (i.e., giving another party your personal key). we tend to develop a brand new cryptosystem for fine-grained sharing of encrypted information that we tend to decision Key-Policy Attribute-Based encoding (KP-ABE). In our cryptosystem, ciphertexts square measure labelled with sets of attributes and personal keys square measure related to access structures that management that ciphertexts a user is ready to decode. we tend to demonstrate the pertinency of our construction to sharing of audit-log info and broadcast encoding. Our construction supports delegation of personal keys that subsumes ranked Identity-Based encoding (HIBE).

EXISTING SYSTEM

In the existing the cloud service supplier, and might share the information with users possessing specific credentials (or attributes). within the current system the quality ABE system doesn't support secure deduplication, that is crucial for eliminating duplicate copies of identical data so on save many space for storing and network system of measure.

Disadvantages of existing system:

-System doesn't support secure de-duplication

-Access policies whereas not revealing the underlying plaintext.

-Existing systems alone succeed it by shaping a weaker security notion

PROPOSED SYSTEM

We gift associate attribute-based storage system with secure deduplication in associate exceedingly} terribly hybrid cloud setting, wherever a private cloud is guilty for duplicate detection and a public cloud manages the storage.

-The auditor is also a honest organization, which might give unbiased auditing results for homeowners.

-TPA(Third Party Auditor) give associate economical secure deduplication theme.
-Regenerate code through proxy server. this method is been developed to supply integrity and regenerating code.

Advantages Of planned system:

-We gift associate attribute-based storage system

-We propose associate approach supported 2 science primitives, equally as a zero-knowledge proof of knowledge and a commitment theme, to comprehend knowledge consistency at intervals the system.

-Time primarily based all and access policy is given by original owner of file social unit transfer the info

ALGORITHMS

(1)**SHA-1:**(Secure Hash Algorithm) is a cryptographic hash function which takes an input and produces a 160-bit (20-byte) hash value known as a message digest – typically rendered as a hexadecimal number, 40 digits long.
This is designed to be computationally infeasible to:
a) Obtain the original message , given its message digest.
b) Find two messages producing the same message digest.
Each round takes 3 inputs-
• 512-bit block,
• The register abcde
• A constant K[t] (where t= 0 to 79)

2) **AES** (Advanced Encryption Standard):

The more popular and widely adopted symmetric encryption algorithm likely to be encountered nowadays is the Advanced Encryption Standard (AES).
1.Derive the set of round keys from the cipher key
2.Initialize the state array with the block data (plaintext).
3.Add the initial round key to the starting state array.
4.Perform nine rounds of state manipulation.
5.Perform the tenth and final round of state manipulation
6.Copy the final state array out as the encrypted data

AES Pseudocode:

```
Cipher(byte in[16], byte out[16], key_array round_key[Nr+1])
begin
{
byte state[16];
state = in;
AddRoundKey(state, round_key[0]);
for( i = 1 to Nr-1 stepsize 1) do
{
SubBytes(state);
ShiftRows(state);
MixColumns(state);
AddRoundKey(state, round_key[i]);
end for
SubBytes(state);
ShiftRows(state);
AddRoundKey(state, round_key[Nr]);
end
}
}
```


PROPOSE SYSTEM ARCHITECTURE

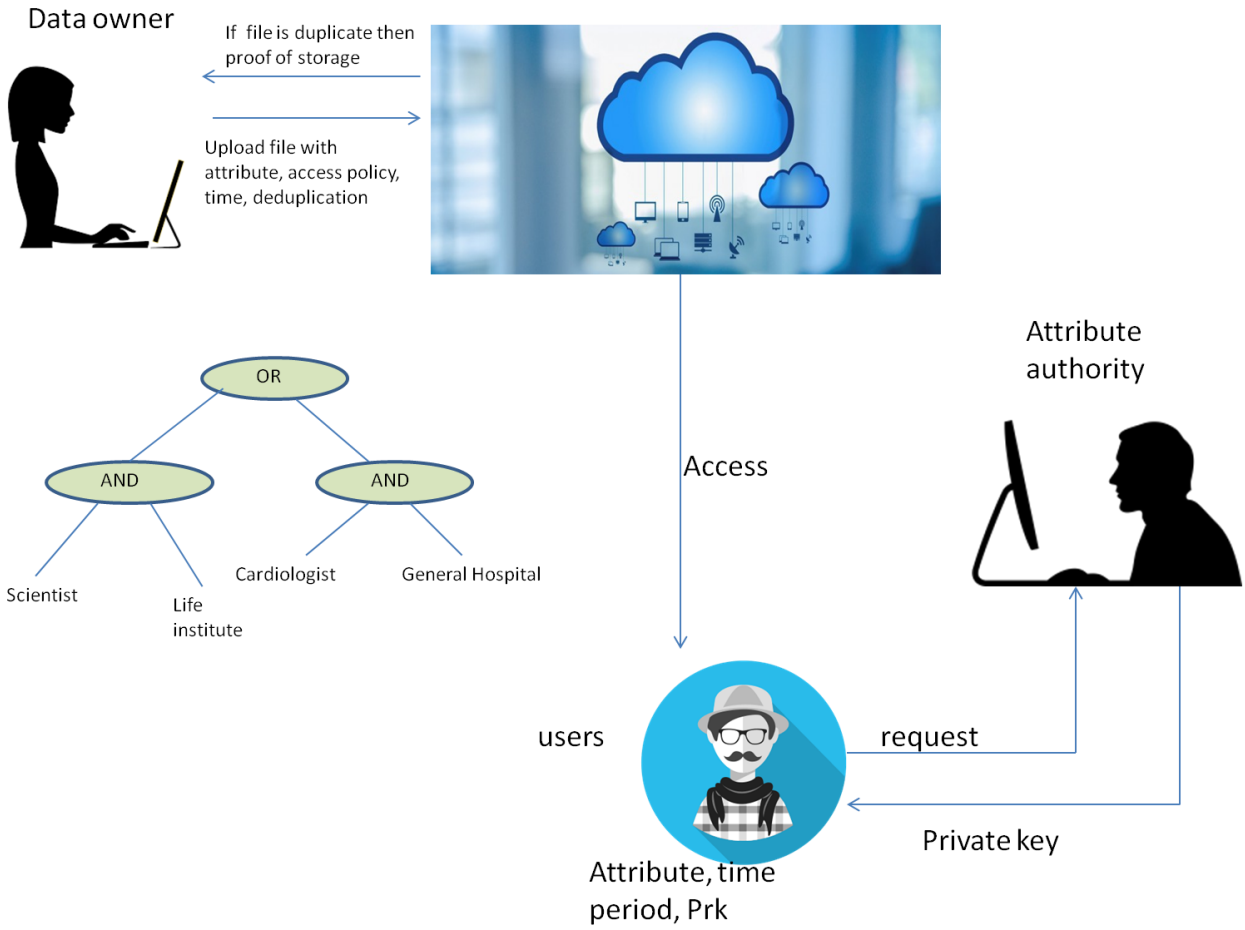
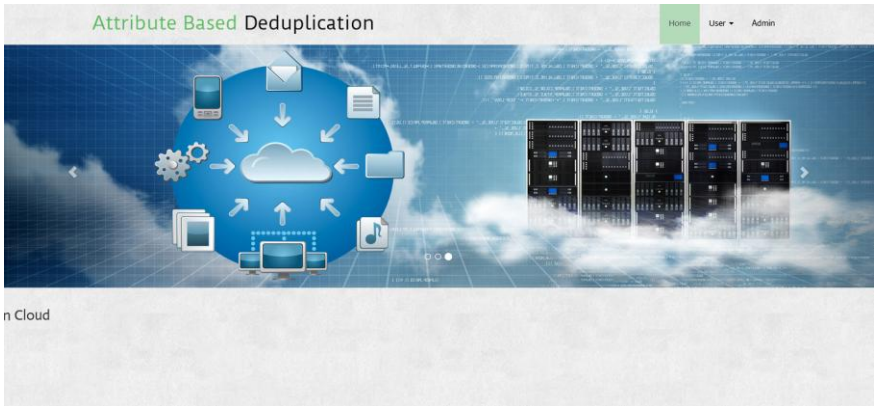


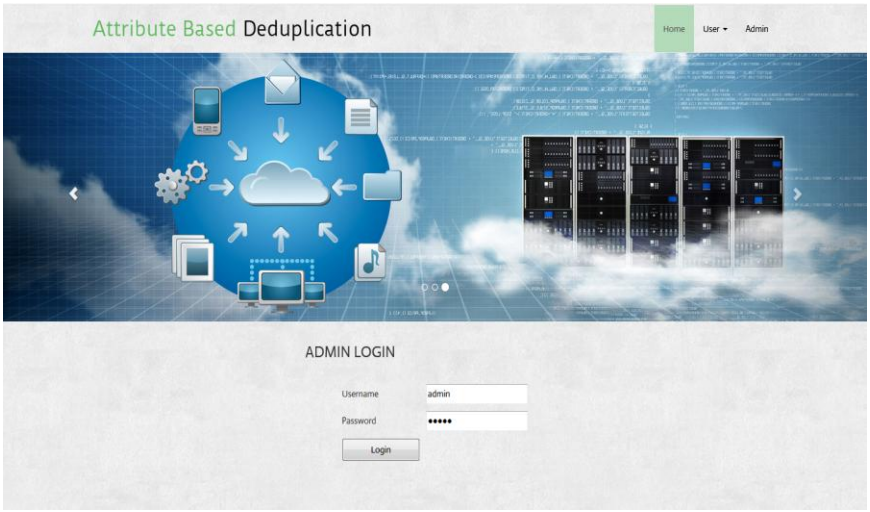
Figure 4.1. System Architecture

The architecture of our attribute-based storage system with secure deduplication is shown in Figure in which four entities are involved: data owner, attribute authority (AA), cloud and users. A data owner wants to outsource his/her data to the cloud and share it with users possessing certain credentials. The AA issues every user a decryption key associated with his/her set of attributes. The cloud consists of a public cloud which is in charge of data storage and a private cloud which performs certain computation such as tag checking. When sending a file storage request, each data provider firstly upload file with the attribute,access policy, time, and then encrypts the data under an access structure over a set of attributes.

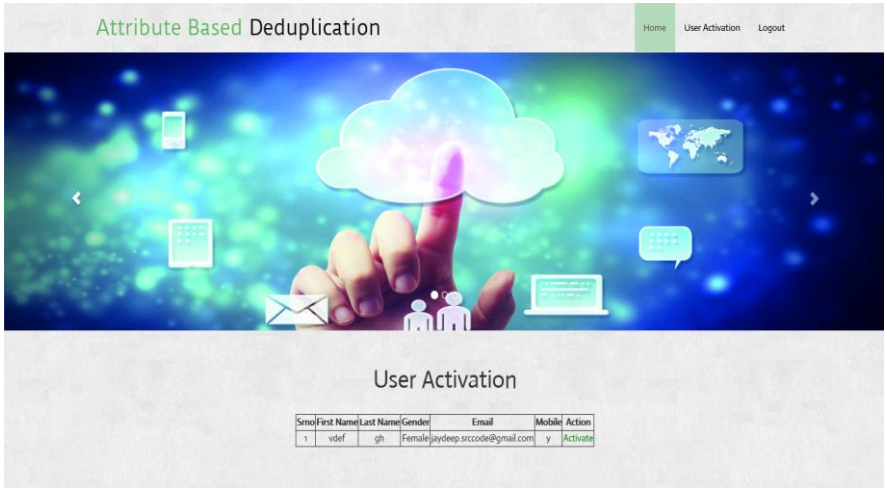
RESULTS



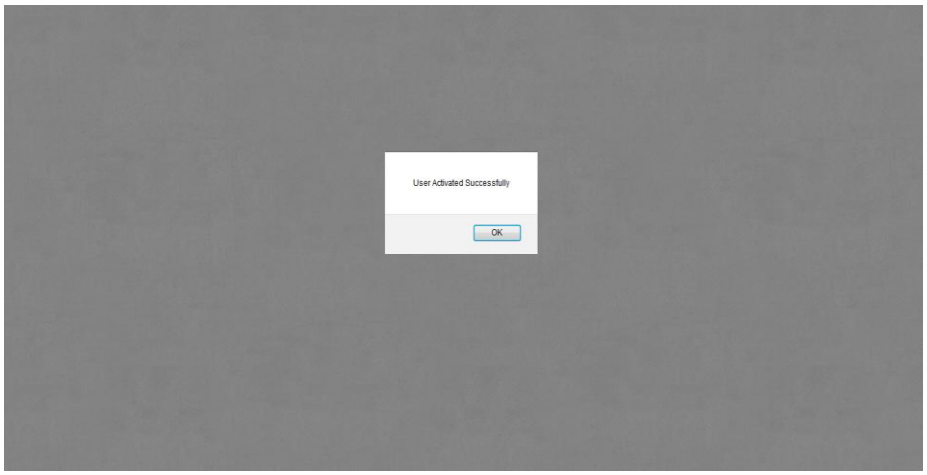
Result 1



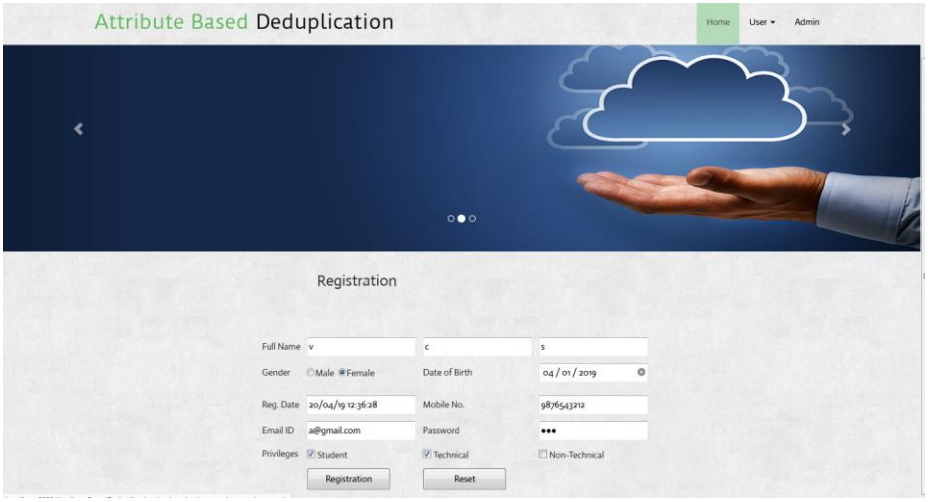
Result 2:Admin login into the system with email id and password.



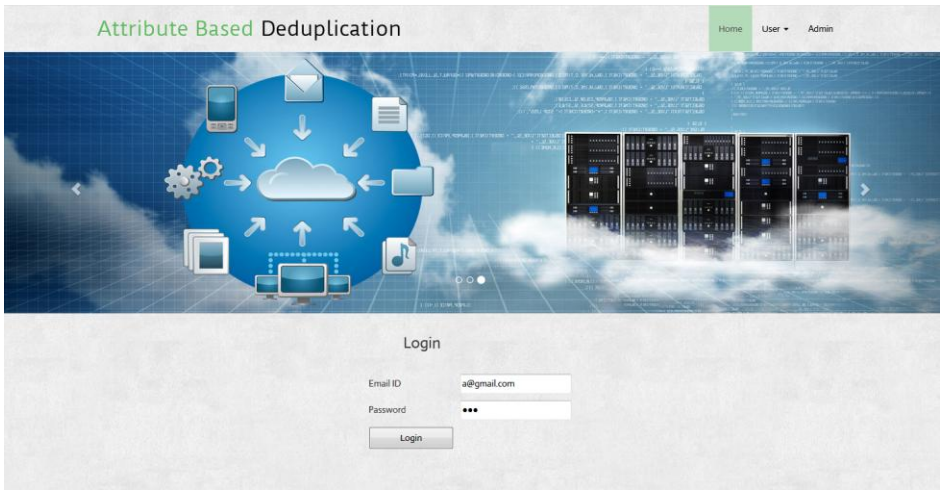
Result 3:In user activation admin activate the user



Result 4:Show pop up message when user successfully activated.



Result 5:Registration page for the new user.New user fill up the registration form.

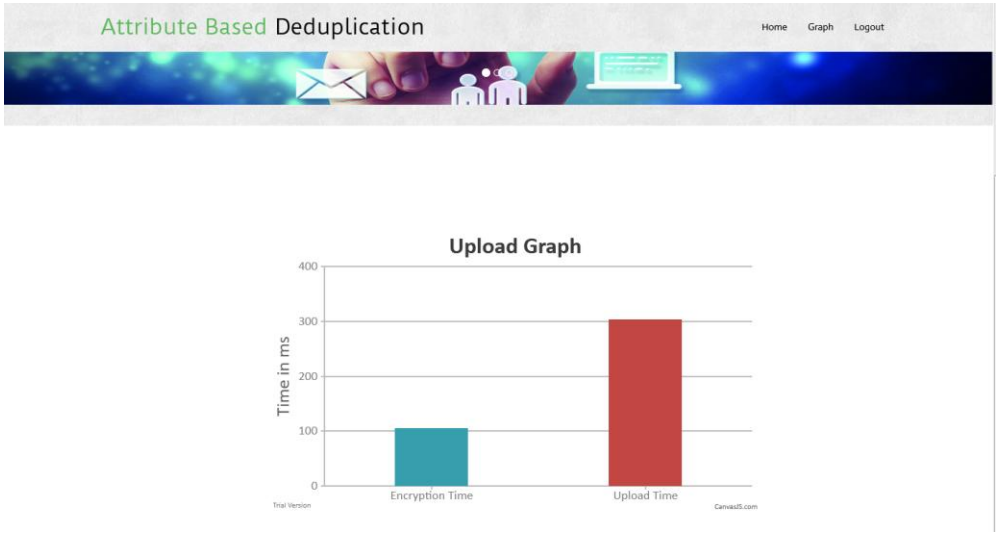


Result 6:User login into the application with registered email id and password.

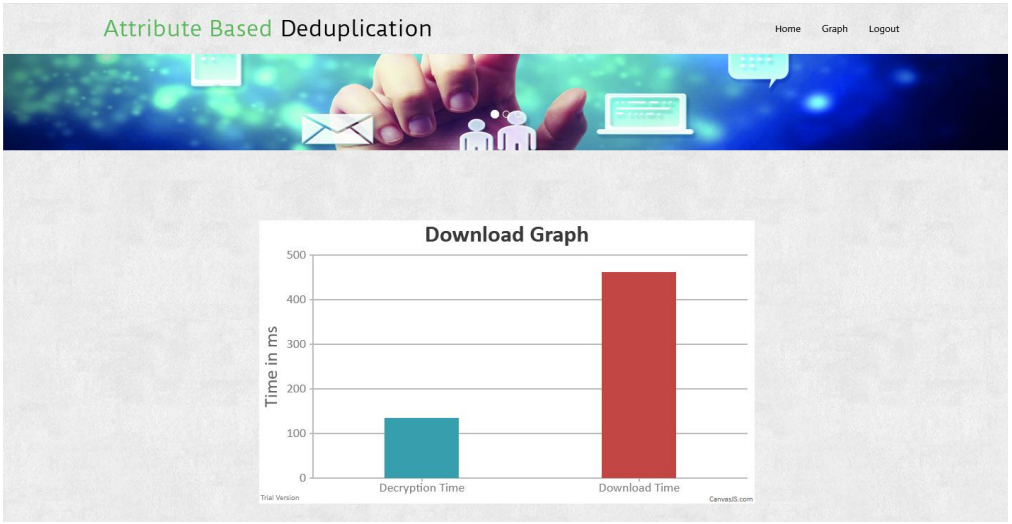
Result 7:User upload file with attributes for access file the uploaded file.

Result 8:User can download the file from cloud.

Result 9:Show the upload graph and download graph of uploaded file.



Result 10:Show the upload graph between upload time and encryption time of file.



Result 11:Show the download graph between download time and decryption time for file.

Result 12:show all details of user.

CONCLUSION AND FUTURE SCOPE

We planned a public auditing theme for encrypted information that may accomplish information integrity auditing and storage Diamond State duplication at identical time. By utilizing the thought of proxy re-encryption, the cloud server entirely ought to store one copy of encrypted information. To Diamond State duplicate the verification tags generated by whole absolutely utterly totally different venders, we have a tendency to tend to tend to tend to mixture the tags. The integrity of Diamond State duplicated information ar on the brink of be properly checked by the checker on behalf of any venders. The inquiry and trial results show that our theme is secure and economical.
The planned storage system enjoys a mix of major blessings. Firstly, it's on the brink of be accustomed confidentially share information with absolutely utterly totally different users by specifying award access policy rather than sharing the key writing key. Secondly, it accomplishs the standard notion of linguistics security whereas existing deduplication schemes entirely win it at a lower place a weaker security notion.

The standard CP-ABE(CIPHER-PLAINTEXT ATTRIBUTE primarily based whole ENCRYPTION) systems do not support secure deduplication, that produces them valuable to be applied in some business storage services.

REFERENCES

[1] D. Quick, B. Martini, and K. R. Choo, Cloud Storage Forensics. Syngress Publishing / Elsevier, 2014. [Online]. Available: http://www.elsevier.com/books/cloud-storageforensics/quick/978-0-12-419970-5

[2] K. R. Choo, J. Domingo-Ferrer, and L. Zhang, "Cloud cryptography: Theory, practice and future research directions," Future Generation Comp. Syst., vol. 62, pp. 51–53, 2016.

[3] K. R. Choo, M. Herman, M. Iorga, and B. Martini, "Cloud forensics: State-of-the-art and future directions," Digital Investigation, vol. 18, pp. 77–78, 2016.

[4] Y. Yang, H. Zhu, H. Lu, J.Weng, Y. Zhang, and K. R. Choo, "Cloud based data sharing with fine-grained proxy re-encryption," Pervasive and Mobile Computing, vol. 28, pp. 122–134, 2016.

[5] D. Quick and K. R. Choo, "Google drive: Forensic analysis of data remnants," J. Network and Computer Applications, vol. 40, pp. 179– 193, 2014.

[6] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings, ser. Lecture Notes in Computer Science, vol. 3494. Springer, 2005, pp. 457–473.

[7] B. Zhu, K. Li, and R. H. Patterson, "Avoiding the disk bottleneck in the data domain deduplication file system," in 6th USENIX Conference on File and Storage Technologies, FAST 2008, February 26-29, 2008, San Jose, CA, USA. USENIX, 2008, pp. 269–282.

[8] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings, ser. Lecture Notes in Computer Science, vol. 7881. Springer, 2013, pp. 296–312.

[9] M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G. Segev, "Message-locked encryption for lock-dependent messages," in Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I, ser. Lecture Notes in Computer Science, vol. 8042. Springer, 2013, pp. 374–391.

[10] S. Keelveedhi, M. Bellare, and T. Ristenpart, "Dupless: Serveraided encryption for deduplicated storage," in Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013. USENIX Association, 2013, pp. 179–194.

[11] M. Bellare and S. Keelveedhi, "Interactive message-locked encryption and secure deduplication," in Public-Key Cryptography – PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 – April 1, 2015, Proceedings, ser. Lecture Notes in Computer Science, vol.9020. Springer, 2015, pp. 516–538.

[12] S. Bugiel, S. N¨ urnberger, A. Sadeghi, and T. Schneider, "Twin clouds: Secure cloud computing with low latency - (full version)," in Communications and Multimedia Security, 12th IFIP TC 6 / TC 11 International Conference, CMS 2011, Ghent, Belgium, October 19- 21,2011. Proceedings, ser. Lecture Notes in Computer Science, vol. 7025. Springer, 2011, pp. 32–44.

[13] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof-systems (extended abstract)," in Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA. ACM, 1985, pp. 291– 304.

[14] M. Fischlin and R. Fischlin, "Efficient non-malleable commitment schemes," in Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings, ser. Lecture Notes in Computer Science, vol. 1880. Springer, 2000, pp. 413–431.

[15] S. Goldwasser and S. Micali, "Probabilistic encryption," J. Comput. Syst. Sci., vol. 28, no. 2, pp. 270–299, 1984.

