Layer-Based Encrypted Deduplication for Secure Cloud Data Storage

Mrm.Dharani Kumar[1], Karthanaparthi Shalini[2]
1Assistant Professor Dept of C.S.E, PVKK Institute of Technology Anantapur, Andhra Pradesh-515001.
2PG Scholar, Dept of C.S.E, PVKK Institute of Technology Anantapur, Andhra Pradesh- 515001.

Abstract:

LSDedup gives a stratified secure deduplication approach for cloud storage, safeguarding information confidentiality while enhancing garage performance. conventional deduplication techniques necessitate get right of entry to to unencrypted information, which gives privacy issues, whereas LSDedup facilitates deduplication on encrypted files without compromising sensitive information. files are labeled into Strictly exclusive (SCFiles) and much less confidential (LSFiles) based totally on their sensitivity ranges. SCFiles are encrypted with the advanced Encryption standard (AES) for better security, whereas LSFiles employ Convergent Encryption (CE), allowing the deduplication of identical data without the want for decryption. A proof of work (PoW) mechanism is implemented to address the misrepresentation of record secrecy in cloud servers, as a result making sure transparency and safety. A safe keyword-based search method is created to facilitate efficient file retrieval while maintaining statistics privacy. This approach produces encrypted trapdoors for key-word-centric searches, obstructing illegal access to search queries and file contents. The counseled system optimizes storage optimization via deduplication and gives a comprehensive approach for secure data control and retrieval in cloud environments. LSDedup achieves an most beneficial equilibrium among security and efficiency in cloud storage systems through the integration of encryption, proof of work validation, and safe search algorithms.

***"Index Terms** -Cloud storage, encrypted data deduplication, layered deduplication, secure deduplication, cloud storage security".*

1. INTRODUCTION

The extensive utilization of cloud storage offerings has ended in a great surge in statistics quantity, requiring powerful garage control techniques to cope with redundant information. a major trouble in cloud storage is the redundant uploading of comparable information via numerous users, leading to sizable storage inefficiency and heightened operational charges for provider vendors [1]. To address this problem, data deduplication has become an essential technology that removes duplicate copies of saved records, assuring the retention of most effective unique documents or data blocks [2]. Deduplication complements system efficiency by means of optimizing save area, for this reason reducing garage and bandwidth costs. Conventional deduplication strategies depend on access to unencrypted statistics, which introduces security vulnerabilities and risks to user privacy [3].

Conventional deduplication methods function via "both file-stage and block-level deduplication". File-degree deduplication keeps an unmarried example of an same file, while block-level deduplication divides a document into smaller elements and eliminates replica blocks to enhance storage performance. Although these solutions drastically diminish redundancy, they necessitate access to unencrypted statistics, thereby doubtlessly exposing touchy consumer statistics to cloud companies and heightening the danger of illegal access. Encryption is substantially employed to shield stored statistics in response to this problem. Encrypting statistics prior to deduplication contradicts standard deduplication methods, as encryption alters equal files into particular ciphertexts, so obstructing conventional redundancy reduction.

To facilitate secure deduplication, academics have suggested encryption methodologies like "Convergent Encryption (CE)", which ensures that identical plaintexts yield comparable ciphertexts, hence allowing deduplication on encrypted documents. Notwithstanding its benefits, CE-based deduplication is vulnerable to frequency analysis assaults, wherein adversaries can deduce stored facts from the recurrence of ciphertexts [6]. A hybrid method that classifies information in step with secrecy stages has been proposed to enhance security. Exceptionally non-public information is encrypted with "advanced Encryption standard (AES), which precludes deduplication, while less sensitive facts is protected by using CE, allowing for deduplication" even as preserving confidentiality [7].

Furthermore, "proof-of-work (PoW) techniques" had been implemented to discourage cloud provider providers from misrepresenting the confidentiality level of a record. through the integration of cryptographic demanding situations, "proof of work (PoW)" ensures transparency in document class and enhances protection in opposition to nefarious alteration. moreover, at ease key-word-based search abilities enable customers to question encrypted documents without the want for decryption, subsequently enhancing usability while maintaining information privateness [8]. This solution guarantees a super equilibrium among protection and performance, overcoming the shortcomings of conventional deduplication strategies and enhancing at ease records management in cloud settings.

2. RELATED WORK

Cloud storage security has been thoroughly examined, emphasizing secure records deduplication strategies to reduce storage redundancy even as preserving secrecy. Conventional deduplication strategies characteristic on the document or block degree, detecting and putting off redundant facts prior to storage. Nevertheless, those answers commonly necessitate access to unencrypted statistics, which raises issues over information privateness and the chance of unlawful access. Encryption-based deduplication solutions had been evolved to provide security and efficiency in cloud environments to address this trouble [9].

"Message-Locked Encryption (MLE)" has emerge as a distinguished approach for safe deduplication, wherein identical plaintext files yield equal ciphertexts, facilitating redundancy removal whilst maintaining anonymity. initial research investigated the viability of MLE in cloud storage structures, revealing its capacity for safe and space-efficient deduplication. though, MLE is at risk of assaults like frequency evaluation, wherein adversaries can deduce saved facts by inspecting ciphertext styles [10]. To address this weakness, numerous modifications to MLE have been cautioned, which include randomized encryption algorithms and hybrid encryption schemes that combine MLE with traditional encryption methods to better safety whilst preserving deduplication efficiency [11].

"Proof-of-ownership (PoW)" strategies had been carried out to enhance protection in encrypted deduplication. those structures allow users to affirm their possession of a document without revealing its contents, thereby thwarting unwanted get right of entry to by means of antagonistic individuals seeking to take advantage of deduplication mechanisms. Evidence-of-work-based deduplication methods make use of cryptographic demanding situations to assure that best authorized customers can assert ownership of saved data. Nonetheless, certain research suggest that PoW approaches impose computational overhead, as a result impacting system overall performance in sizeable cloud systems [12]. To tackle this issue, researchers have investigated lightweight evidence of work systems that reconcile security and efficiency by way of diminishing computing complexity while maintaining strong authentication assurances [13].

Another technique for making sure secure deduplication employs threshold-primarily based encryption techniques, allowing deduplication by myself while a designated number of users upload same information. This approach mitigates the risk of frequency evaluation attacks by ensuring that a record is encrypted uniquely unless numerous customers have access to the identical content. Thru the implementation of threshold encryption, cloud companies can alleviate privacy troubles even as keeping the blessings of deduplication. This method necessitates more coordination amongst cloud users, ensuing in extra complexity in substantial deployments [14].

to augment safety, certain research have counseled deduplication-earlier than-encryption techniques, wherein data is first of all deduplicated and eventually encrypted with distinct encryption keys. This technique minimizes storage overhead while safeguarding information confidentiality. Despite the fact that, it raises troubles regarding the exposing of brief plaintext data previous to encryption, potentially developing protection flaws. To remedy this problem, researchers have included secure hardware enclaves to permit truthful deduplication-before-encryption, thereby lowering the threat of statistics leaking even as keeping efficiency [15].

Encrypted deduplication offers issues in metadata management, as it's far complicated to build indexes for encrypted files at the same time as assuring secure get admission to control. Some of research have investigated privateness-maintaining metadata control techniques, inclusive of searchable encryption and oblivious facts structures, to facilitate at ease deduplication with effective query competencies. These methodologies allow users to behavior keyword-based totally searches on encrypted deduplicated files even as safeguarding touchy facts. Despite the fact that effective, those strategies regularly encompass supplementary computational fees, affecting system scalability [16].

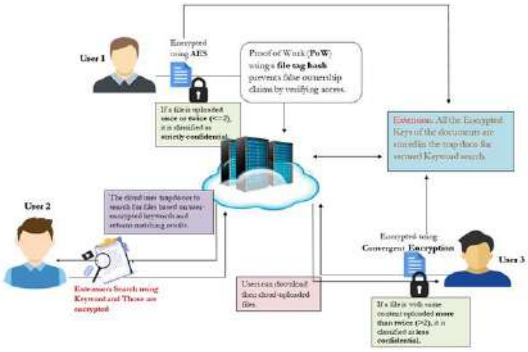
Alongside protection issues, effective storage control is still a primary emphasis in encrypted deduplication research. Research has examined most appropriate storage architectures making use of hierarchical deduplication methods, with primary deduplication completed at the purchaser facet previous to records upload to the cloud. This method optimizes network bandwidth use and diminishes computing burden on cloud servers. Customer-facet deduplication necessitates massive processing electricity on user devices, rendering it less suitable for resource-confined environments [17]. To mitigate this constraint, hybrid deduplication frameworks were hooked up, integrating consumer-aspect and server-facet deduplication to beautify overall performance in step with fluctuating workload situations. Those frameworks make use of adaptive algorithms to check the only deduplication technique for a particular dataset, optimizing storage performance and computational expense [18].

A nascent studies avenue includes the combination of blockchain generation with secure deduplication. Blockchain gives a decentralized and immutable report for monitoring deduplicated documents, subsequently enhancing transparency and safety. Utilising clever contracts, blockchain-primarily based deduplication solutions assure that only authorized users can get admission to and control the stored statistics. Furthermore, blockchain obviates the need for a centralized authority, for this reason diminishing the probability of insider threats. Nevertheless, blockchain-primarily based deduplication introduces latency as a result of consensus procedures, requiring improvements to decorate real-time performance [19].

Future developments in encrypted deduplication aim to tackle scalability issues at the same time as retaining robust safety assurances. Machine learning methodologies were investigated to enhance deduplication efficacy by way of forecasting redundancy patterns in considerable datasets. these methodologies make use of deep studying models to categorize documents according to their chance of duplication, facilitating enhanced storage control. moreover, investigations into publish-quantum cryptography are seeking to establish encryption methodologies that are impervious to quantum assaults, so safeguarding the long-lasting safety of encrypted deduplication structures. As cloud storage advances, at ease deduplication strategies can be essential in optimizing records security, garage efficiency, and computing overall performance in present day cloud settings [20].

3. MATERIALS AND METHODS

LSDedup improves secure deduplication in cloud storage by facilitating deduplication of encrypted files without the want for decryption, subsequently maintaining security and performance. It classifies information according to their confidentiality, making use of AES encryption for particularly exclusive files and “Convergent Encryption (CE)” for files of lesser confidentiality, so optimizing storage whilst preserving privateness. A “proof of work (PoW)” mechanism is employed to authenticate confidentiality tiers, thwarting deception via cloud provider carriers. LSDedup further enhances its capabilities with at ease Cloud search, permitting encrypted key-word-primarily based searches at the same time as concealing record content material. This characteristic creates a trapdoor with encrypted key phrases, facilitating green report retrieval even as retaining confidentiality. by means of safeguarding both stored information and search queries, it improves metadata control and guarantees integrity. The dual encryption method, collectively with PoW verification and privateness-keeping search functionalities, efficiently tackles cloud protection problems, balancing deduplication performance with robust encryption techniques.


User 1
Encrypted using AES
Proof of Work (PoW) using a file tag hash prevents false ownership claim by verifying access.
If a file is uploaded once or twice (<=2), it is classified as strictly confidential
Extension: All the Encrypted Keys of the documents are stored in the trap door for secured Keyword search.
User 2
The cloud uses trapdoor to search for files based on user-encrypted keywords and returns matching results.
Extension: Search using Keyword and Those are encrypted
Users can download their cloud-uploaded files.
Encrypted using Convergent Encryption
User 3
If a file is with same content uploaded more than twice (>2), it is classified as less confidential.

“Fig.1 Proposed Architecture”

This diagram (Fig.1) depicts a relaxed cloud storage system prepared with privacy protections. users add encrypted files, specified as "strictly personal" if uploaded two times or fewer, and "less personal" if submitted extra than twice. making use of record tag hashes in "evidence of labor" mitigates fraudulent ownership assertions. AES and Convergent Encryption are employed for protection purposes. Trapdoors facilitate key-word searches within encrypted statistics. users may additionally download their documents, and encrypted keys are properly saved for retrieval. the answer integrates strong encryption with looking abilities, safeguarding records privacy and integrity.

i) Data Classification and Encryption Mechanisms

LSDedup classifies documents in keeping with secrecy tiers to implement suitable encryption techniques, consequently assuring security and storage efficiency. Files detailed as strictly personal are subjected to “advanced Encryption standard (AES) encryption”, ensuring strong protection via symmetric key cryptography. Conversely, much less personal files make use of “Convergent Encryption (CE), facilitating deduplication” even as retaining anonymity. CE produces encryption keys based on document content material, permitting same documents to offer the same ciphertext for optimized storage. This bifurcated strategy bolsters protection while maximizing storage performance. The classification method safeguards towards undesirable get entry to via ensuring the utilization of suitable encryption techniques. This method harmonizes confidentiality and deduplication performance, reducing the chance of information publicity even as preserving cloud storage resources.

ii) Secure Deduplication Process

LSDedup utilizes a cozy deduplication technique that eliminates unnecessary files while maintaining data security. A cryptographic hash set of rules is utilized to perceive replica files and to save you unwanted get right of entry to. To prevent the publicity of exclusive information at some stage in deduplication, “proof of work (PoW)” is employed to evaluate the secrecy level of files previous to the deduplication manner. This method inhibits cloud service companies from misrepresenting files to make the most garage efficiencies. Deduplication is implemented on each the customer and server aspects, optimizing storage performance. The method preserves data integrity at the same time as minimizing garage prices, rendering LSDedup a scalable answer for secure cloud storage. LSDedup integrates encryption with deduplication to safeguard statistics while improving storage performance.

iii) Secure Cloud Search Implementation

LSDedup complements its competencies through integrating secure Cloud seek, enabling customers to find files thru encrypted keyword-primarily based searches. The system develops a trapdoor feature that encrypts search key phrases, so keeping confidentiality as an alternative of exposing plaintext searches. The trapdoor is transmitted to the cloud server, which correlates the encrypted question with indexed information without decrypting the stored files. The method safeguards privacy by way of obstructing query leaks and making sure that the cloud provider can't deduce sensitive seek styles. Moreover, secure search techniques facilitate fuzzy matching, allowing customers to access data despite minor key-word discrepancies. This plugin improves LSDedup's usability with the aid of permitting secure, efficient, and privateness-maintaining file retrieval in cloud settings.

iv)Storage Optimization and Performance Evaluation

LSDedup enhances storage efficiency by integrating encryption-based deduplication with computationally efficient techniques. The method ensures the removal of unnecessary files without jeopardizing security, resulting in vast garage savings. Overall performance is classified using measures like deduplication ratio, encryption overhead, and retrieval put off. The generation reduces computational costs by using enhancing encryption and deduplication procedures. Moreover, protection studies evaluates LSDedup's resilience to data leakage, illegal access, and ciphertext attacks. UsingPoW verification guarantees the prevention of data misrepresentation, hence safeguarding against the exploitation of deduplication techniques. LSDedup attains an ideal equilibrium among security, efficiency, and performance, rendering it a powerful choice for safe cloud storage.

v) Modules:

New User Signup: This module permits new users to set up an account in the system. Customers input their information “(including username, email, and password)” into a registration form. Upon submission of the information, the device statistics the brand new user in the database, granting them access to the application.

User Login:This module permits user authentication for system access. users authenticate by inputting their credentials (username and password). The system authenticates these credentials with the database. Upon successful authentication, users obtain access to their personal dashboard and file management functionalities.

Outsource File:This module oversees the uploading and categorization of files. users transfer files to the cloud. The system creates a distinct document Tag for every uploaded record and verifies the database for pre-existing files with analogous tags to assess duplication. It determines the proper encryption set of rules based totally on the quantity of duplicates and a predetermined threshold (“AES for Strictly private files or Convergent Encryption for less exclusive documents”). The encryption specifications and tags of the document are eventually archived within the cloud database. The software program also collects and encrypts key phrases from the report to establish a secure trapdoor for subsequent searches.

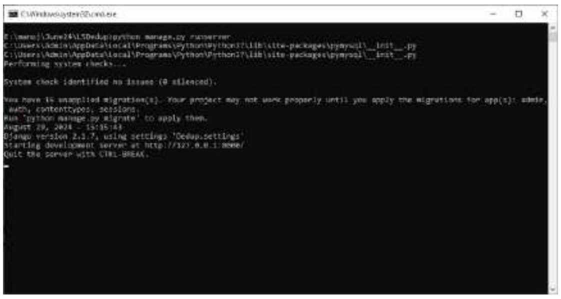
View Files:This module permits users to view and administer their uploaded files. Users can view a comprehensive listing of all files they have uploaded to the cloud. They are able to decrypting and downloading any of those documents as required. This module assists users in monitoring their data and regulating file access.

Secured Cloud Search:This module lets in customers to search for documents with encrypted key phrases. Customers enter keywords into a seek form. The key phrases are encrypted and transmitted to the cloud. The cloud executes the hunt via the encrypted trapdoor and affords a list of documents that include the search keywords. This guarantees that each the hunt process and the results uphold facts confidentiality.

Storage Graph:This module offers a graphical depiction of storage expenses. It produces a graph illustrating the storage expenses linked to the proposed LSDedup method in evaluation to modern deduplication methods. The x-axis denotes the amount of documents uploaded, whilst the y-axis suggests the related storage cost. The graph allows users to assess the efficiency and value-effectiveness of the counseled answer in contrast to traditional procedures.

4. RESULTS AND DISCUSSION

To execute the project, copy all material from the 'Database.txt' file and paste it into the MySQL console to construct the database. finally, double-click the 'runServer.bat' document to initiate the server and access the following page.



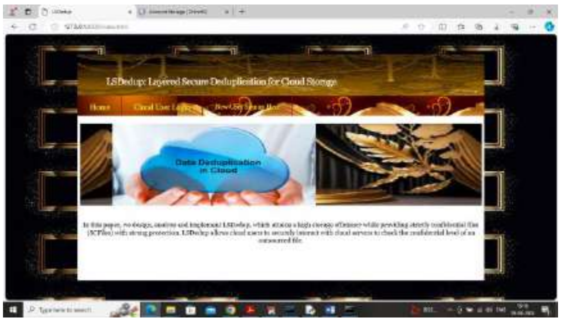
```
C:\Windows\system32\cmd.exe

E:\manoj\June24\LSDedup>python manage.py runserver
C:\Users\Admin\AppData\Local\Programs\Python\Python37\lib\site-packages\pymysql\__init__.py
C:\Users\Admin\AppData\Local\Programs\Python\Python37\lib\site-packages\pymysql\__init__.py
Performing system checks...

System check identified no issues (0 silenced).

You have 15 unapplied migration(s). Your project may not work properly until you apply the migrations for app(s): admin,
 auth, contenttypes, sessions.
Run 'python manage.py migrate' to apply them.
August 29, 2024 - 15:35:43
Django version 2.1.7, using settings 'Dedup.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CTRL-BREAK.
```

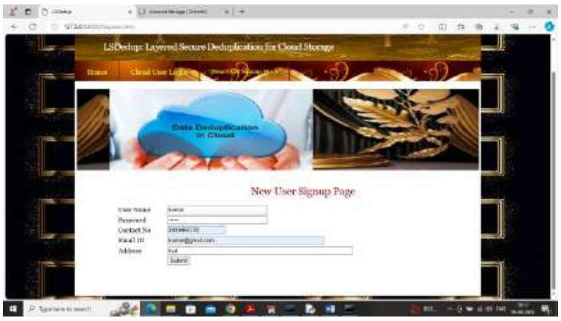
The Python server has commenced operation; now, open a browser and enter the "URL http://127.zero.0.1:8000/index.html", then click the enter key to display the subsequent page.



LSDedup: Layered Secure Deduplication for Cloud Storage

In this paper, we design, analyse and implement LSDedup, which attains a high storage efficiency while providing strictly confidential files (SCFiles) with strong protection. LSDedup allows cloud users to securely interact with cloud servers to check the confidential level of an outsourced file.

click on on the "New user sign up" alternative on the higher display screen to access the following web page.

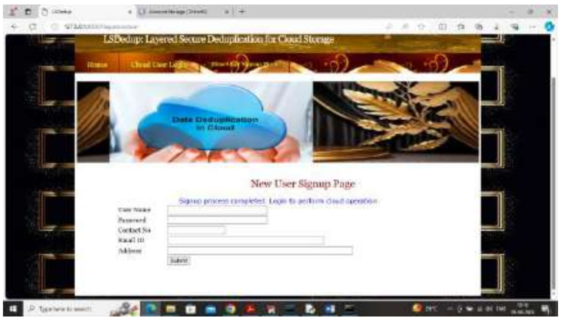


LSDedup: Layered Secure Deduplication for Cloud Storage

New User Signup Page

User Name	kumar
Password	*****
Contact No	[illegible]
Email ID	kumar@gmail.com
Address	hyd

The person inputs sign-up information on the top display and in the end presses the button to get right of entry to the page below.



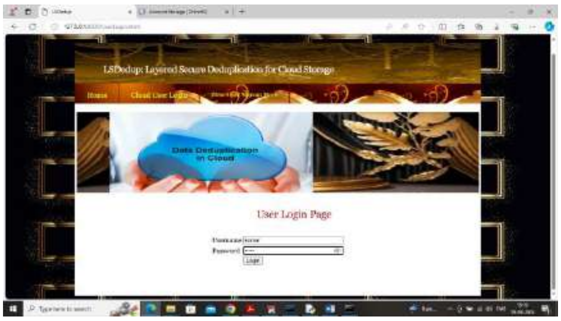
LSDedup: Layered Secure Deduplication for Cloud Storage

New User Signup Page

Signup process completed. Login to perform cloud operation

User Name	
Password	
Contact No	
Email ID	
Address	

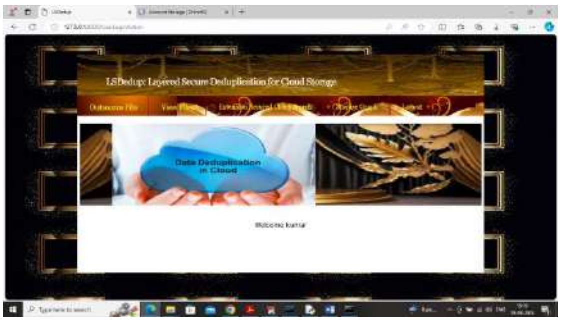
The user has completed the sign-up system and have to now click on at the "Cloud person Login" link to get admission to the following login web page.


LSDedup: Layered Secure Deduplication for Cloud Storage
Home
Cloud User Login
Data Deduplication in Cloud

User Login Page

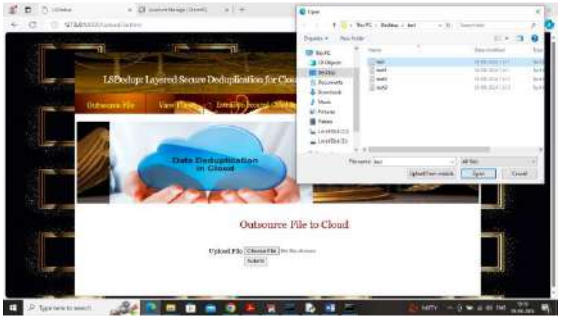
Username
Password
Login

Upon logging in, the user will be directed to the subsequent page displayed above.


LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Data Deduplication in Cloud
Logout

Welcome kumar

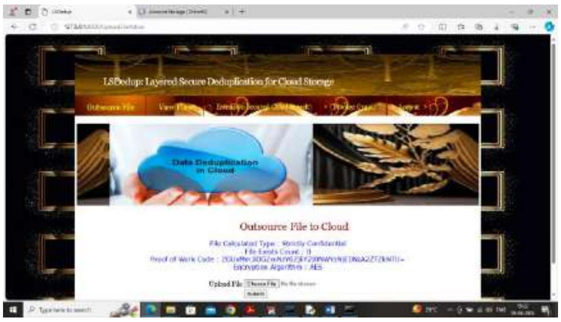
Click at the “Outsource record” hyperlink at the pinnacle screen to get admission to the following web page.


LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Data Deduplication in Cloud
Open
Open
Cancel

Outsource File to Cloud

Upload File: Choose File No file chosen
Submit

Upload the take a look at.txt record on the above screen and then press the button to get admission to the following web page.

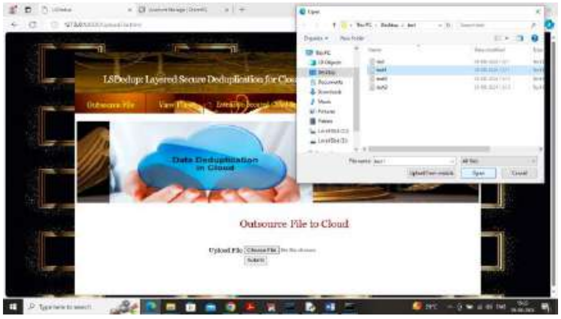

LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Data Deduplication in Cloud
Logout

Outsource File to Cloud

File Calculated Type : Strictly Confidential
File Exists Count : 0
Proof of Work Code : [illegible]
Encryption Algorithm : AES

Upload File: Choose File No file chosen
Submit

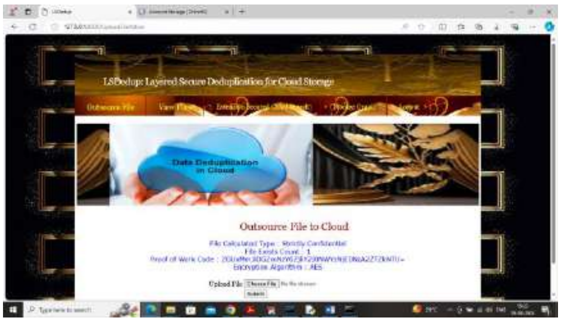
The display displays that record confidentiality is assessed as 'Strict' with a count number of zero. Moreover, it suggests the proof of labour file Tag and the chosen encryption algorithm as AES. Every other document with same content is currently being uploaded.


LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Outsource File to Cloud

Upload File: Choose File No file chosen
Submit

The identical report is being uploaded at the pinnacle screen, and the output is displayed below.

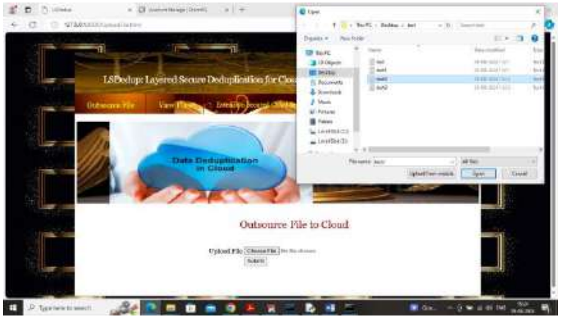

LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Outsource File to Cloud

File Calculated Type : Strictly Confidential
File Exists Count : 1
Proof of Work Code : [illegible]
Encryption Algorithm : AES

Upload File: Choose File No file chosen
Submit

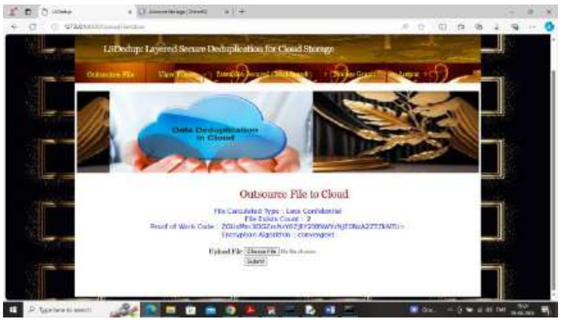
The record content be counted displayed above is 1, and the algorithm employed is AES. Importing a further record with equal content material.


LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Outsource File to Cloud

Upload File: Choose File No file chosen
Submit

Inside the above display, the third file is being uploaded, and the output is displayed under.

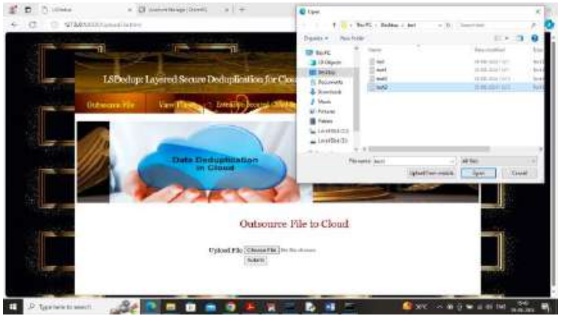

LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Outsource File to Cloud

File Calculated Type : Less Confidential
File Exists Count : 2
Proof of Work Code : [illegible]
Encryption Algorithm : convergent

Upload File: Choose File No file chosen
Submit

The content matter within the above screen climbed to 2, surpassing the edge, and the set of rules employed turned into “Convergent Encryption.”Importing the fourth document with distinct content material.

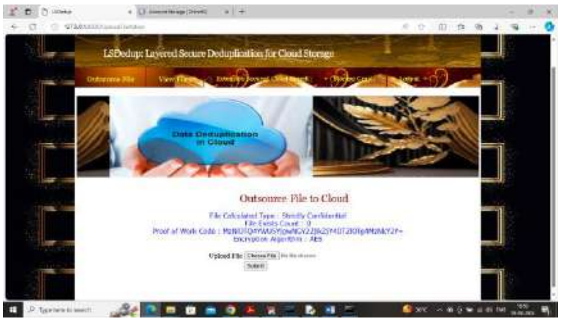

LSDedup: Layered Secure Deduplication for Clou
Outsource File
View Files
Data Deduplication in Cloud
Open
Open
Cancel

Outsource File to Cloud

Upload File: Choose File No file chosen

Submit

In the screen above, the fourth file is being uploaded, and the output is displayed underneath.


LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Data Deduplication in Cloud

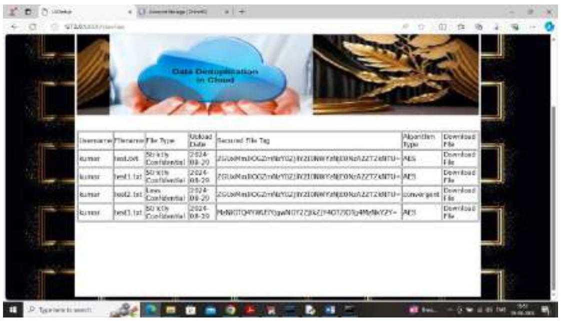
Outsource File to Cloud

File Calculated Type : Strictly Confidential
File Exists Count : 0
Proof of Work Code : MzNkOTQ4YWU5YjgwNGY2ZjlkZjY4OTZlOTg4MzNkY2Y=
Encryption Algorithm : AES

Upload File: Choose File No file chosen

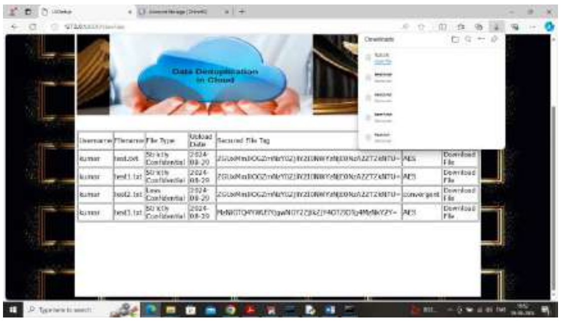
Submit

The fourth document at the aforementioned display has awesome fabric; as a result, the AES technique is employed, and the file is classed as Strict. Click on at the "View documents" hyperlink to get the list of documents.



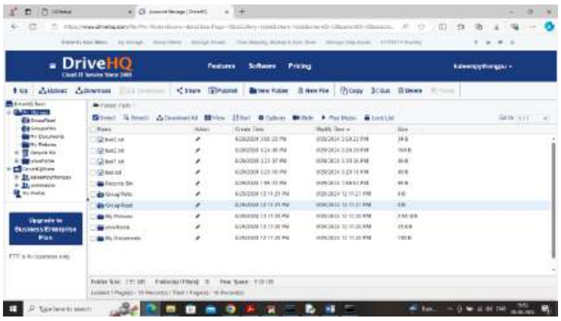
Username	Filename	File Type	Upload Date	Secured File Tag	Algorithm Type	Download File
kumar	test.txt	Strictly Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	AES	Download File
kumar	test1.txt	Strictly Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	AES	Download File
kumar	test2.txt	Less Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	convergent	Download File
kumar	test3.txt	Strictly Confidential	2024-08-29	MzNkOTQ4YWU5YjgwNGY2ZjlkZjY4OTZlOTg4MzNkY2Y=	AES	Download File

The above display presentations all uploaded files, accompanied through a record tag indicating same and varied content material. when duplicate content material is diagnosed, Convergent Encryption is employed; in any other case, AES is utilized. Click on on the 'down load document' link to provoke the down load.

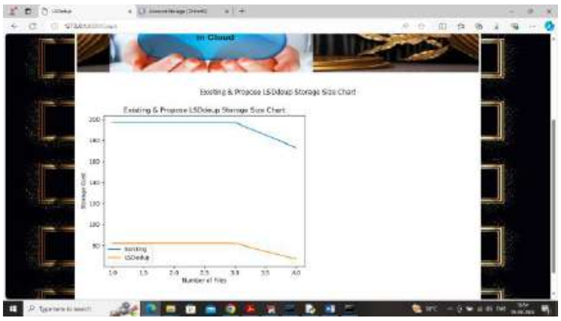


Username	Filename	File Type	Upload Date	Secured File Tag	Algorithm Type	Download File
kumar	test.txt	Strictly Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	AES	Download File
kumar	test1.txt	Strictly Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	AES	Download File
kumar	test2.txt	Less Confidential	2024-08-29	ZGUxMmI0OGZmNzYOZjIlY2I0NWYzNjE0NzA2ZTZkNTU=	convergent	Download File
kumar	test3.txt	Strictly Confidential	2024-08-29	MzNkOTQ4YWU5YjgwNGY2ZjlkZjY4OTZlOTg4MzNkY2Y=	AES	Download File

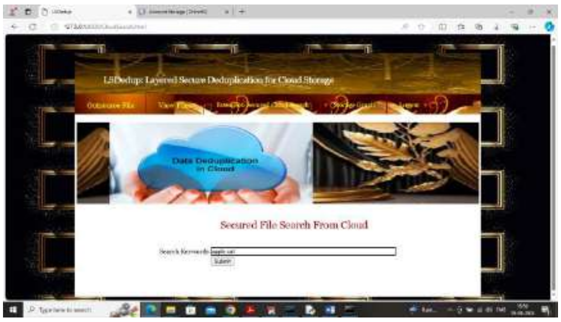
The downloaded indicator above indicates that the file has been downloaded with decrypted content, and all documents are seen in the "DriveHQ cloud" underneath.



in the preceding screen, all four documents named 'take a look at' have been uploaded to "DRIVEHQ". Now, click on at the 'storage Graph' hyperlink to obtain the graph displayed under.



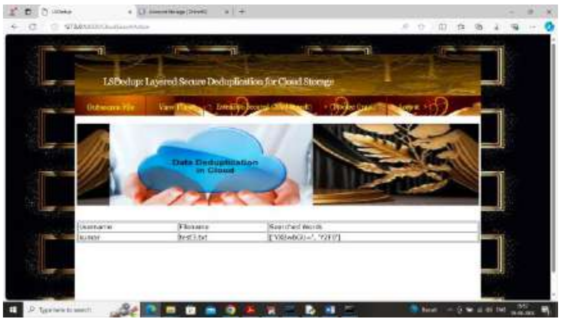
The x-axis of the graph denotes the wide variety of documents uploaded, at the same time as the y-axis suggests the garage price. The orange line illustrates the proposed "LSDEDUP method", and the blue line displays the present approach. The proposed answer requires less garage; now click on the "Extension relaxed Cloud seek" link to get the following page.


LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Extension Secured Cloud Search
Storage Graph
Logout
Data Deduplication in Cloud
Submit

Secured File Search From Cloud

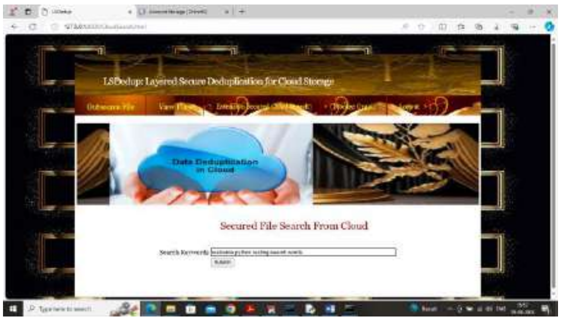
Search Keywords: apple.txt

Enter search keywords in the top box and then press the button to access the page below.


LSDedup: Layered Secure Deduplication for Cloud Storage
Outsource File
View Files
Extension Secured Cloud Search
Storage Graph
Logout
Data Deduplication in Cloud

Username	Filename	Searched Words
kumar	test1.txt	['YXBwbGU=', 'Y2F0']

The above screen presentations the names of searched files alongside encrypted phrases, and you may further look for any key phrases. below is every other instance.

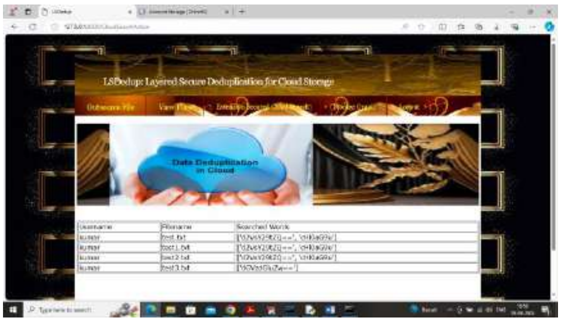

LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Secured File Search From Cloud

Search Keywords: welcome python testing secret words

Submit

The terms furnished within the above display screen are displayed inside the search output underneath.


LSDedup: Layered Secure Deduplication for Cloud Storage
Data Deduplication in Cloud

Username	Filename	Searched Words
kumar	test.txt	['d2VsY29tZQ==', 'cHl0aG9u']
kumar	test1.txt	['d2VsY29tZQ==', 'cHl0aG9u']
kumar	test2.txt	['d2VsY29tZQ==', 'cHl0aG9u']
kumar	test3.txt	['dGVzdGluZw==']

The above screen displays a list of files that contain the specified keywords.

Utilizing the aforementioned result, we have developed a Secured Cloud De-duplication system, incorporating secured cloud keywords that were searched.

5. CONCLUSION

In conclusion, the LSDedup task gives a complete framework for shielding user facts even as facilitating effective deduplication of encrypted files, thereby keeping confidentiality. The system enhances storage efficiency by classifying files as 'Strict' and 'less personal,' utilizing AES and Convergent Encryption to decrease redundancy and lower expenses. The implementation of a proof of work mechanism improves the veracity of confidentiality assertions, deterring cloud provider carriers from exaggerating record sensitivity. Users obtain more advantageous control over their file confidentiality and management, facilitated by clear verification procedures. Additionally, the project's extension consists of a at ease, encrypted keyword-based totally seek capability, allowing users to discover their documents without jeopardizing important information. LSDedup correctly tackles crucial problems in cloud storage security, imparting a holistic solution that harmonizes secrecy, performance, and user autonomy, representing a notable development in at ease cloud data management.

Future improvements may contain AI and machine learning to enhance deduplication and improve key-word seek precision, facilitating extra clever data control. Augmenting multi-cloud competencies could improve accessibility across numerous cloud structures. Enhancing the user interface can render the machine extra intuitive, subsequently augmenting reputation amongst non-technical users. Moreover, integrating compliance functionalities for rules such as GDPR and HIPAA would provide instruments for information governance, auditing, and reporting, thereby ensuring comfy and legally compliant cloud storage solutions.

REFERENCES

[1] M. Song, Z. Hua, Y. Zheng, T. Xiang and X. Jia, "FCDedup: A two-level deduplication system for encrypted data in fog computing", IEEE Trans. Parallel Distrib. Syst., vol. 34, no. 10, pp. 2642-2656, Jul. 2023.

[2] J. Li, P. P. Lee, C. Tan, C. Qin and X. Zhang, "Information leakage in encrypted deduplication via frequency analysis: Attacks and defenses", ACM Trans. Storage, vol. 16, no. 1, pp. 1-30, Mar. 2020.

[3] P. Puzio, R. Molva, M. Önen and S. Loureiro, "ClouDedup: Secure deduplication with encrypted data for cloud storage", Proc. 5th Int. Conf. Cloud Comput. Technol. Sci., vol. 1, pp. 363-370, 2013.

[4] M. Bellare, S. Keelveedhi and T. Ristenpart, "Message-locked encryption and secure deduplication", Proc. 32nd Annu Int. Conf. Theory Appl. Cryptographic Techn., pp. 296-312, 2013.

[5] J. R. Douceur, A. Adya, W. J. Bolosky, P. Simon and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system", Proc. 22nd Int. Conf. Distrib. Comput. Syst., pp. 617-624, 2002.

[6] J. Gantz and D. Reinsel, "The digital universe in 2020: Big data bigger digital shadows and biggest growth in the far east", IDC iView IDC Analyze Future, vol. 2007, no. 2012, pp. 1-16, 2012.

[7] J. Gants, "Digital universe decade—Are you ready?", 2010, [online] Available: https://www.mendeley.com/catalogue/235ded01-b2cb-3e9f-83ab-009f793f887c/.

[8] D. T. Meyer and W. J. Bolosky, "A study of practical deduplication", ACM Trans. Storage, vol. 7, no. 4, pp. 1-20, Feb. 2012.

[9] L. Wei et al., "Security and privacy for storage and computation in cloud computing", Inf. Sci., vol. 258, pp. 371-386, Feb. 2014.

[10] M. Abadi, D. Boneh, I. Mironov, A. Raghunathan and G. Segev, "Message-locked encryption for lock-dependent messages", Proc. 33rd Annu. Cryptol. Conf. Santa Barbara CA USA, pp. 374-391, 2013.

[11] J. Li, G. Wei, J. Liang, Y. Ren, P. P. Lee and X. Zhang, "Revisiting frequency analysis against encrypted deduplication via statistical distribution", Proc. IEEE Conf. Comput. Commun. (INFOCOM), pp. 290-299, 2022.

[12] J. Li et al., "Enabling secure and space-efficient metadata management in encrypted deduplication", IEEE Trans. Comput., vol. 71, no. 4, pp. 959-970, Mar. 2021.

[13] S. Keelveedhi, M. Bellare and T. Ristenpart, "DupLESS: Server-aided encryption for deduplicated storage", Proc. 22nd USENIX Secur. Symp., pp. 179-194, 2013.

[14] J. Stanek and L. Kencl, "Enhanced secure thresholded data deduplication scheme for cloud storage", IEEE Trans. Dependable Secure Comput., vol. 15, no. 4, pp. 694-707, Aug. 2016.

[15] Z. Yang, J. Li and P. P. Lee, "Secure and lightweight deduplicated storage via shielded deduplication-before-encryption", Proc. USENIX Annu. Tech. Conf., pp. 37-52, 2022.

[16] J. Stanek, A. Sorniotti, E. Androulaki and L. Kencl, "A secure data deduplication scheme for cloud storage", Proc. 18th Int. Conf. Financial Cryptogr. Data Secur., pp. 99-118, 2014.

[17] M. S. Jawed and M. Sajid, "A comprehensive survey on cloud computing: Architecture tools technologies and open issues", Int. J. Cloud Appl. Comput., vol. 12, no. 1, pp. 1-33, 2022.

[18] J. Xu, E.-C. Chang and J. Zhou, "Weak leakage-resilient client-side deduplication of encrypted data in cloud storage", Proc. 8th ACM SIGSAC Symp. Inf. Comput. Commun. Secur., pp. 195-206, 2013.

[19] R. Di Pietro and A. Sorniotti, "Boosting efficiency and security in proof of ownership for deduplication", Proc. 7th ACM Symp. Inf. Comput. Commun. Secur., pp. 81-82, 2012.

[20] S. Halevi, D. Harnik, B. Pinkas and A. Shulman-Peleg, "Proofs of ownership in remote storage systems", Proc. 18th ACM Conf. Comput. Commun. Secur., pp. 491-500, 2011.