A Secure Backup System Using Multi Cloud and Fog Computing

[1]KothakotaSatya Supriy, [2]D.S.Ramkiran
[1,2]Dept. of Computer Science & Engineering, KIET, Kakinada, AP, India

Abstract
Data backup is essential for disaster recovery. Current cloud-based solutions offer a secure infrastructure. However, there is no guarantee of data privacy while hosting the data on a single cloud. Another solution is using multi-Cloud technologies. Although using multiple clouds to save smaller pieces of the data can enhance data privacy, it comes at the cost of the need for the edge device to manage different accounts and manage communication with different clouds.
These drawbacks made this technology rare to use technology. In this paper, we propose DropStore to provide an easy-to-use, highly secure, and reliable backup system using state-of-the-art multi-Cloud and encryption techniques. DropStore adds an abstraction layer for the end-user to hide all system complexities using a locally hosted device,,,,,the Droplet"""", that is fully managed by the user. Hence, the user does not rely on any untrusted third party. This was achieved using Fog Computing technology. The uniqueness of DropStore comes from the convergence of MultiCloud and Fog Computing principles. The system implementation is open-source and available online. Performance results show that the proposed system improves data protection in terms of reliability, security, and privacy preservation while maintaining a simple and easy interface with edge devices.

I. Introduction

The widespread use of digital storage in networking and computing has led to an increase in the importance of data backup. However, digital storage also poses several threats, including security attacks, hardware failure, and operation errors. To prevent these threats, data backup is crucial, and cloud backup systems offer protection and disaster recovery.
With the increased use of cloud computing technology, it has become challenging to ensure data protection. Although many cloud service providers offer their services at low costs or even for free, they often lack uniform policies for data protection and privacy preservation. This poses a significant risk to organizations and individuals who rely on cloud services to store their data.
To address this issue, researchers have developed the multi-Cloud concept, which uses a heterogeneous architecture with various cloud computing and storage facilities. This architecture offers increased data protection, flexibility, and cost optimization. When using multi-Cloud architecture, users can either manage resources and services themselves or enlist the help of a third-party service provider. By doing so, they can ensure that their data is protected from various threats while taking advantage of the benefits of cloud computing technology. The Number of rounds Nr is based on key length of Nk and words. Nb is steady for all forms. Cryptography is the portion of science which bargains with data security which has gotten to be exceptionally basic in present day computing framework to secure information transmission and capacity. The significance of security has gotten to be a major need as broad utilize of individual communication gadgets. The trade of advanced information in cryptography comes about completely different calculation classified into two cryptographic components: symmetric key in which same key issue for encryption and decoding which are quick and less demanding to actualize than topsy-turvy key calculation.

B. Related Work

Rongxing Lu et.al[1] The mobile healthcare (mHealthcare) system has been envisioned as a significant computing application for improving health care quality and saving lives. In an m-Healthcare emergency, an opportunistic computing paradigm can be used to overcome the challenging dependability issue in the PHI process. To solve this problem, we present SPOC, a novel secure and privacy-preserving opportunistic computing paradigm. Marlena Ning Cao et.al [2] Cloud computing is the longawaitedrealisation of computing as a utility, in which cloud customers can store their data remotely in the cloud and access high-quality apps and services on demand from a shared pool of programmable computer resources. Individuals and businesses are both motivated to outsource their local complicated data management system to the cloud because of its excellent flexibility and cost savings. Sensitive data must be protected in the cloud and beyond to prevent unauthorised access. Jing Chen et.al[3] Wireless mesh networks are used in a variety of applications, including industrial control, environmental monitoring, and military operations. Network coding is a potential technology that can help wireless mesh networks run better. Because the fixed backbone of wireless mesh networks is usually unlimited energy, network coding is appropriate. It effectively addresses the flow coding collision problem by introducing the information process, which effectively lowers the decoding failure rate. Ruiying Du et.al[4]This problem is addressed by integrating broadcast encryption techniques with ABE schemes in a user-revocable ABE system. To enable direct user revocation, the data owner should accept complete responsibility for maintaining the entire membership list for each attribute group in this scheme. This strategy is not relevant to the data sharing system since after saving their data on an external storage server, the data owners will no longer have direct control over their data. User revocation in the ABE-based data sharing system was also recently addressed by Yu et al. The data server performs user revocation using proxy encryption in this scheme. However, in order to revoke users, the KGC should generate all secret keys on behalf of the data server, including the proxy key. To prevent revoked users from decrypting the ciphertext, the server would reencrypt it using the proxy key received from the KGC. ZhengxiaZou et.al [5] Many realistic ABE-based systems have recently recognised the need of immediate user revocation. Ostrovky et al. recommended employing ABE, which enables negative clauses, for user revocation. To accomplish so, simply combine the AND of negation of revoked user identities together. One disadvantage of this technique is that the size of the private key grows by a factor of login, where n is the maximum number of attributes. For no monotonic ABE, Lewko et al. presented more efficient instantiations of the Ostrovky et al. architecture, where public parameters are only O(1) group elements and private keys for access structures including t leaf attributes are of size O. (t). However, these userrevocable schemes have a constraint in terms of the number of users

III. Previous Implementations

Researchers have been concerned about the concept of big data. Using big data to gather useful information has become a popular trend in recent years. Big data research's main purpose is to process massive amounts of data in order to extract useful information. Furthermore, in the long run, a proper approach to large data processing is crucial. However, dealing with large amounts of data requires more than a single computer or server. As a result, in the building of big data, the distributed structure is extremely crucial.Cloud computing originated from distributed computing, which may provide a variety of big data-related services such as distributed processing, virtualization, and distributed databases.

- It is impossible to cope with huge data on a single computer or server.
- A great number of security issues can occur when all data must be uploaded to the cloud.
- They encrypt real data with a symmetric key encryption process and utilize a deniably encrypted plan-ahead symmetric data encryption key[4].
- Most decryption error concerns exist in deniable encryption methods. The designed decryption mechanisms are at blame for these problems.
- Decryption is performed using the subset decision process. According to the subset decision result, the receiver selects the decrypted message.
- An error occurs if the sender selects an element from the universal set while the element is found in the specialized subset.
- All transparent set-based systems have the same issue[7].

IV. Proposed Methodology

Singh et al. presented a secure data deduplication technique that uses secret sharing schemes. In this technique, data is sliced based on the Permutation Ordered Binary (POB) numbering system and stored on multiple cloud servers. The key information is divided into multiple random shares based on the Chinese Remainder Theorem (CRT) and saved to multiple servers. However, data can be restored only if all the shares are available, while the key can be restored from k servers out of n servers, where k is less than n. Therefore, the system will not be able to survive in the event of cloud service provider lockouts.

Trivia is a chunking-based backup system that minimizes storage needs by using the sec-cs data structure for deduplication of flat contents. This system offers multi-Cloud storage for the generated backups, making storage efficient. However, this comes at the expense of data reliability and immunity against lockouts.

Trusty Drive is a document storage system that utilizes multiple cloud providers to store documents while preserving user anonymity and document anonymity. The focus of the system is to save and secure document files only. However, the system does not provide an interactive or easy way to share and view saved documents. MultiCloud storage systems offer high availability, strong security, and prevent service provider lockouts. However, these systems have limitations. To overcome these limitations, future research on multi-Cloud storage systems should focus on implementing redundancy techniques to ensure data reliability and improving interactivity and sharing features. Additionally, security measures must be in place to mitigate potential security risks associated with the use of thirdparty providers.

The below diagrams depict the entire system architecture of the droplet in network and Dropstore software

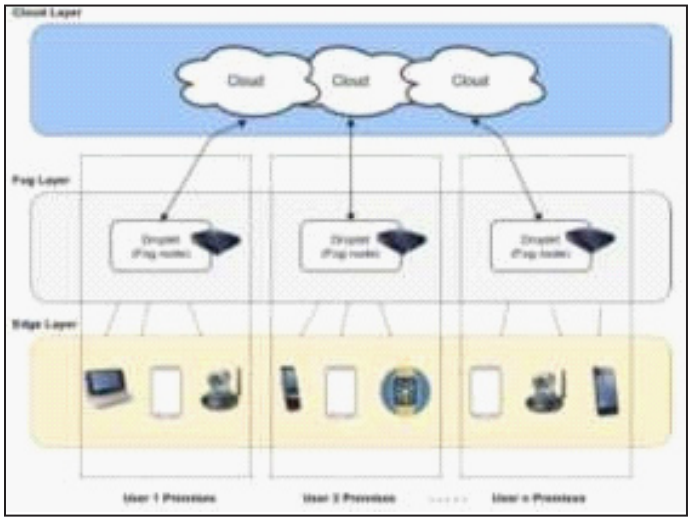

Fig. 1: System Architecture of the droplet in network

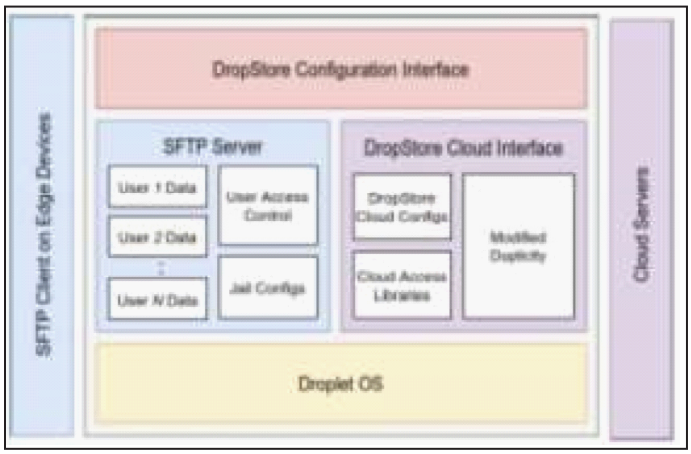

Fig. 2: Dropstore software architecture

V. System Implementation

There are 4 modules

- Edge Nodes
- Droplet
- Public Cloud
- DropStore- System

A. Edge Nodes

- Register
- Login
- Register Device
- Upload Data
- View Data
- My Profile
- Logout

B. Droplet

Login

- User management
- Fog layer
- Logout

C. Public Cloud

Login

- Edge devices
- Droplet Fog Layer

D. DropStore-System: -

- Login
- Edge Nodes
- Droplet Fog Layer
- Logout

VI. Results

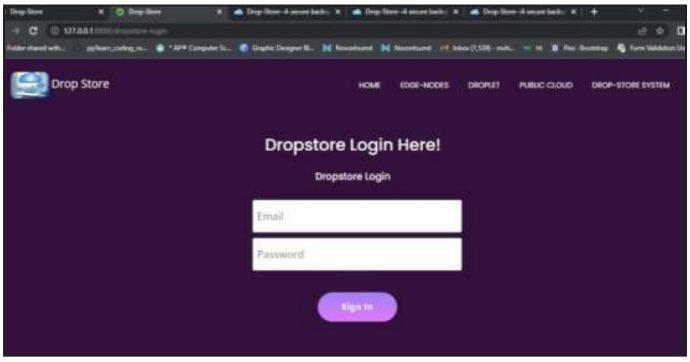

Fig. 3: DropStore Login Page

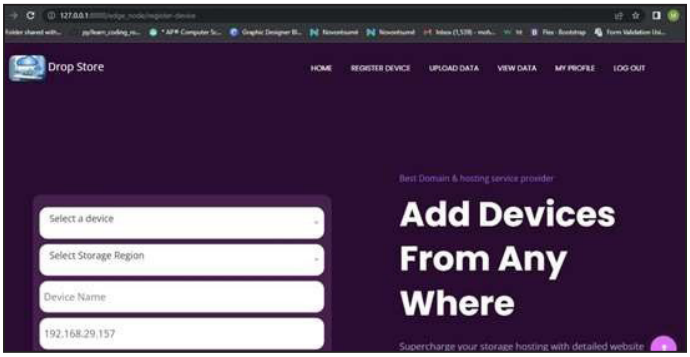

Fig. 3: Adding Devises to Dropstore

VII. Conclusion and Future Work

In conclusion, DropStore presents a novel solution to address the challenges of data security and reliability by utilizing the Multi-Cloud and Fog Computing paradigms. The system is designed to provide a seamless backup experience for individual users while abstracting them from the system's complexities. DropStore ensures data security and user privacy through encryption and data partitioning on Multi-Cloud Storage. The system's efficiency and reliability were validated through real-world experiments using two different implementations. The results demonstrate that DropStore is capable of storing and retrieving data reliably with minimal complexity at the edge side.

As future work, we plan to explore better scheduling strategies for data uploading to the cloud. New scheduling strategies will consider QoS parameters and the remaining storage at each CSP to optimize the system's performance. Additionally, we will develop linear block codes for data replication to enhance the system's error detection and correction capabilities. These advancements will further improve the system's reliability and efficiency while reducing complexity, making it an even more attractive backup solution for individual users.

References:

[1] L. Wang, J. Tao, M. Kunze, A. C. Castellanos, D. Kramer, and W. Karl, „„Scientific cloud computing: Early definition and experience,"" in Proc. 10th IEEE Int. Conf. High Perform. Comput. Commun., Sep. 2008, pp. 825– 830.

[2] Y. Singh, F. Kandah, and W. Zhang, „„A secured cost-effective multi-cloud storage in cloud computing,"" in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2011, pp. 619–624.

[3] P. Habibi, M. Farhoudi, S. Kazemian, S. Khorsandi, and A. Leon-Garcia, „„Fog computing: A comprehensive architectural survey,"" IEEE Access, vol. 8, pp. 69105–69133, 2020.

[4] R. K. Naha, S. Garg, D. Georgakopoulos, P. P. Jayaraman, L. Gao, Y. Xiang, and R. Ranjan, „„Fog computing: Survey of trends, architectures, requirements, and research directions,"" IEEE Access, vol. 6, pp. 47980–48009, 2018.

[5] S. Yi, C. Li, and Q. Li, „„A survey of fog computing: Concepts, applications and issues,"" in Proc. Workshop Mobile Big Data, New York, NY, USA, Jun. 2015, pp. 37–42, doi: 10.1145/2757384.2757397.

[6] S. Sarkar and S. Misra, „„Theoretical modelling of fog computing: A green computing paradigm to support IoT applications,"" IET Netw., vol. 5, no. 2, pp. 23–29, Mar. 2016.

[7] B. Tang, Z. Chen, G. Hefferman, T. Wei, H. He, and Q. Yang, „„A hierarchical distributed fog computing architecture for big data analysis in smart cities,"" in Proc. ASE BigDataSocialInform., New York, NY, USA, 2015, pp. 1–6. [Online]. Available: https://dl.acm.org/doi/10.1145/2818869.2818898

[8] K. Kai, W. Cong, and L. Tao, „„Fog computing for vehicular ad-hoc networks: Paradigms, scenarios, and issues,"" J. China Universities Posts Telecommun., vol. 23, no. 2, pp. 56–96, Apr. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1005888516600213

[9] S. U. Zaman, R. Karim, M. S. Arefin, and Y. Morimoto, „„Distributed multi cloud storage system to improve data security with hybrid encryption,"" in Intelligent Computing and Optimization, P. Vasant, I. Zelinka, and G.-W. Weber, Eds. Cham, Switzerland: Springer, 2020, pp. 61–74.

[10] P. Singh, N. Agarwal, and B. Raman, „„Secure data deduplication using secret sharing schemes over cloud,"" Future Gener. Comput. Syst., vol. 88, pp. 156–167, Nov. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X17327474

[11] A. Sreekumar and S. B. Sundar, „„An efficient secret sharing scheme for n out of n scheme using POBnumber system,"" Hack, vol. 33, pp. 1–88, Mar. 2009.

[12] V. J. Katz, A. Imhausen, E. Robson, J. W. Dauben, K. Plofker, and J. L. Berggren, The Mathematics of Egypt, Mesopotamia, China, India, and Islam: A Sourcebook. London, U.K.: Princeton Univ. Press, 2007. [Online]. Available: https://books.google.com.eg/books?id=3ullzl036UEC.