Secure Cloud Backup for Data Sources Based on Blockchain

S K Mouleswaran1*, R Aruna2, J Visumathi2, and S Gurusubramani3
1Department of Computer Science and Engineering, Dayananda Sagar University, Bangalore, Karnataka, India
2Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India
3Department of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, Tamil Nadu, India
Email: *meetmoulee@gmail.com

Abstract. Users can upload, save and download their data with cloud storage. The cloud does however handle the user's generation of data, as well as devices created by the fast Internet development of stuff, including such webcam, and IoT and edges cameras. This growth in outsourced data volume has posed two main problems in cloud infrastructure cloud management: optimum data storage and data protection. Cloud storage employs a deduplication strategy to prevent redundant data to solve the former problem and uses cryptography to solve the latter problem. Both of these questions however are orthogonal. The convergent encryption is concurrently tackles the two problems. The conventional systems rely on web server, however which put some confidence in a private entity that may malfunction in one stage. In this paper we propose a new, transparent, authenticated online storage framework, which removes data replication from the security of the file and enhances the credibility of blockchain software. Two different approaches to data deduplication secrecy are discussed, in particular: double hazing is used and symmetrical encryption is used. We have implemented the suggested architecture in order to show stability and accessibility in spite of the increase request throughput requirements for cloud data processing.

Keywords: Block chain, encryption model, deduplication, data secrecy, cloud.

1. Introduction
Cloud technology is among the most general cloud technologies that allow consumers to supply large volumes of cloud information overweighting their infrastructure. The creation and production of IoT data by Internet of Thing (IoT) ensures that cloud storage dependence is further increased. IoT becomes an important component in the design of mobile devices and related facilities. The Fog age provides, in conjunction with IoT, a massive amount of data. This would lead to growing volumes of cloud storage data becoming redundant in several specifics including IoT and cloud computation.
Figure 1: Cloud model Overview

Cloud computing uses identical data to overcome redundant data collection and maximise the efficiency of storage. The deduplication of data is a method that avoids the storage of several file copies by storing only one file copy [1], [2]. Therefore, both bandwidth and availability are increased. It also raises storage space by the block level saving of one copy of a file [3], [4]. Cloud storage especially prevents numerous copies of existing files at file level, i.e. when client C needs a F file uploaded, to Cloud Server CS testing if F is usable. When CS detects F, then A is prohibited from uploading the file. At the block stage, CS tests deduplication on multiple file blocks. In order to conserve the bandwidth of the transfer, the user often uploads a single file tag to validate that the file is available [5], [6] instead of a file. Figure 1 illustrates in detail the deductibility mechanism. Sadly, when a user encrypts his/her files, CS can’t allow deduplication.

Normally, before outsourcing to CS a security-relevant user encrypts his code. Since encryption makes discriminating between a single text's encrypted texts from the domain generated value difficult for anyone, encryption clashes with the deduplication mechanism [7]. When AE in addition BE users encode a single text PF file using KB besides KB, for example, they create two distinct cipher-text files, i.e., AEF in addition BEF. Thus, it does not check the presence of the file while CS uses deduplication procedure. The cloud computing can’t then saves storage space with deduplication. Convergent Encryption (CE) is worked for deduction of encrypted data. When converging, a user first produces the file hash too then habits it to encode the file. The records are encoded besides secured in this way.

However, the hash of the file, namely the suffix, the user uses to encrypt the cloud for checking deduplication. Therefore, CS besides used the tag to decipher the encoded files. The current works in the literature require external servers to also use information deduplication lengthways with meeting encryption [8]. The denial of the user's uploading of the file can also be an adverse CS problem. In the beginning, for example, a user loads a CS file. However, occasionally, the CS only castoffs the file out of stock to save interplanetary, also once the user revenues CS repudiates the user has actually submitted the dossier. It is regarded as a retrievability resistant [9], in which a CS indicates that the user already has the file imported. However, in order to accommodate the large load produced on the cloud storage, the solution for the above-described issues needs to be lightweight, as the cloud load often increases with the number of IoT nodes.

Influenced by the above difficulties, in these studies we suggest a new, integrated sturdy, cryptographic cloud storage infrastructure, which removes data warehouse deductions, uses customised file confidentiality and file completeness with Blockchain [10]. The client encrypts data using probabilistic encryption to guarantee file privacy in our unorthodox idea and uses the quality of the data multipathing as a phrase. In the second solution, symmetrical encryption and protected information are used to create a tag that encourages various entities, while also maintaining user records, to access cloud storage.

2. Related Work
This segment focuses on convergence cryptography and blockchain to explore relevant works in secure cloud computing [11]. Then we address the issue assertion by stressing in an adversarial cloud
scenario the device model and hazard model. Convergence encryption is a means to encrypt a password with a key that is related with a file[12]. A user produces the file hash first and uses the hash output for encrypting the file. The convergence encryption method is as given in Figure 2.

The formal description of fusion encryption consists of three algorithms. Blockchain is a distribution system to store the sequence list simply. The same distributed data order applies to each node in Blockchain[13]. Since Blockchain is a decentralized storage, the validity of the results between the nodes is established by a consensus algorithm. Centered on the clustered, immobilisation nature, is becoming a viable technology for unfaithful peers in the peer network. In reality, Blockchain is marketing various channels. Examples include Hyperledger [16], Bigchaindb [15], Ethereum [14] and Bitcoin[13]. Examples include: Blockchains remain primarily used in the electronic currencies, Bitcoin including. Different kinds of transaction computing systems for other sectors or e-commerce operations are supported in Ethereum and Hyperledger beyond crypto currencies. Though, these networks deficiency high efficiency besides the research community must boost their performance urgently.

![Encryption Model](image)

**Figure 2: Encryption Model**

We're using a Mystiko Blockchain [11] framework for Yugala deployment, which is directed primarily at big data operations. It was developed using the distribution of storage age from Apache Cassandra [17] and from Apache Kafka [18]. Mystiko Blockchain's key explanation for using Yugala is its high business efficiency and scalability. Since the Mystiko storage template supports massive data consumption loads, encrypted file loads have been placed on the Blockchain. In [19] discussed about privacy of the healthcare system using cloud and blockchain trending techniques for content Deduplication. The Block Chain Based technique discussed for applying the security on Food Beverages [20]

Mystiko uses dynamic programming from Scala and from Aplos smart Contract Service to support parallel transaction executions. All intelligent contracts with Aplos Mystiko Blockchain on the Yugala network, written-ten 10. The cloud storage system, which consists of a server for the cloud storage and community of users, has been used in the proposed environment (as seen in Fig. 1). The storage of encoded files is possible in the cloud. Users encrypt file so that the presence of files can only be verified by others who have the same files, whereas the cloud storage system is unable to.

Danger model: The cloud storage is allegedly semi-honest and does not balance the customers. The CS recognises, but does not know the contents of the encrypted file, the ex-instance of the files the user is demanding. CS can also easily uninstall the file that is loaded into the cloud by a customer. Therefore, information confidentiality and file privacy are the two security priorities that the proposed architecture offers for the threat model.

Initially, anonymity is accomplished by all users encrypting files before downloading, so CS cannot read the file material. Second, CS cannot delete the document which itself implies devices stability and confidentiality, since the presence of files within BlockChain is now accessible (because the data in BlockChain is spread through many repositories and the details is redundancy). The new solution is to optimise encryption with an intelligent combination of integration authentication and spontaneous puzzle strategies.

3. Proposed System
In this segment we will first deliberate the architecture by naming all major players. Then we clarify two ways of secrecy in code deduplication: one in dual hash and one synchronous. There are essentially three players in the architect: users, Blockchain and cloud infrastructure. Take into account
that two clients use the same file. The Blockchain identifier is first submitted by Client1. Blockchain verifies that there is already a file with the tag. The file is fresh if there is no file. This is a reserve file. Using the Blockchain search tag Blockchain to import the same file to the Yugala cloud by the client and do not save a new file. For all replication control operations, the smart agreement on the Mystiko Platform is used. Then the two planned data deduction files are discussed in secrecy.

In instruction to illustrate the multi-user situation, two users have been measured, which often applies to more users. The method is detailed below.

a) At the outset, the user first determines the threat of the file he needs to submit before attempting to upload the file. Then it uses the hash output to construct a hash output again that it considers to be the name. He then transfers the tag to the cloud.

b) The storage in the cloud sends the user negative or positive) when the file is not in the cloud (or otherwise).

c) If a positive cloud response is obtained to the use; i.e., the cloud supplies the encrypted tag file, the user can do nothing.

d) If not, the user uses a hash output for the key to encrypt the file.

In this we talk about the basic Blockchain Mystico instead and start moving on to dissimilar Yugula modules besides their Mystico touch. Mystiko is an easily deployable Blockchain technology for huge datasets. Apache Cassandra uses a different database [17] (with Paxos agreement as the underneath-lying agreement podium). Mystiko utilises Hadoop Framework besides Akka for reverse compression activities on big data. Mystiko used the Elastic search API based on Apache Lucene to allow complete Blockchain data text searches. Mystiko Blockchain is a Micro services scattered framework. These Micro services can be found in extremely scalable systems with Docker and Kubernetes. On current Blockchain systems, Mystiko tackled three key bottlenecks, namely, architecture to run order, full data reproduction and Imperative smart contracts style. Redis cache-based validate-execute-group architecture provides for dealing with issues relevant to conventional order-execute Blockchain architecture [16]. This architecture helps transactions to be authenticated and completed while a client
transfers to the network. The customer must not delay until the transaction is committed by a block. This new architecture offers Mystiko Blockchain high degree of scalability and transaction performance.

This approach prohibits mystiko in traditional blockchains through reproducing complete nodes. As an antidote to the essential format of the existing Blockchains, Mystiko provides the reference implementation platform from Aplos. The infrastructure is based with the dialect Scala and Akkators that is completely regulated by competition. The Aplos platform facilitates concurrent transaction execution with Akka actors and Practical Programming-based competition control. Both Mystiko transfers are done by performers. Transactions are sent to the Customer through apache kafka [18] to Mystiko aplos service. Mystiko is provided with the name and message of the actor. Actor absorbs messages; then it will write/update/quest on the directory besides reappearance the answer to the customer on the basis of the post.

4. Results and Discussion
We did a Yugala cloud ser-vice performance assessment. In AWS 2x huge examples (16GB RAM besides 8 CPUs), we used the Mystiko Blockchain Cluster besides Yugala facilities and obtained the performance. The assessment is rendered using the Yugala dual-hair method, which was based on Mystiko Blockchain. The findings are obtained in 5 areas: transaction efficiency, scalability of transactions, search performance, double-hasing performance, and Yugla customer encryption performance. We reported the amount of transactions to be performed in any participant of Mystiko Blockchain in Yugala’s cloud for transaction throughput in Figure 4.

![Figure 4: Performance of the proposed system](image-url)

We saturated transfers with each Blockchain member and recorded transactions. Each Blockchain combination has consistent transaction efficiency. The three key sound systems for the performance of high transactions are the fundamental Apache Cassandra repositories in Mystiko, Mystiko’s Framework for Validating and Executing Community. Three key sons of the rea are here. In addition, the number of concurrent users was recorded in the Yugala network compared with a number of Blockchain pairs, to track scalability. We inundated a concurrent settlement in any Blockchain peer and recorded the volume of money made.

The transaction throughput increases linearly as we increase the number of Blockchain peers. Cassandra's Storage on which Mystiko Blockchains were based is the key explanation for this line scalability. Cassandra practises master less ring architecture, such that all cluster nodes are able to write. The cluster is added to a network, the transaction throughput increases linearly. The search efficiency of the Yugala cloud service based on Mystiko Blockchain was measured next. For this research, the Mystiko Blockchain was booted and the time needed to locate a record was collected.
In 4 milliseconds, we might check for a database of two million transactions. Mystiko Blockchain uses Cassandra as the storing framework and Elastic search as the indexing method of Apache Lucene. This storage based on Elastic search and Cassandra results in very fast searches in Yugala cloud storage. Lastly, we assessed Yugla client SDK’s encryption efficiency and complexity analysis in Figure 5. We used the Yugla client SDK from Golang for this assessment. It uses asymmetric GCM mode AES encryption. These SDKs encoding feature operates in contradiction of varying file sizes besides records the number of encryption ions within seconds.

5. Conclusion
This paper explores the current encrypted Yugula cloud storage architecture, which promises secrecy and integrity of files using modified convergent encryption and blockchain. Two approaches to secrecy are discussed in particular: one is double hazing and the other is symmetrical encoding. We have used Yugala as proof of principle. The cloud service Yugala is focused on the extremely scalable Blockchain Mystiko Blockchain. Since Mystiko’s key goal for big data, we will provide Yugala cloud service to support IoT requirements. The result segment displays the formatting of the architecture that illustrates the specifications of high transaction quantity with admiration to the upload file size and Blockchain load in IoT setup.

References
[1] Baktir, A. C., Ozgovde, A., & Ersoy, C. (2017). How can edge computing benefit from software-defined networking: A survey, use cases, and future directions. IEEE Communications Surveys & Tutorials, 19(4), 2359-2391.
[2] Sharma, P. K., Chen, M. Y., & Park, J. H. (2017). A software defined fog node based distributed blockchain cloud architecture for IoT. Ieee Access, 6, 115-124.
[3] Liu, M., Yu, F. R., Teng, Y., Leung, V. C., & Song, M. (2018). Computation offloading and content caching in wireless blockchain networks with mobile edge computing. IEEE Transactions on Vehicular Technology, 67(11), 11008-11021.
[4] Chaudhary, R., Jindal, A., Aujla, G. S., Aggarwal, S., Kumar, N., & Choo, K. K. R. (2019). BEST: Blockchain-based secure energy trading in SDN-enabled intelligent transportation system. Computers & Security, 85, 288-299.
[5] Jindal, A., Aujla, G. S., & Kumar, N. (2019). SURVIVOR: A blockchain based edge-as-a-service framework for secure energy trading in SDN-enabled vehicle-to-grid environment. Computer Networks, 153, 36-48.
[6] Tang, F., Ma, S., Xiang, Y., & Lin, C. (2019). An efficient authentication scheme for blockchain-based electronic health records. IEEE access, 7, 41678-41689.
[7] Cha, S. C., Chen, J. F., Su, C., & Yeh, K. H. (2018). A blockchain connected gateway for BLE-based devices in the internet of things. IEEE Access, 6, 24639-24649.

[8] Yao, Y., Chang, X., Mišić, J., Mišić, V. B., & Li, L. (2019). BLA: Blockchain-assisted lightweight anonymous authentication for distributed vehicular fog services. IEEE Internet of Things Journal, 6(2), 3775-3784.

[9] Andreev, S., Galinina, O., Pyattaev, A., Gerasimenko, M., Tirronen, T., Torsner, J., ... & Koucheryavy, Y. (2015). Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap. IEEE Communications Magazine, 53(9), 32-40.

[10] Punithavathi, P., Geetha, S., & Shanmugam, S. Building cloud-based biometric services. (2017). IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), pp. 35-38.

[11] Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017, March). Blockchain for IoT security and privacy: The case study of a smart home. In 2017 IEEE international conference on pervasive computing and communications workshops (PerCom workshops) (pp. 618-623). IEEE.

[12] Xu, L. D., He, W. & Li, S. (2014, November). Internet of Things in Industries: A Survey[J]. IEEE Transactions on Industrial Informatics, 10(4), 2233-2243.

[13] Shen, W., Liu, L., Cao, X., Hao, Y., & Cheng, Y. (2013). Cooperative message authentication in vehicular cyber-physical systems. IEEE Transactions on Emerging Topics in Computing, 1(1), 84-97.

[14] Chien, H. Y. (2018). Group-oriented range-bound key agreement for Internet of Things scenarios. IEEE Internet of Things Journal, 5(3), 1890-1903.

[15] Li, N., Liu, D., & Nepal, S. (2017). Lightweight mutual authentication for IoT and its applications. IEEE Transactions on Sustainable Computing, 2(4), 359-370.

[16] Pustišek, M., Kos, A., & Sedlar, U. (2016, October). Blockchain based autonomous selection of electric vehicle charging station. In 2016 international conference on identification, information and knowledge in the Internet of Things (IIKI) (pp. 217-222). IEEE.

[17] Li, L., Liu, J., Cheng, L., Qiu, S., Wang, W., Zhang, X., & Zhang, Z. (2018). Creditcoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles. IEEE Transactions on Intelligent Transportation Systems, 19(7), 2204-2220.

[18] Khan, S., & Khan, R. (2018). Multiple authorities attribute-based verification mechanism for Blockchain microgrid transactions. Energies, 11(5), 1154.

[19] Pandey, A., & Prakash, G. (2019). Deduplication with Attribute Based Encryption in E-Health Care Systems. International Journal of MC Square Scientific Research, 11(4), 16-24.

[20] Prakash, G. and Nagesh Y., (2019). Secure and Efficient Block Chain Based Protocol For Food Beverages. International Journal of MC Square Scientific Research,10(3):19-30