6th International Conference on Multidisciplinary Research (ICMR-2019)
Osmania University Campus, Hyderabad (India)
30th-31st May 2019 www.conferenceworld.in
Conference World
ISBN : 978-93-87793-89-7

A Blockchained Trusted Access Controller for Data Sharing in Decentralized Storage System

Anil V. Deorankar[1], Pratik S. Nichit[2]

[1,2]*Computer Science and Engineering, Government College of Engineering, Amravati (MH), India.*

ABSTRACT

With the development of attribute based encryption (ABE) the problem of privacy data leakage and fine grain access control is solved. But the problem with these schemes is that it requires a third party for generating keys. This third party (intentionally) can decrypt all of the data stored in the cloud server. Again the cloud server mostly make use of centralized storage. There is always a risk of collapse of system with single point of failure in centralized storage. A decentralized storage can overcome the risk of single point of failure. In this paper we designed an architecture that combine Hyperledger blockchain, decentralized storage IPFS and ABE scheme to achieve fine grained access control over the data stored in decentralized storage. In this paper we implemented attribute based keywords search functionality. This avoid an unauthorized user from access data. Finally we implemented the scheme on Linux system and Hyperledger fabric blockchain with decentralized storage, IPFS. The experimental result shows that our scheme is realizable.

Keywords: ABE, attribute revocation,attribute based keyword search,blockchain, Hyperledger fabric, IPFS.

I. INTRODUCTION

In today's world cloud storage has emergedas an important business model for storing data. It has made possible to store, share and access data at anywhere and anytime on internet resources. Such technology has been globally accepted and proved to be a very successful storage system. These storage technology require large amount of storage space, workers for maintaining records and guarantee the privacy and availability of data at any time. Thus such kind of systems service is provided only by large organization. Such organization are assume to be trusted third party for storing important data. Even if such system willing to guarantee the availability of data at any time, cloud service provider still has a risk of force majeure. In force majeure a genuine user is interrupted from his liability, which lead to inability of user from accessing its own data. Again traditional cloud storage make use of centralized storage system. A single point of failure in system leads to collapse of the entire system. Again this storage system require large amount of cost of development and maintenance. The cost of centralized storage system is mainly

because of legal buy cost, rents and employee wages. Employee wages and rents keep on increasing with time thus the cost of centralized cloud storage system also increases. With the advancement in technology the cost storage devices are getting cheaper. Thus there is a need of decentralized storage.

Decentralized storage systems do not rely on a central service provider. It allow users to store files at storage nodes that rent out free storage space. These systems make use of blockchained core structure. It guaranteed the availability of data at any time. User need not to worry about data being inaccessible.

The blockchain technology used in Bitcoin brought sophisticated implementation to decentralized storage system. The blockchain technology has enabled the peer to peer connection of bandwidth, storage space and processing power. With this technology one can rent free space of hard disk for storage over an internet. User don't have to worry of inaccessibility of data because the chaincode deployed on the blockchain guaranteed the availability of data. The only thing that the user have to do is to pay fee regularly for the data they have stored.

Imagine if an users want to secretly share data stored in traditional cloud storage systems, a technology is needed to achieve fine grain access control over data such that the data can only be decrypted and accessed by a specific user. In need of this demand, the attribute-based encryption mechanism [1] was proposed and rapidly developed. In order to achieve fine grained access control over the datain attribute based encryption, the data owner has ability to specify the access policy for data based on the user's identity and attributes.In all ABE schemes a trusted private key generator (PKG) is required to setup the system and distribute the corresponding secret key to users [2]. There are number of problems with this a system. Firstly, a private key generator PKG is not trustworthy in real sense. Secondly, it has a problem of key abuse, the data owner cannot control its own data. The PKG has potential to decrypt all data storedat the server, and PKG may misuse data for illegal gains. Once the data owner loses its own control over data, he can't even decrypt his own data, and PKG can still decrypt the data.

There is need to allow only data owner to control their own data and generate secret key for users. For example, a data manager in a bank should be able to generate secret key for clerks, branch manager, and specialist officer so they can access data according to different levels of their position.

To guarantee the availability and the privacy of data, the data storage centers should be transferred from the centralized cloud storage systems to the decentralized storage systems, which have advantages such as lower prices than traditional cloud storage, the high data throughput, and to stop worrying about single point of failure of system.

In this paper, we designed an architecture to achieve fine grained access control over data in decentralized storage systems, and perform attribute based search with keywords on data in decentralized storage. The contributions of this paper are as follow: (1) we proposed an architecture that combines attribute based encryption technology, the hyperledger blockchain, and decentralized storage system IPFS to achieve finegrained access control over the data. The trusted private key generator (PKG) is not needed in our system. Theonly dataowner DO can controlshis own

data, and the Admindistribute private key for data users.The hyperledger blockchainisused for storing the user private key, the problem of key storage and retrieval in the traditional ABE schemes is solved. Whenever a user loses his own private key, he only needs to fetch the key from blockchain by passing the corresponding transactionsdata. (2) The chaincode is deployed on the hyperledger blockchain to implementthe attribute based keywordsearchinthedecentralizedstorage systems. Once the chaincode is deployed, it will work in good faith according to the logic of the chaincode. (3) Under the windows 10 pro system with docker desktop, a simulation of the system scheme is carried out through the hyperledger fabric, and the corresponding performance and cost were analyzed.

The paper is organized as follows. Section 2 consist of literature survey, Section 3 consists of proposed system model, the performance and security analysis is given in Section 4. Finally, the conclusion is presented in section 5.

II. LITERATURE SURVEY

A. Blockchain Technology

In recent years, cryptocurrencies such as Ethereum, Zcash, Bitcoin, etc. are getting more attention. These currencies are making use of distributed blockchain technology.Nowadays, financial sector has a major implementation of blockchain. Non-financial sector also has many implementation of blockchain such as decentralized storage, decentralized IOT [3], decentralized supply chain [4] etc. The need that only data owner should controls his own data, a blockchain based personal data management system has been proposed [5], the system provide better protection to the user data. In order to provide privacy to data in IOT system, a blockchained data privacy system for IOT was proposed [6], in which ABE (attribute based encryption) technology was implemented to control data access. In order to solve the privacy and security issues that hamper the development of big data, a blockchained access controller framework [7] had been proposed for increasing security in big data.

Decentralized storage systems such as Sia, IPFS, Storj, etc. do not depends on a centralized storage system, and allow users to store their files to different storage nodes that rented out free storage space. These systems make use of blockchain in their core structure. A Filecoin is used in a content-addressed decentralized storage platform, IPFS. In our study we found that IPFS provides a weak privacy cryptographic algorithm interface to user for uploaded files. Whereas, the Storj platform makes an end-to-end encryption of file, generate cryptographic hash fingerprint of file and stores it on the blockchain by providing a method of verifying file integrity. The Sia decentralized storage combines a peer-to-peer storage network with blockchain technology. It splits the file into multiple segments, and then each segment is encrypted individually. The fileciphertextisstored at the storage node which provide service of storage through smart contracts. The user have to pays Siacoin to the storage service provider, and the storage node service provider has toregularly submits a proofofstorage of file to stop thestoragenodefromdeletingthe stored file.

B. Attribute Based Encryption technology

In traditional cloud storage, a fine-grained access control [1] can be achieve over data can with the help of attribute based encryption technology.In attribute based encryption technology, attributes are used as parameters instead of identities of user. The users groups can be assign by Data owner that can access the data only if the attributes set meet the access policy.

Many research works have been done based on actual needs and significant research results are achievedsince ABE technology was proposed.For example, in a practical implementation, if the attributes of users changes, then theusers corresponding secret key must also be changed. Driven by this demand revocable attribute-based encryption schemes are proposed.The important privacy information of users may be revealed because of access policies, sothe attribute-based encryption schemes with hidden access policy [8], were proposed. In some commercial application, multiple authoritiesarerequiredforattribute approval, so multi-authority attribute-based encryption schemes [9] has been proposed. In case of application used in mobile devices having limited storage and computing power, outsourced decryption in attribute-based encryption schemes [10] are proposed.

Atpresent,attribute-basedencryptiontechnology can well control the access of revoked user in traditional cloud storage systems, but there is no system that can achieve fine-grained access control and control the access of revoked user over data in decentralized storage systems.

C. Keyword Searchable Technology in decentralized storage Systems

Song et al. [11] in 2000 proposed a one-to-one symmetric searchable encryption mechanism. Searchable encryption technology has many research results in traditional cloud storage systems. Some of them are: public key searchable[12] encryption, multiple keywords supporting searchable encryption [13] and so on.

With the advent of the blockchain technology, many scholars tried toconvert thetraditionalstorage system into decentralized storage systems and studiedsharing in decentralized storage systems. A searchable symmetric encryption scheme[14] was proposed by Li et al using the Bitcoin blockchain system. In this scheme, encrypted keyword indexes and users data are split and stored on the blockchain. A trustworthyprivacykeywordsearch [15] inencrypted decentralized storage scheme was proposed by Cai et al.

A blockchain based keyword search scheme[16] in decentralized storage was proposed by Jiang et al. In this scheme the keyword search is performed inthedecentralizedstoragesystems.

These keyword search technique returns the results to user without checking user attribute. Due to this a revoked user still canget the cipher text of inaccessible file, which he should not. A new attribute based search scheme in decentralized storage system is proposed.

D. Hyperledger Fabric

Hyperledger Fabric is an enterprise-grade permissioned open source distributed ledger technology (DLT) platform, designed for use mainly in enterprise contexts. Hyperledger was established under the Linux Foundation.

Fabric has a configurable and highly modular architecture enabling optimization, innovation, and versatility for a broad range of industry use cases including finance, banking, healthcare, insurance, digital music delivery, supply chain and even human resources.

Fabric supportgeneral-purpose programming languages such as Node.js, Go and Java. This means that no additional training is needed to develop smart contracts.

The Fabric platform is permissioned network, this mean that, unlike permissionless public network, the participants are known to each other. Thus a network is built such that trust does exist between participants. Legal agreement or framework handle disputes among participants.

Fabric support for pluggable consensus protocols enable the platform to be customized more effectively to fit particular use cases and trust models.

The combination of these features makes Fabric one of the best performing platforms available today in terms of transaction confirmation latency and transaction processing. It enables confidentialityand privacyof transactions

E. IPFS

Interplanetary File System (IPFS) [17] is a P2P distributed file system that aims to connect all computing devices with the same file system. IPFS uses content addressed block storage model with high-throughput and content addressed hyperlinks. It uses technologies such as self-certifying namespaces, distributed hash tables (DHT), incentivized block exchange etc. In IPFS storage nodes need not to trust each other also it need to worry of single point of failure. The advantage of IPFS over existing cloud storage is that the data is distributed and stored in different places of the world. There is no central server.

A file in IPFS system is identified by a unique file cryptographic hash string.The hash string is sufficient to publicly access the file. Blockchainarenotdesigned to store largefiles such as audio, video etc. In the proposed system, we use IPF for storing file. Afewmetadata of fileisstoredonblockchain. The users will able to read the data from the blockchainonly when the attributes set of users meets the access policy defined. Theuser will download theencryptedfilefromtheIPFS,andthen decrypt the file by getting the private key from blockcahin.

III. PROPOSED SYSTEM MODEL

We propose a framework that combines the decentralized storage system IPFS, the Hyperledger blockchain, and attribute-based encryption (ABE) technology to achieve finegrained access control over data in decentralized storage systems. The Hyperledger blockchain is used to manage the user private key. In this system we are using our own distributed hash table rather than IPNS distributed hash table. This restrict data user to access only the requested version of file. Thus the attribute revoked user cannot access the inaccessible file in this system.

The system consists of the following two entities:

Data owner (DO): DO is an organization or person that owns files that is to be shared.

Data user (DU): DU are clients that are authorized to view some of the file of DO.

The storage nodes in IPFS and the miners on the blockchain are not considered here. The double arrow pointed between the blockchain and smartcontract indicates that the smartcontract is deployed on the blockchain. The system framework is shown below:

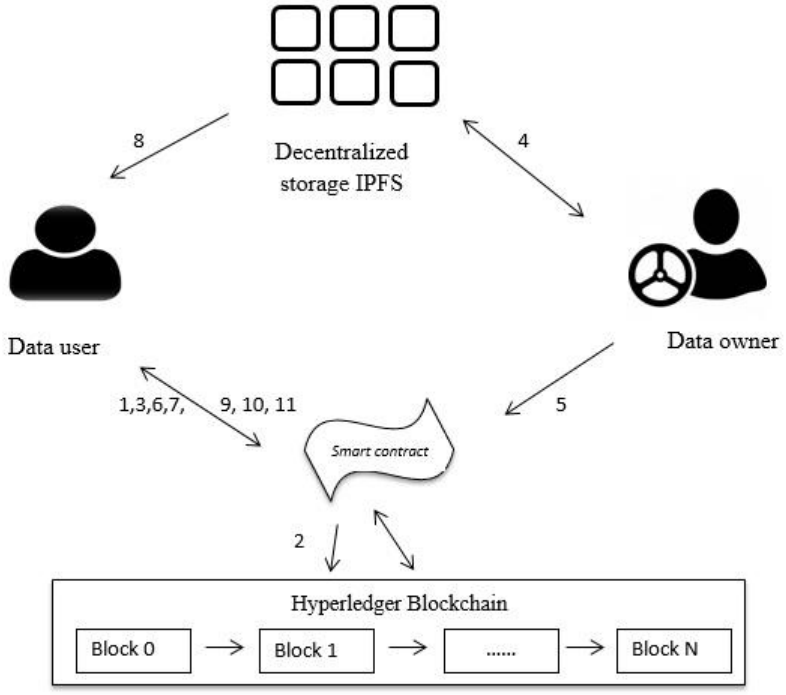

Fig 1. AProposed System Model

Admin set up the system. The system master key is encrypted and then it is embedded into hyperledger blockchain. Admin deploys a smart contract on the hyperledger blockchain.

The corresponding description of each step number are shown as follows:

1 DU sends a registration request to system.

2 Admin approve the attribute and system generates secret key for DU and embed the secret key into hyperledger transaction.

3 System sends the transaction id related to secret key to DU through a secure channel.

4 DO selects an attribute and uses ABE algorithm to encrypt the file and uploads it to IPFS.

5 DO stores the file location returned by IPFS in system.

6 DU generates a search request to search a file. System display all search results.

7 DU make a request to access a file. System allow access only to those file which have same attributes as DU have.

8 DU downloads encrypted file from IPFS.

9 DU sends transaction id to system to get private key from hyperledger transaction.

10 System get the key from hyperledger blockchain and return to DU.

11 DU uses private key to decrypt the file.

IV. PERFORMANCE AND SECURITY ANALYSIS

A. Performance Test

We implemented the scheme to analyze performance and feasibility of the system.The configuration of the platform used for implementing system: Intel core i5-M460 @ 2.53GHz processor, 8 GB RAM, and the system is Windows 10 pro with Docker Desktop. The programming language used is Go, Java and Python. In the ABE program, we use the cryptography library Miracl in attribute based encryption programs. SHA-256 algorithm is used for hash function.

Fig 2. Time required to store Private key in blockchain

We analyze the time required for generating private key based on varying number of attributes, generating hash of key and storing the key in blockchain. The time required to store private key with single attribute is 59 seconds, whereas time required for storing a private key having eight attributes is 74 seconds. Thus time required to store private key in blockchain increases with the increase in attribute set. This can be well illustrated by the graph shown in fig Figure 2. The graph is plotted based on number of attributes on X-axis and time in seconds on Y-axis.

B. Security and Privacy Analysis

In our paper we propose a framework that combines the decentralized storage system IPFS, the Hyperledger blockchain, and attribute-based encryption (ABE) technology to achieve finegrained access control over data in decentralized storage systems. After implementing system we get number of benefits over traditional data storing and sharing system. These benefits are illustrated below:

1. Data owner controls their own data

In proposed scheme, data owner generate keys for files and user. Thus the PKG is not needed in our scheme. Also keys are stored in blockchain thus misuse of keys is not possible. Also files are stored in decentralized storage. The security, availability and reliabilityof data is guaranteed by the smart contract.

2. Avoid single point of failure

In proposed scheme, we make use of decentralized storage system, IPFS. The nodes in IPFS are located though out the world. IPFS ensure the availability of data by replicating the files. As the number of access to file increases the number of replica of file also increases. Thus a single point of failure does not result in failure of the system in returning file.

3. Search fairness

In proposed scheme we ensure fairness of search action by letting process to be handled by smart contract. This does not let user to get the link for inaccessible file.

4. Security and privacy

In our scheme we use decentralized storage, IPFS for storing file. The nodes in IPFS are distributed in nature. Thus the data stored on node need to be secured. We achieve this by encrypting files using attribute based encryption technique. This avoid unauthorized access to data.

V. CONCLUSION

In traditional cloud storage user data may be unavailable because of force majeure. The ABE technology is an important tool for achieving fine grain access control and data privacy. There is always need of PKG in all ABE scheme. The PKG is not flexible enough and end into key abuse. The risk of single point of failure in traditional storage can be overcome by using decentralized storage. In addition decentralized storage has advantages such as high throughput, low price etc.

In this paper we propose a system to combine decentralized storage system, blockchain and ABE technology. In this system there is no need of PKG. Thus the problem of key abuse is solved. We also study interplanetary naming system used in IPFS. In IPNS there is a pointer to current version of file. So whenever a user make request to previous version of file in IPNS, it automatically redirect to latest version of file. This make user to download unauthorized file. This results in failure of attribute revocation functionality. An attribute based search function helps to overcome this.

From performance and security analysis, we conclude that our system is feasible to implement.

REFERENCES

[1] A.Sahaiand, B.Waters,‘‘Fuzzyidentity – basedencryption,’’inProc.Annu. International Conference Theory Application Cryptography Technology Berlin, Germany: Springer, 2005, pp. 457–473.

[2] J. Zhang, X. A. Wang, and J. Ma, ‘‘Data owner based attribute based encryption,’’ in Proc. International Conference Intell. Network Collaborative System (INCOS), Sep. 2015, pp. 144–148.

[3] A Decentralized Network for Internet of Things. Accessed: Mar. 25, 2018. [Online]. Available: https://iotex.io

[4] Blockchain for Supply Chain. Accessed: Mar. 25, 2018. [Online]. Available: https://www.ibm.com/blockchain/supply-chain.

[5] G. Zyskind, O. Nathan, and A. S. Pentland, ‘‘Decentralizing privacy: Using blockchain to protect personal data,’’ in Proc. Security Privacy Workshops (SPW), May 2015, pp. 180–184.

[6] Y. Rahulamathavan, R. C.-W. Phan, M. Rajarajan, S. Misra, and A. Kondoz, ‘‘Privacy-preserving blockchain based IoT ecosystem using attribute-based encryption,’’ in Proc. IEEE International Conference Advance Network Telecommunication System, Dec. 2017, pp. 1–6.

[7] H. Es-Samaali, A. Outchakoucht, and J. P. Leroy, ‘‘A blockchain-based accesscontrolfor bigdata’’International Journal of Computer Network and Communication Security,vol.5, no. 7, pp. 137–147, 2017.

[8] A. Kapadia, P. P. Tsang, and S. W. Smith, ''Attribute-based publishing withhiddencredentialsandhidden policies''inProc. NDSS,vol.7,2007, pp. 179–192.

[9] M. Chase, ''Multi-authority attribute based encryption'' in Proc. Theory Cryptography Conference Berlin, Germany: Springer, 2007, pp. 515–534.

[10] J. Li, Y. Wang, Y. Zhang, and J. Han, ''Full verifiability for outsourced decryption in attribute based encryption,'' IEEE Tranaction Services Comput.

[11] D. X. Song, D. Wagner, and A. Perrig, ''Practical techniques for searches on encrypted data,'' in Proc. IEEE Symp. Security Privacy, May 2000, pp. 44–55.

[12] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, ''Public key encryption with keyword search,'' in Proc. International Conference Theory Application Cryptography Techn. Berlin, Germany: Springer, 2004, pp. 506–522.

[13] D. Boneh and B. Waters, ''Conjunctive, subset, and range queries on encrypted data,'' in Proc. Theory Cryptography Conference Berlin, Germany: Springer, 2007, pp. 535–554.

[14] H. Li, F. Zhang, J. He, and H. Tian. (2017). ''A searchable symmetric encryption scheme using blockchain.'' [Online]. Available: https://arxiv.org/abs/1711.01030.

[15] C.Cai,X.Yuan,andC.Wang,''Towardstrustworthyandprivatekeyword search in encrypted decentralized storage,'' in Proc. IEEE International Conference Communication (ICC), May 2017, pp. 1–7.

[16] H. G. Do and W. K. Ng, ''Blockchain-based system for secure data storage with private keyword search,'' in Proc. IEEE World Congr. Services (SERVICES), Jun. 2017, pp. 90–93.

[17] J. Benet. (2014). ''IPFS-content addressed, versioned, P2P file system.'' [Online]. Available: https://arxiv.org/abs/1407.3561.