Scientific Research Publishing

Analysis of Blockchain Architecture in File Sharing Management for Tertiary Institution

Kanayo Kizito Uka, Stanley Ikechukwu Oguoma, Udochukwu Princewill Chuma-Uba

Department of Computer Science, Imo State University, Owerri, Nigeria
Email: kizzyict@gmail.com, oguomaikechukwu@gmail.com, michaelvalidprince@gmail.com

Abstract

The aim of this paper is to analyze the sharing and management of files in a tertiary institution using blockchain architecture. It is expected to enable an online system that could provide a decentralized architecture for multiple transfers and sharing of files amongst participants, to design a system that can provide data integrity and security of files using IBM Blockchain technology, to provide a system that can allow multiple user and multiple transactions at once. The research was motivated because of the security challenges associated with existing system which include: delay in transfer and sharing of files, much authority and task given to a single user in a centralized system, high risk of attack and loss of files, reduced speed of file transfer and file access/retrieval, limited number of users per time. Methodology adopted was Object Oriented Analysis Design Methodology (OOADM) in conjunction with Unified Modeling Language (UML) and IBM Blockchain Technology while the programming language used was HTML, CSS, Java and Node Js. The result after design was a decentralized cloud based file sharing and management system that enables multi-shared, replicated and permissioned transactions amongst participants in a network.

1. Introduction

File sharing is the practice of distributing or providing access to digitally stored information. Though, various users and organizations have long relied on the internet as a medium of sharing or publishing data with each other. Most often, data warehouse is provided to store large files and is adopted by organization, individuals, but has been identified as associated with huge risk, because it is

managed by a single user. Examples of such repository technologies include Blogs [1]. Other technologies that spring up because of the rise of Web 2.0 encourage multiple data sharing among individuals. Flicker permits its users to upload, organize and share photos. Wikipedia as one of the most popular online encyclopedias permits its users to share knowledge by modifying the encyclopedic entries. Also, Google Groups provides repository where users can post messages and share files. Hence various organizations require a more secured technology that could grant confidence and trust among participants. Blockchain technology stances the best option as it contains a ledger which is often described as decentralized, because it is replicated across many network participants, each of whom collaborate in its maintenance. Its decentralization and collaboration are powerful attributes that mirror the way businesses exchange goods and services in the real world. Information recorded to a Blockchain is append-only, using cryptographic techniques that guarantee that once a transaction has been added to the ledger it cannot be modified hence granting information immutability. However, centralized solutions are unwelcome sometimes because of privacy concerns and censorship, which are problems that can be alleviated by switching to decentralized solutions. Therefore, this paper contributes an Analysis of Blockchain Architecture in File Sharing Management for Tertiary Institution to provide a secured data repository that can support a variety of collaborative file sharing and management among various participants in the Institution.

In summary, the proposed system contributed immensely in the aspect of security and confidentiality of files within an organization by the adoption of IBM Blockchain architecture. The use of IBM Blockchain architecture makes it possible for the Network to have two keys (private and public keys) for each participant on a network. Before a participant could use the platform in sharing or reading file from one department to another, an encryption or decryption process must be done by the participant with one of the provided keys which could serve as a block in the network, once all the participants in the network receive the file (block), an endorsement must be granted by all the participants and the file could be read by each of them with the key (private key) in this case. All files are stored in the IBM Blockchain ledger which is referred as the storage. The organization of the rest of this paper includes: literature review, methodology adoption, analysis of the existing system, analysis of the proposed system, proposed system architecture, process of the proposed system in sharing file by users, consensus algorithm (Proof-of-work) use case model of various users, high level model of the proposed system, conclusion and recommendation.

2. Related Literature

The Internet has changed the way people communicate and distribute information in all parts of the world. In 2005, about one billion people used the Internet [2] to exchange emails and instant messages, access information on web servers,

and ever increasingly share files through various networks [3]. The emergence of P2P networks makes it possible for Internet users to be simultaneously sharing and accessing digital content in many forms like audio and video files. File sharing techniques like the Peer-to-Peer (P2P) file sharing networks started with the Napster phenomenon, reaching 20 million users who downloaded hundreds of millions of files before it was shut down in 2000 [4], due to copyright infringement [5]. Napster had a centralized directory with the list of shared files (music files) of every IP address (host) in the network. This centralized directory was updated constantly as hosts connected or disconnected from the network. Hosts connected to the network used the centralized directory to search for a file. The file search result contained the list of hosts from which the searched file could be directly downloaded. Therefore, the searched file was directly downloaded from the host sharing the file. Another File sharing device enveloped containing the architecture of a decentralized system for sharing music files known as Gnutella in the year 2000 [6], to replace Napster. Gnutella offered a new distributed search paradigm that many popular P2P file sharing networks like Limewire [3], use today. These networks are overlays constructed on top of the Internet and account for 20% of the total aggregated traffic of the Internet [3]. An overlay network is a logical network on top of another network (e.g., on top of the Internet).

2.1. Spreading and Sharing Personal Information

With technology developing, people can collaborate with more devices, regardless of different timing and locations. Hence, it is potentially troublesome to a team project if files are scattered in different places, with many different versions. How to handle this complexity is an emerging issue both in personal information management (PIM) and related technology design. [7] research suggested out that file synchronization would be important in future technical developments because of users' needs for file integration and information sharing among individuals and groups. Since the human factor was getting involved more and more in system design, they suggested that we should consider and focus on users, instead of devices, because it is humans who create and then adopt the devices. [8] indicated that in order to help users working in collaboration, the system and shared workspace should provide awareness information to remind users and allow them work together with high flexibility. On the other hand, because of the complexity of cooperation, [9] discussed how to present awareness information efficiently in a sharing workplace. They found that some current systems might not clearly show indications of what is being shared to users, and it confused people when they could not distinguish shared items in the workspace. Without clear indicators of sharing, people usually felt frustrated when they recognized and managed shared files. Therefore, how to deal with shared and non-shared items on the same platform has become a rising topic, and it deeply impacts technology development and interface design in personal

information management.

2.2. Peer-to-Peer File Sharing Networks

In recent works, much attention has been focused on peer-to-peer applications. A review by [10] suggested that P2P applications be divided into three dimensions: file sharing, distributed computing and communication & collaboration. Nonetheless, distributed computing and communication & collaboration have not been developed well due to some technical problems which are hard to be conquered. On the other hand, P2P file sharing systems are very successful. They introduced the distributed hash table [11], technology to their system to provide the lookup operation: given a key, Chord system maps the key to a node. Data location is easily implemented in Chord by associating a key with each data item and storing the key-data pair at the mapped node. However, Chord does not support keyword lookup. Recognizing this limitation, [12] established an architecture for content-based publish/subscribe services that perfectly solved the problem. For resource transportation function, [13] investigated a novel streaming architecture consisting of home-to-home online devices that collaborated with one another to provide on-demand access to large repositories of continuous media such as audio and video clips. Furthermore, [14] proposed a hybrid P2P file sharing system called An Efficient Implementation of File Sharing Systems on the Basis of WiMAX and Wi-Fi which was aware of the underground network topologies. End devices under the same SS are connected with Wi-Fi technology, *i.e.*, every two end hosts are one-hop connected and each end host in the same local area network is connected directly to the SS. At the same time, they propose a changed dynamic source routing (CDSR) lookup protocol on SSs mesh networks where the number of SSs under a BS is not large enough to efficiently construct a DHT based file sharing architecture. Moreover, by implementing CDSR on SSs mesh networks, there layered file sharing system can support keyword lookup within a SSs mesh network and also designed DHT based file lookup architecture on BSs networks, because BSs networks are cable connected and able to tolerate the extra traffic loads. [15] proposed a Friend share architecture which was a decentralized multiple-writer data repository where the data is distribute and replicated across multiple nodes in the network. [1] proposed an android file sharing system which was implemented with Java a general-purpose, concurrent, class-based, object-oriented computer programming language that is specifically designed to have as few implementation dependencies as possible, Android SDK which enables developers to build powerful programs in a neat, object oriented design.

2.3. Overview of an IBM Blockchain Architecture

Blockchain network contains a distributed ledger that records all transactions that take place on a network. Therefore a blockchain ledger is often described as decentralized because it is replicated across many network participants, each of

whom collaborates in its maintenance. It is also noted that decentralization and collaboration are powerful attributes that mirror the way businesses exchange goods and services in the real world. In addition to being decentralized and collaborative, the information recorded to a blockchain is append-only, using cryptographic techniques that guarantee that once a transaction has been added to the ledger it cannot be modified. This property of "immutability" makes it simple to determine the provenance of information because participants can be sure information has not been changed after the fact. It's why blockchains are sometimes described as systems of proof. IBM Blockchain is a fully integrated enterprise-ready blockchain platform that is designed to accelerate the development, governance, and operation of a multi-institution business network: With IBM Blockchain, developers could use the tool called Hyperledger Composer to quickly build a blockchain application which is highly secure and per formant on IBM Cloud environment.

2.3.1. The Ledger

A distributed ledger is a type of database, or system of record, that is shared, replicated, and synchronized among the members of a network: The distributed ledger records the transactions, such as the exchange of assessor data, among the participants in the network. This shared ledger eliminates the time and expense of reconciling disparate ledgers. The IBM Hyperledger is shown in Figure 1. Every record in the distributed ledger has a time stamp and unique cryptographic signature, thus making the ledger an auditable history of all the transactions in the network. In the context of Hyper-ledger Fabric, the ledger is composed of two components: The blockchain itself, and the world state database [16].

Table 1 illustrates the major difference between the proposed system and the related literatures. A close look into the above table, one could find out that the existing researchers adopted a different technology while the proposed system improved the research gap by adopting the Hyperledger Fabric IBM Blockchain technology.
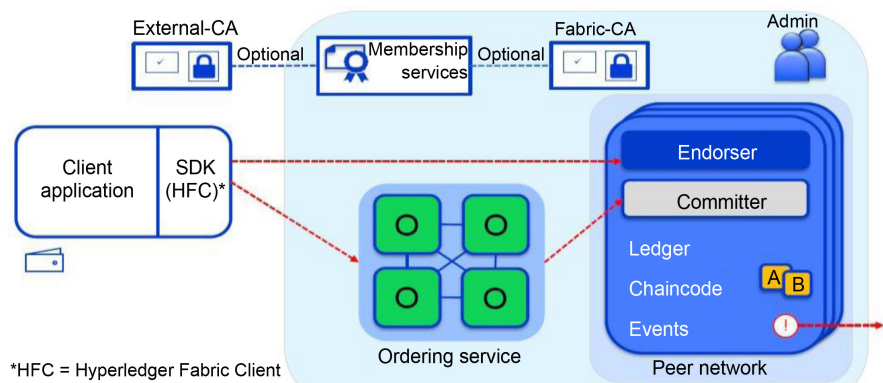

Figure 1. IBM hyperledger fabric architecture. Source: IBM Blockchain 2019 (Source: [17]).

Table 1. Major difference between the proposed system and the related literatures.

S/N	Related Literatures	Used Technology	Proposed System	Used Technology
1	Peer-To-Peer File Sharing Networks	• Introduced a distributed hash table. • Centralized directory with the list of shared files.	Analysis of Blockchain Architecture in File Sharing Management For Tertiary Institution.	• Hyperledger Fabric IBM Blockchain Network architecture • Decentralized Architecture
2	Hybrid P2P file sharing system	• Operates on the Basis of WiMAX and • Wi-Fi which work on the underground network topologies.	Proposed System	• Operates in a Decentralized Platform so as to effectively manage users stored files.
3	Friendshare architecture	• A decentralized multiple-writer data repository	Proposed System	• A decentralized multiple-writer data repository but with consensus. • Append-only, using cryptographic techniques that guarantee that once a transaction has been added to the ledger it cannot be modified
4	android file sharing system which was implemented with Java a general-purpose	• Android Mobile Technology. • Accessed only on a mobile platform.	Proposed System	• Enhanced security measures as every record in the distributed ledger has a time stamp and unique cryptographic signature. • Could be accessed remotely with both mobile and large devices

3. Methodology

Object-Oriented Analysis and Design Methodology (OOADM) was adopted for the analysis and design of this work following the three development life cycles.

    1) Object-Oriented Analysis

    2) Object-Oriented Design

    3) Object-Oriented Implementation

It uses the unified modeling language (UMLs) to model the various users and

their privileges, data flows diagram and high level model diagram.

3.1. Analysis of the Existing System

File sharing at the management of tertiary institution require efficient security and enhanced measure to ensure that file leakage or spread of relevant file are prevented. Imo State University Owerri was adopted as the case study, during the investigation on how files are shared among the institution management; it was observed that the institution adopted a traditional way of processing and sharing files, which hence automatically affects productivity. Most times staffs uses email attachment or other files sharing technology like flash-share, zender android application etc. to send files. But during the analysis of this work, it was identified that there is security challenge in sharing sensitive files to members within the institution, sometimes, shared files are manipulated by members or staff of the institution. On the other hands, the institution employed a lot of messengers in various offices; they are the ones in charge of sending message or carrying files from one office to another. Most times these files are mishandled by them, some vital information that does not need to get out are exposed through this form. Hence there is need to design a secured file sharing and management system that will improve the distribution or sharing of documents (files) within different departmental heads in the institution. Figure 2 below shows the processes of the existing file sharing system within the institution.

3.2. Analysis of the Proposed System

This paper proposed an analysis of Blockchain architecture in file sharing management for tertiary institution which works in a peer-to-peer network. The
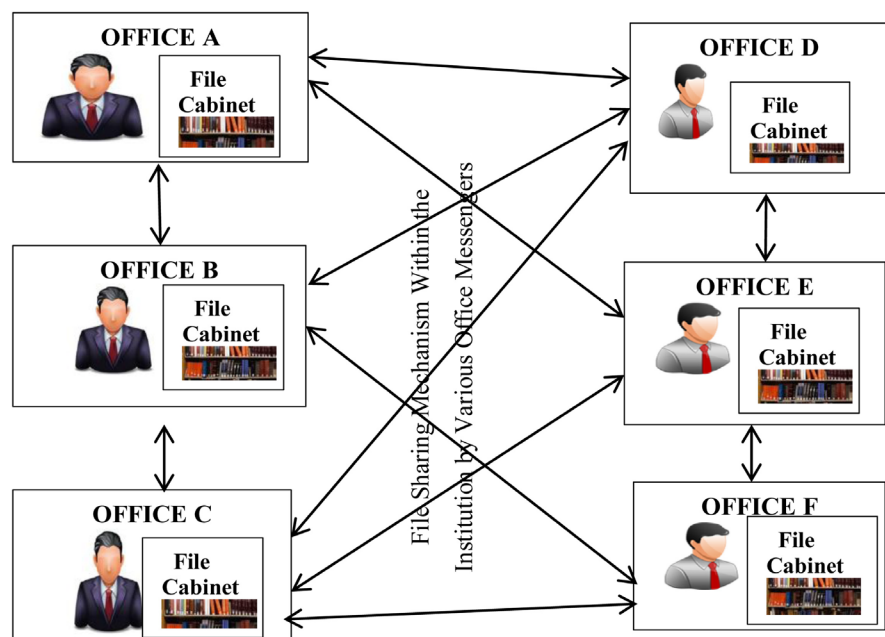

Figure 2. Existing file sharing mechanism from various offices.

software will enable organization to transfer confidential files or document among participatory employees to communicate remotely, and hence improve service delivery and productivity. The proposed system will adopt IBM Blockchain architecture so that file security could be securely achieved and to ensure that only required participants concerned in the institution have access to those files and before a file could be altered, it must be validated and append permission by the participant in the network. Further, Blockchain architecture works in a decentralized ledger, whereby information storage is done on a global network which is publicly available and safe from tampering by non-participants. The IBM Blockchain Network provides two keys to any participants in the network (private and public keys). Once a participant wants to share a file to another department or office (participants), it will be encrypted first with the public key and this process is treated as a block, and then the file is broadcasted to all permissioned participants in the network in this case, approved staffs which could read the shared file with their individual private keys. Participants in the network approve the file as valid (consensus) then the file is appended or stored in the Blockchain shared ledger, then the file shared to the other department is updated in the world state. Figure 3 illustrates the Proposed System Hyperledger Fabric IBM Blockchain Network architecture while Figure 4 shows the stored files in a Blockchain shared ledger.

3.3. Proposed System Architecture

3.3.1. Process of the Proposed System in Sharing File by User

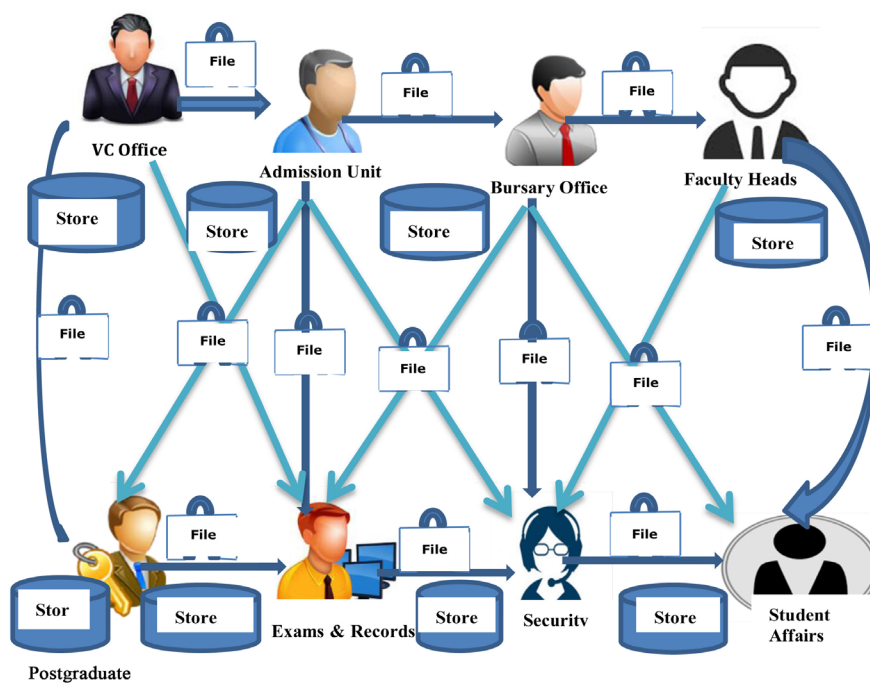The user clicks the icon on the android phone or desktop computer to runs the

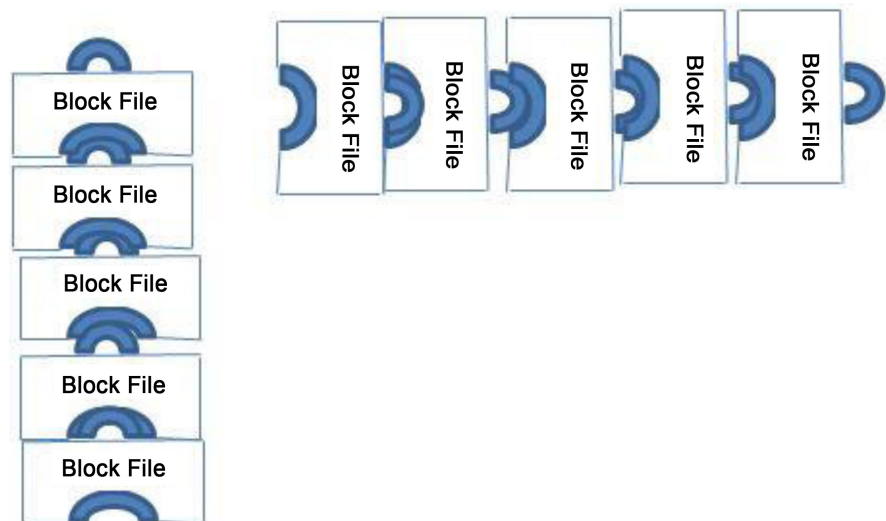**Figure 3.** Proposed system hyperledger fabric blockchain network architecture.

Figure 4. Stored files in blockchain shared ledger.

application, after selecting a file to share and encryption processes with the help of the REST API which then invoke the ordering service that checks and agrees on the transaction sequence and distributes file blocks to permissioned management peers in the hyper-ledger network. Then the participants in the network validate (consensus) the files as valid then the file is appended to the chain within the shared ledger and the world state is updated. **Figure 5** illustrates the process of the Proposed System architecture.

3.3.2. Using Psuedocode to Explain the Proof-of-Work (Pow) Algorithm

✓ Program Start;
✓ User runs application;
✓ REST API Lunched;
✓ Users Login Module Opened;
✓ If (User Login Details Correct?){
✓ Open Dashboard;
✓ User select file to upload;
✓ Encrypt file;
✓ Share file to all participants in the Network;
✓ If(Shared files is Valid){
✓ All Participants Append Validation and Decrypt file
✓ Consensus by all Participants in Network Done;
✓ }else{
✓ Network Collapse;
✓ Message "Files tampered";
✓ }
✓ }else{
✓ Error Message;

- ✓ Invalid User Login Accounts;
- ✓ End if;
- ✓ End program;

Therefore, in other to effectively explain how the Consensus Algorithm (Proof-Of-Work) shown in **Figure 6** was applied into the proposed system architecture, the PSUEDOCODE algorithm was adopted while **Figure 7** illustrates the proposed system uml case diagram for various on the platform and their privileges.
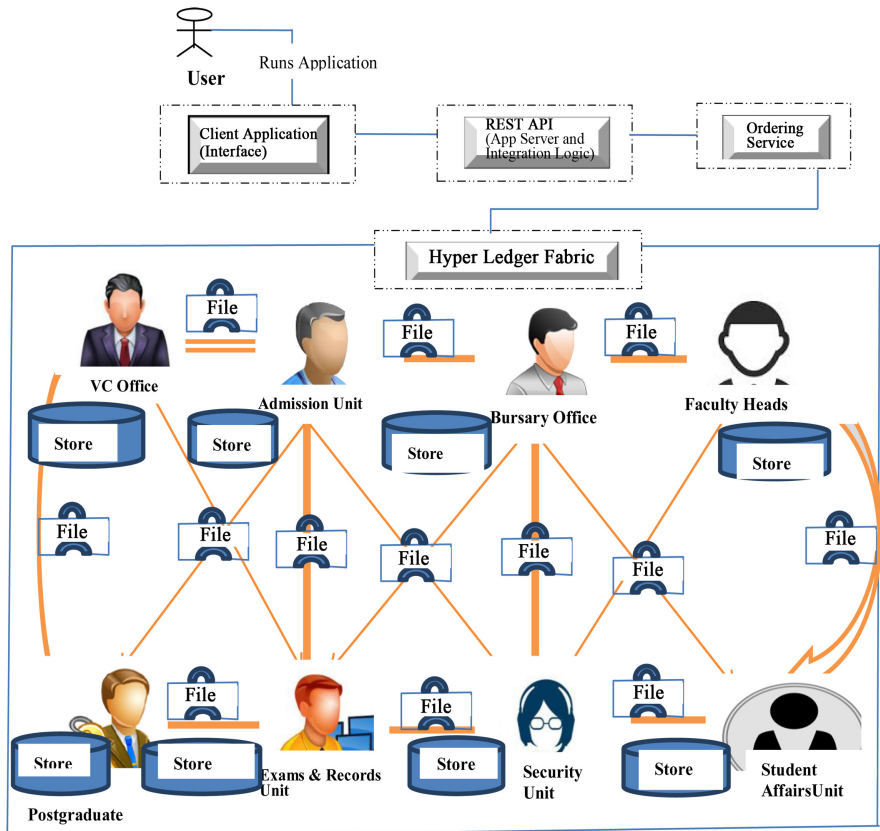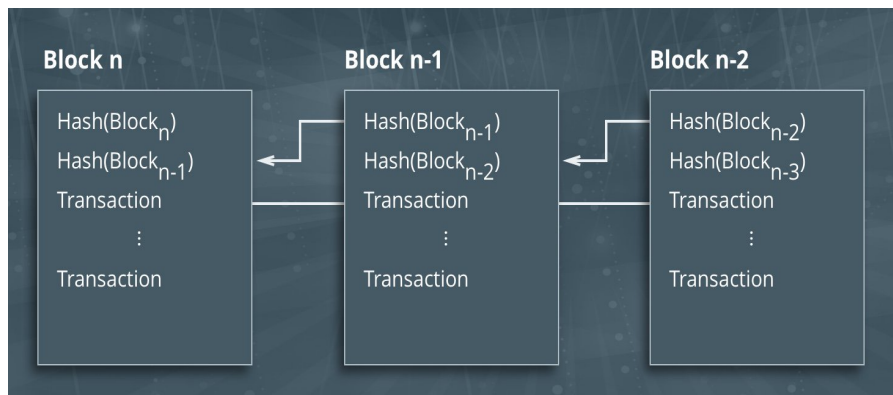

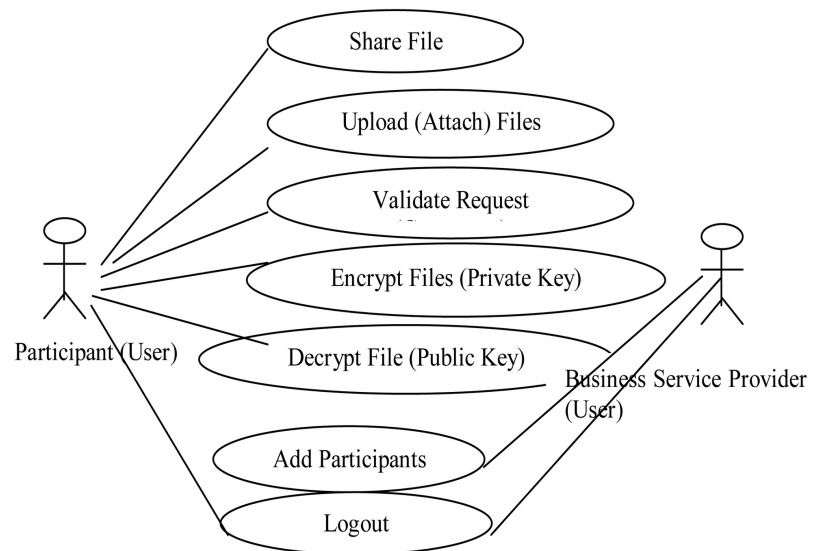**Figure 5.** Proposed system architecture.

Figure 6. Consensus algorithm (Proof-of-Work) (Source: [18]).

Figure 7. The user UML case diagram for proposed system.

4. System Design and Implementations

System design deals with putting solutions to the problems with the available resources. It is the blue print of the system that is to be developed. Therefore, software life cycle requires some basic activities which include: Decomposition of the entire system into components: Identification of the system architecture including various menu and screen displays, how the system can be installed, the hardware and software specifications/requirements for the installation of the system and how to start up the system, Component of the proposed system and their relationship as well as its dependencies are covered, and Inter component communication mechanism are also achieved.

4.1. High Level Model of the Proposed System

The high level model of the proposed system shown in Figure 8 explains the modules and sub-module of the new system. The design adopted a top-down design approach where every module, sub-module and sub-sub module could be shown clearly.

4.2. Choice and Justification of Programming Language Platform

The new system is a decentralized cloud based file sharing and management system that enables multi-shared, replicated and permissioned transactions amongst participants in a network. The front end was designed with hypertext markup language (HTML), Cascading Style Sheet (CSS), Bootstrap, Java Script (JS) while the backend was done with, Object Oriented Java, Node JS and in IBM Cloud Blockchain environment. The system has an easy and simple friendly interface design such as the user only enter the URL in the browser to access the application.
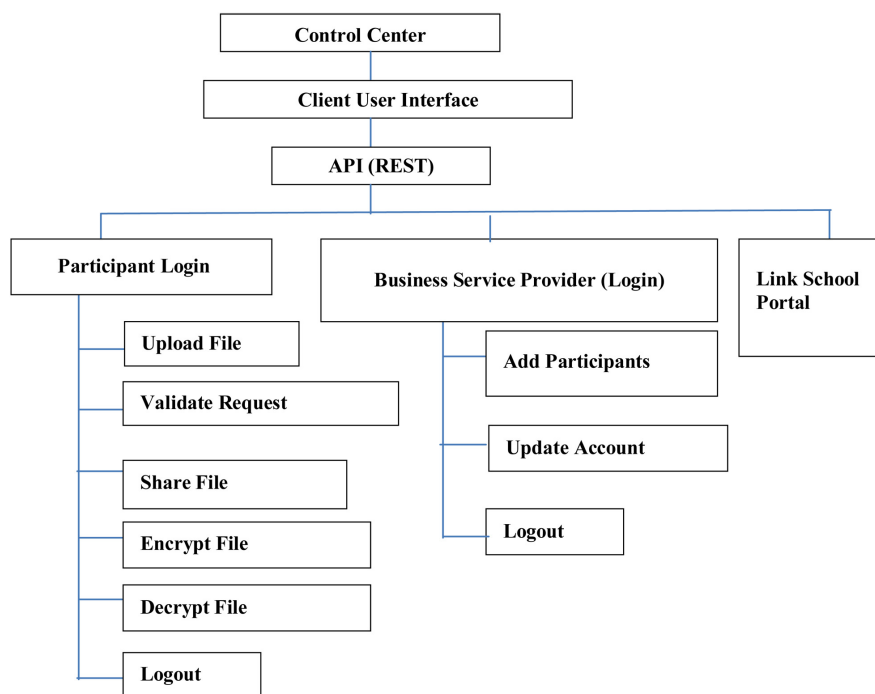
Figure 8. High level model of the proposed system.

5. Results

Figure 9 shows the control center of the new system, this interface shows that before any user of the system could access the content of the new system, an authentication process must be guaranteed.

This shows the dashboard of the system administrator, as stated earlier in this paper, all the participants of the system have no access to update any store data in the shared ledger unless a consensus is mutually done. Hence the role of the administrator in this case only acts when a consensus has been achieved by various users on the platform, shown in Figure 10.

Figure 11 shows the designed interface of the participant dashboard, it is on the users dashboard the file sharing, file encryption and decryption, validation of files request by various participants on the network are done.

Users of the platform can share files by using the interface shown in Figure 12, these interface provides an input for attachment of encrypted files, file title, and body of the text to be shared.

In case of sharing more than one files at a time, the platform provide an interface for users to attach multiple files at a time shown in Figure 13.

As stated in the analysis of the proposed system, files must be first encrypted before sharing; therefore Figure 14 illustrates the file sharing or decryption interface.

Every user or participants on the network can check for shared files but cannot alter or edit the shared files. There activities could be achieved by the displayed interface in Figure 15.
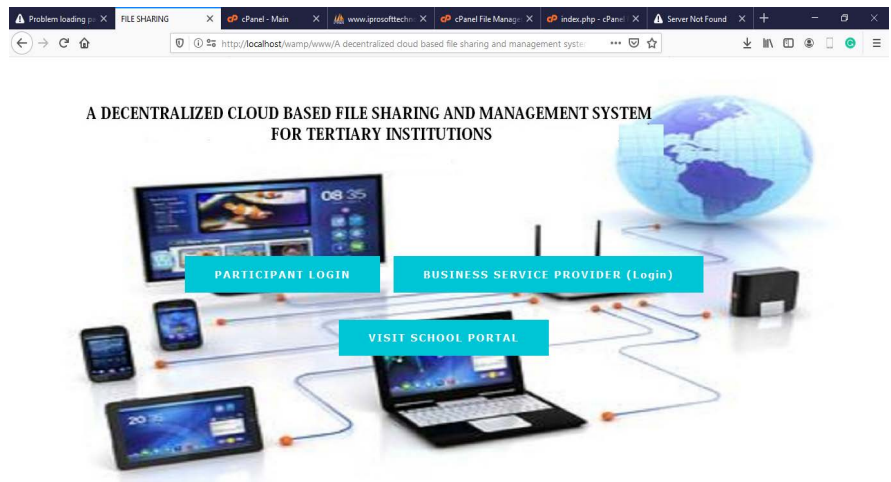
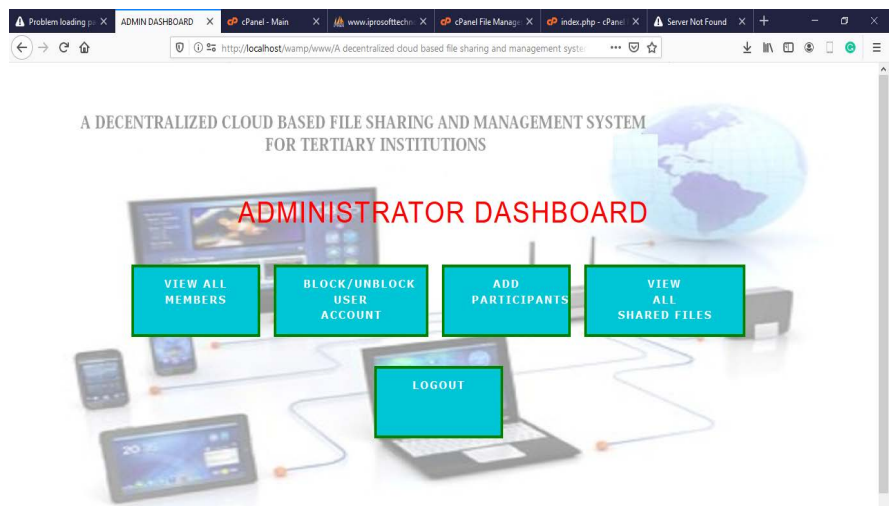**Figure 9.** System control center interface.

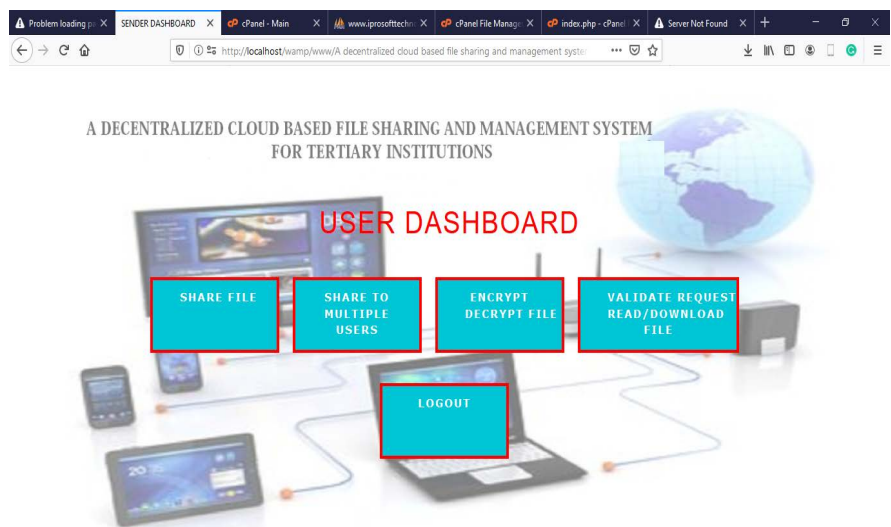**Figure 10.** Administrator user dashboard interface.

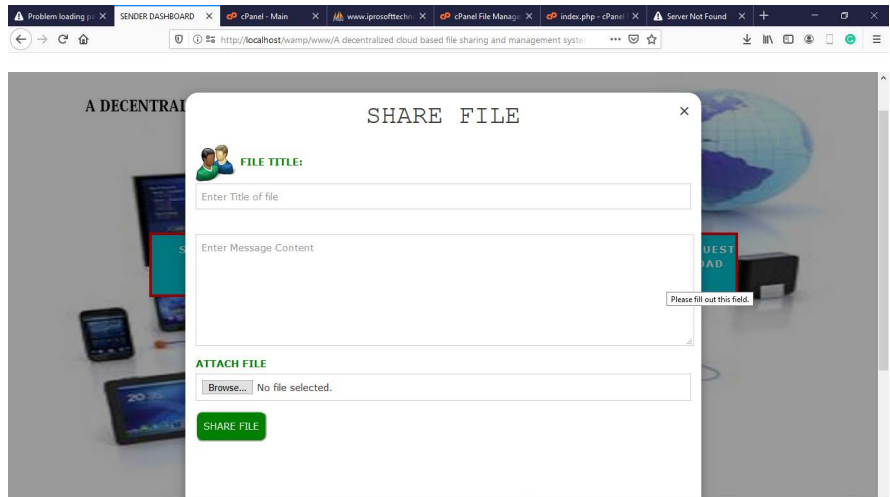**Figure 11.** Participant dashboard interface.

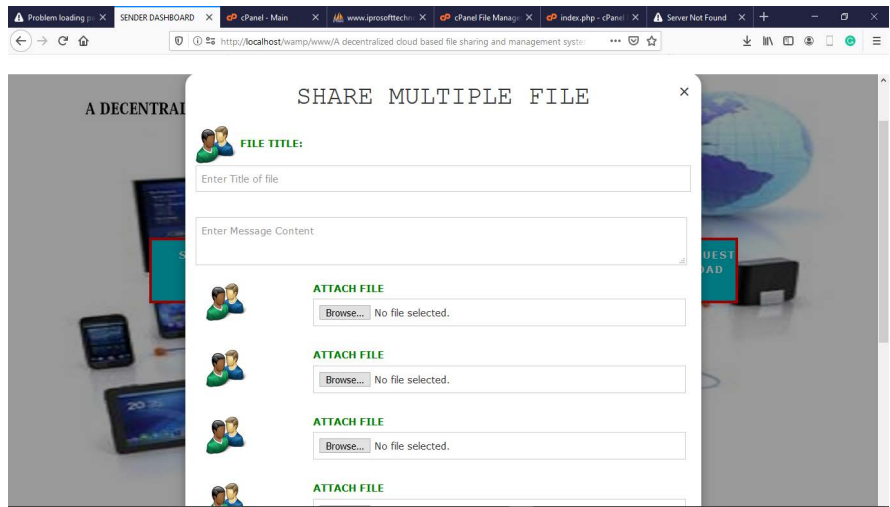**Figure 12.** User file sharing interface.

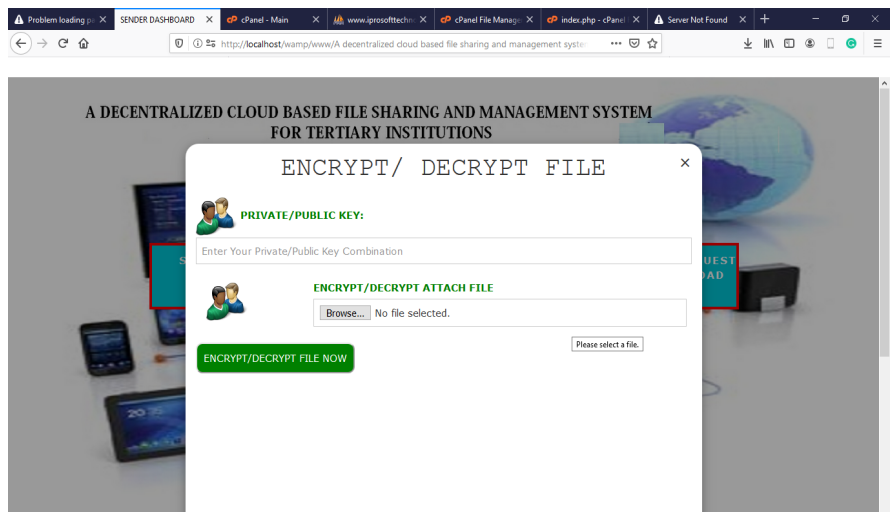**Figure 13.** User multiple file sharing interface.

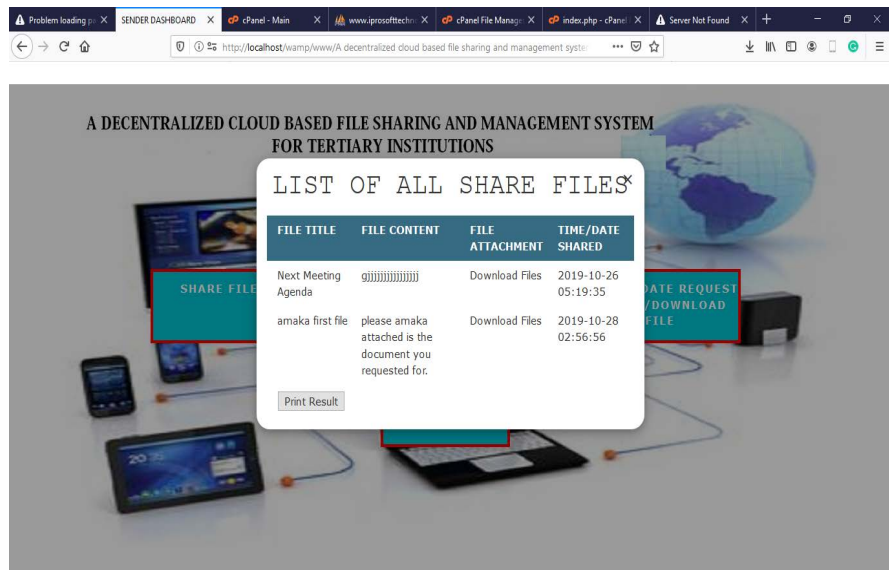**Figure 14.** User file encryption and decryption interface.

Figure 15. List of all share files (user validation request) interface.

To access the platform, user authentication must be done, which means that every user must provide login details. **Figure 16** illustrates the login access to the platform.

Figure 17 shows the interface used to add participants into the system during the creation of the system.

6. Conclusions

IBM Blockchain integration in various organizational activities will encourage trust; improve security, and confidence among participants in business and transactions. It is off uttermost interest of various higher educational institutions to improve the security of their files or record and hence prevent it from unauthorized access and fraudulent personnel's at all times. Therefore implementing Blockchain architecture in the educational activities to secure files and sharing activities among staffs of the institution is highly important.

Hence, this analysis of Blockchain integration in the file sharing and management of tertiary institution yield a great result not in the aspect of data security only but in strengthening trust and confidence in decision making at the management level. This analysis also shows that virtually, Blockchain technology could be adopted in other educational areas such as in school fees payment, accurate and secured record keeping, teaching and learning and student's transcript and grading system. In summary, the researcher proposed that Blockchain technology be employed in the educational system so as to improve the standard of information processing and sharing in the sector.

Recommendations for Further Studies

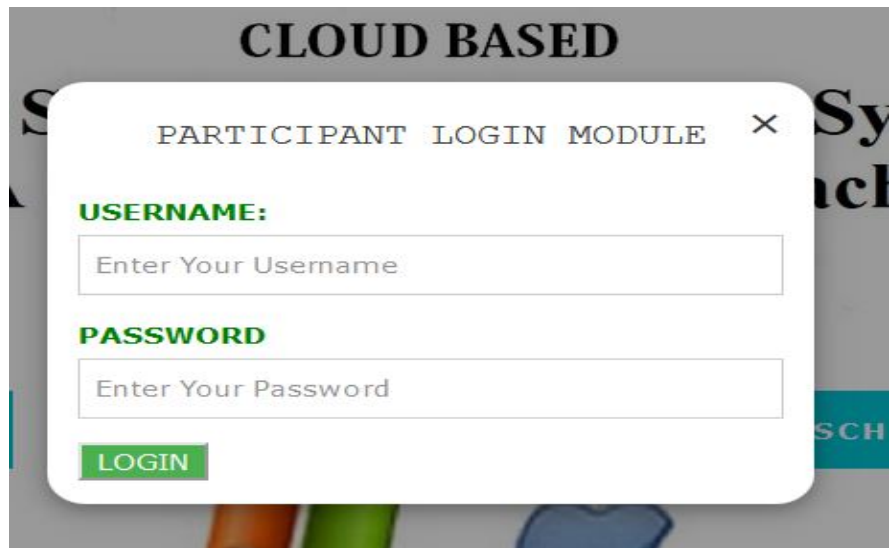Therefore, this study recommends the following:

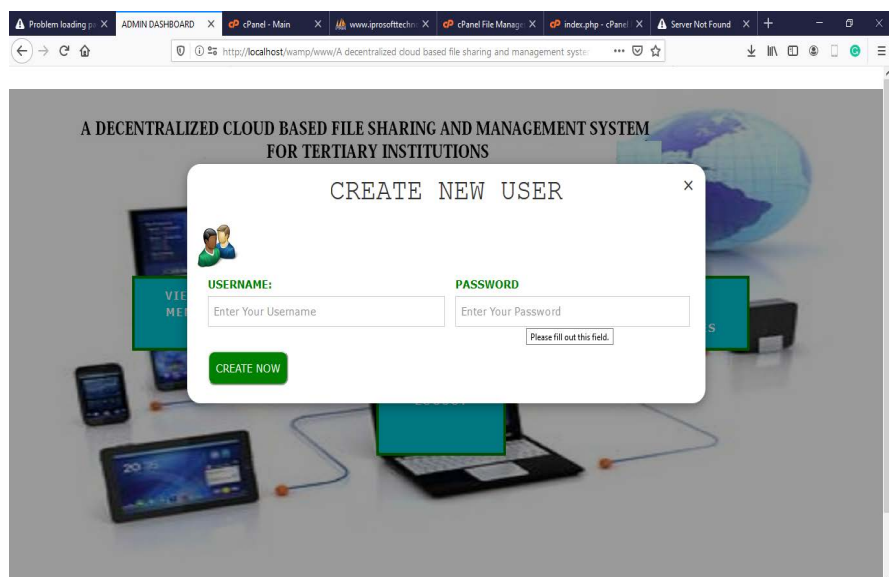**Figure 16.** Participants login interface.

Figure 17. Add Participant interface.

1) Full implementation of IBM Blockchain cloud based System be developed using node JS, and Angular.

2) Other researchers could design the system using a different Blockchain Consensus Algorithm Model such as proof-of-Activity (POA), etc.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Alatishe, A.A., Adegbola, M.A. and Dike, U.I. (2013) File Sharing Application for Android. *International Journal of Computer Science Engineering*, **2**, 20.

[2] Rajendran, R.K. and Rubenstein, D. (2004) Optimizing the Quality of Scalable Video Streams on P2P Networks. *Proceedings of GLOBECOM*, **2**, 953-959. https://doi.org/10.1145/1005686.1005735

[3] Karagiannis, T., Broido, A., Brownlee, N., Claffy, K. and Faloutsos, M. (2004) Is P2P Dying or Just Hiding? *Proceedings of GLOBECOM*, **3**, 1532-1538.

[4] Cable News Network (2000) Napster Shutdown Seen as Potential Boon for Competitors. http://archives.cnn.com/2000/LAW/07/27/napster.backlash

[5] Kurose, J.F. and Ross, K.W. (2001) Computer Networking. 2nd Edition, Addison Wesley, Boston, MA.

[6] Yang, B. and Garcia-Molina, H. (2001) Comparing Hybrid Peer-to-Peer Systems. In *Proceedings of the* 27*th International Conference on Very Large Databases*, Roma, 11-14 September 2001, 561-570.

[7] Dearman, D. and Pierce, J. (2008) It's on My Other Computer! Computing with Multiple Devices. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (*CHI* 2008), ACM Press, New York, 767-776. https://doi.org/10.1145/1357054.1357177

[8] Dourish, P. and Bellotti, V. (1992) Awareness and Coordination in Shared Workspaces. In: *Proceedings of the* 1992 *ACM Conference on Computer-Supported Cooperative Work* (*CSCW*, 1992), ACM Press, New York, 107-114. https://doi.org/10.1145/143457.143468

[9] Whalen, T., Toms, E.G. and Blustein, J. (2008) Information Displays for Managing Shared Files. In: *Proceedings of the* 2*nd ACM Symposium on Computer Human Interaction for Management of Information Technology*, ACM Press, New York, 5. https://doi.org/10.1145/1477973.1477980

[10] Milojicic, D.S. and Kalogeraki, V. (2002) Peer-to-Peer Computing. Technical Report HPL-2002-57, HP Lab.

[11] Burr, W.E. (2006) Cryptographic Hash Standards: Where Do We Go from Here. *IEEE Security and Privacy Magazine*, **4**, 88-91. https://doi.org/10.1109/MSP.2006.37

[12] Zhu, Y. and Hu, Y. (2004) Ferry: An Architecture for Content-Based Publish/Subscribe Services on P2P Networks. 2005 *International Conference on Parallel Processing*, Oslo, Norway, 14-17 June 2005, 427-434.

[13] Ghandeharizadeh, S., Krishnamachari, B. and Song, S. (2005) Placement of Continuous Media in Wireless Peer-to-Peer Networks. *IEEE Transactions on Multimedia*, **6**, 335-342. https://doi.org/10.1109/TMM.2003.822787

[14] Li, J.Y., Huang, L.S., Jia, W.J., Xiao, M.J. and Du, P. (2006) An Efficient Implementation of File Sharing Systems on the Basis of WiMAX and Wi-Fi. Joint Advanced Research Institute of City University of Hong Kong and Suzhou University of Science and Technology, Suzhou, China, 9-20. https://doi.org/10.1109/MOBHOC.2006.278658

[15] Frank (2008) Overview of the Friendshare Architecture.

[16] IBM Blockchain (2019) The Ledger. IBM Skills Academy. https://www.ibmskillsacademy.com/

[17] IBM Hyperledger Fabric Architecture. Source: IBM Blockchain. https://www.ibmskillsacademy.com/

[18] Consensus Algorithm (Proof-of-Work). https://cointelegraph.com/explained/proof-of-work-explained