Blockchain-enabled Big Cloud File System(B2CFS): An Inventive System for Secured Record Keeping

Mrs. V. L. Badadare[1]
Assistant Professor
D Y Patil Agriculture and Technical University Talsande
Kolhapur (Maharashtra).
vidya.badadare@dyp-atu.org

Dr. Varsha P. Desai[2]
Assistant Professor
D Y Patil Agriculture and Technical University Talsande
Kolhapur (Maharashtra).
varshadesai@dyp-atu.org

Dr. A. M. Shaikh[3]
Head
Dept. of Electronics
The New College,
Kolhapur (Maharashtra).
amshaikh7@gmail.com

ABSTRACT

In this digital era, it's a very astonishing pace for amalgamation of Information technology in all the sectors. One of the novel IT applications implemented in almost all the sectors is Recordkeeping. But implementation might be using different technologies as per the user requirement. So, every recordkeeping application has its own pros and cons. Those pros and cons actually come from technology used or the way it has been implemented. Nowadays the major concern in the context of computerized recordkeeping is security. So, just having electronic recordkeeping system is not enough. The recordkeeping system must be secure, it must be efficient. In the present paper, authors have proposed the recordkeeping system by means of integration of various cutting-edge technologies: Blockchain, Big Data and Cloud Computing.

Keywords— Blockchain Technology, Secured Recordkeeping, B2CFS

I. INTRODUCTION

At the bottom-line recordkeeping system revolves around capturing, organizing, accessing, and protecting the data as well as retaining historical, and destroyed data. Any Secure recordkeeping system indicates: The basic CIA triad that is basic pillars for any sort of security, non-repudiation, Authorization. Apart from these security pillars, some other crucial factors are there affecting efficiency of recordkeeping. Duplication of records is one of those factors.

Duplication of records which will lead to inefficient storage management. Because duplication of records leads to memory wastage. And the novel solution is content addressing. In content addressing, all contents are uniquely identified by their cryptographic hash, including links. So, content addressing will ensure tamper proof data, prevention of double spending. The proposed system B2CFS will be the most efficient solution to overcome various challenges regarding Record management. The system is an integration of different technologies like: Cloud Computing, Big Data and Blockchain [1].

II. LITERATURE REVIEW

Authors put into effect the concept of Trusted TSS (Time Stamping Service), linking and Distributed Trust. For secure Record-Keeping, TSS maintains the date and time the record was received. This is the naive solution for tamper-proof data storage and access. To overcome some issues raised by TSS of digital documents like privacy, bandwidth and storage, Incompetence, and Trust; the authors illustrated two enhancements of TSS. One by adjoining Hash functionality while the other by utilizing a digital signature. Subsequently, the authors presented another two Time-Stamping Schemes: linking and Distributed Trust. The linking mechanism allows the linking of the TSS of a new digital document with the previous one. In the Distributed trust technique, all users can sign the documents by using a Standard secure pseudorandom generator. These techniques will avoid fake timestamps [2].

After reviewing the application of blockchain technology in the Healthcare sector for Secured and Efficient

Patient Record Management, authors concluded that the development of a blockchain network for EHR management would bring the Indian healthcare profession to a viable phase [3].

In the present article authors introduced the incorporation of Blockchain technology into Big Data. They attempted to raise a systematic review of the interactions between big data and cryptocurrency. The authors reviewed numerous applications of big data and cryptocurrency as below:

1. Integrating the cryptography method of cryptocurrency into the E-payment system can effectively revamp security and privacy preservation issues.
2. Decentralization retains direct transactions between buyer and seller[4].

described numerous fundamentals of Blockchain and Big Data technologies. Also explained blockchain use cases in big data with the following specifications:

1. The immutability feature will ensure trustworthiness.
2. Verification of transaction validity as well as computational power by using a consensus protocol
3. Big Data offers predictive analysis and Real-time data analysis
4. Blockchain supports efficient data sharing [5]

brought the spotlight on one of the important aspects in the healthcare sector which is Big Data Analytics. A huge amount of health data is generated through various resources like mobile biometric sensors, and wearable, and smartphone apps. So, it is essential to store and analyze such a large amount of data as this analysed information will provide further insights. They also reported some benefits of EHR management, like proper EHR management, which will provide enhanced access to the Patient's medical history [6].

(Atkuri, 2022) As per the Statistical report stated by Statista, Data loss or leakage is perceived by 64% of respondents as the main cloud security threat in 2021. Data encryption is accountable for the protection and secure databases that blockchain technology offers. The same Statista survey indicates that 44% of the respondents have witnessed visibility and transparency as a severe risk to cloud security [7].

(Rovnaya, 2023) Cloud Computing provides on-demand, anytime, and anywhere metered services to users. Along with such characteristic features, the biggest challenges of cloud computing are as below:

- Cloud services are more prone to cyber attacks
- Centralized management of Cloud Resources [8]

Integration of blockchain with Cloud computing strengthens the potential of cloud capabilities. Blockchain technology amplifies the security by means of cryptography and immutability. It improves visibility as well as traceability.

The above-portrayed literature survey interprets that, the major challenges in existing Record Management Systems are:

- Handling Overloaded Information
- Data Security, Data accuracy & integrity.
- Managing digital data across multiple locations.

III. CRAFTING B2CFS ARCHITECTURE

The present section includes role of various technologies incorporated for designing architecture of B2CFS.

A. Role of Cloud Computing

The conventional approach for recordkeeping is storing data on local storage. This conventional approach restricts broader accessibility, scalability, and recovery of data. There can be different cloud options as per the requirement: public, private, hybrid

Storing encrypted data on the cloud, for easier access of data from anywhere using any device. This technology supports data storage on a cloud i.e., remotely placed storage resources for improved efficiency. If the organization is dealing with sensitive data that cannot be stored remotely means they want on-campus storage, then they can opt for the private cloud; where all storage resources are deployed on campus and maintained on

campus. If there are hurdles to in-campus maintenance of the storage resources and data; due to lack of manpower or any other reason, one can go for managed private cloud. Here, infrastructure is launched on campus but cloud operations and maintenance tasks are outsourced

B. **Interpretation by statista**

As per the survey report by Statista, which is a well-known German company (database company) specializing in market and consumer data. According to the company, its platform contains more than 1,000,000 statistics on more than 80,000 topics from more than 22,500 sources and 170 different industries and generates a revenue of about €60 million.

As of 2022, surveyed technical executives, managers, and practitioners of cloud technologies from around the world indicated that the biggest challenges of using cloud computing technology within their organizations were related to security. Around 85 percent of respondents found security to be a significant challenge. Other commonly cited challenges included managing cloud spend, governance, and lack of resources/expertise. And Blockchain can overcome this challenge.

C. **Role of Big data**

High volumes of data present a challenge to the cloud environment. It seems like cloud computing and big data are an ideal combination for this. Together, they provide a solution that is both scalable and accommodating for big data and business analytics.

To surmount scalability issues big data technology is utilized. Data storage and processing are distributed. The data file gets split into small chunks will be called blocks, blocks are encrypted and stored across the cluster.

Now issues regarding on demand availability, scalability, easy data recovery, broad accessibility, efficient handling of massive amount of data get resolved [9].

D. **Amalgamation of Blockchain**

The Blockchain will resolve security and double spending issues. And the crucial blockchain component supporting these is the Merkle tree. Nowadays, blockchain technology is ruling the IT industry. Not surprisingly but the Blockchain market is remarkably growing. The growth is due to prominent features of blockchain like improved trust, security, traceability, and prevention of double spending. Blockchain also uses a distributed record-keeping system called a ledger that keeps track of changes to assets within the chain[10][11].

E. **Potential of Merkle tree**

Merkle tree is nothing but a tree of hash values. Merkle trees undoubtedly have a place in the process of securely and effectively storing data. A single block in the blockchain can contain multiple transactions.

Instead of applying the hash algorithm directly on a bunch of transactions, get the hash value of each unique transaction, and the same transaction hash is hashed several times to get Merkle root. This Merkle root will be a part of Blockchain's header part[12].

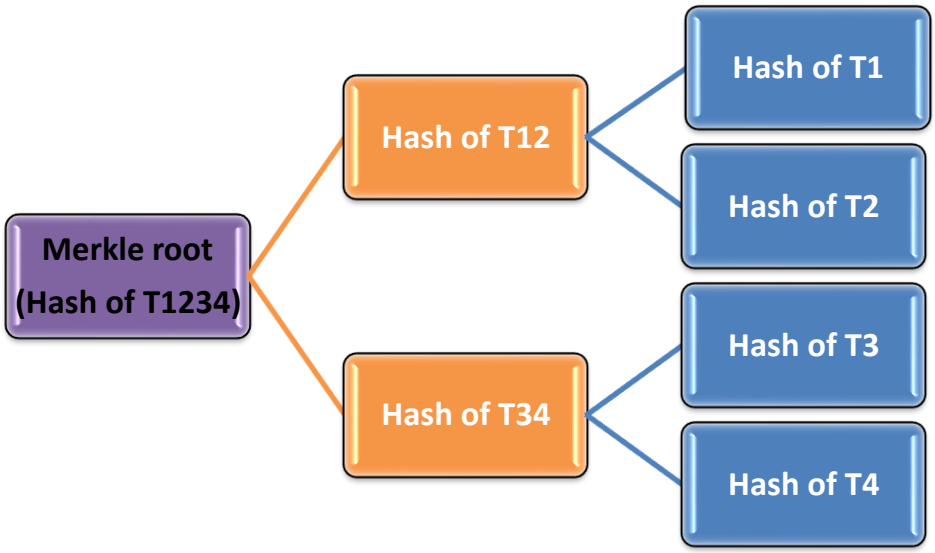

Figure 1: Generation of Merkle tree

In the given example, there are four transactions (T1, T2, T3, T4) to be included in the blockchain. Generation of Merkle root starts with getting the hash of each transaction. At the next level, all the transactions are again hashed by grouping the hash values of two transactions. The number of levels will differ based on the number of transactions. In the end, there will be a single hash called Merkle root, which is linked to the hash of all the transactions in the block. So, if an intruder tries to alter any transaction, it will reflect in the Merkle root and the alteration will get rejected. The generated Merkle root is the part of Block header.

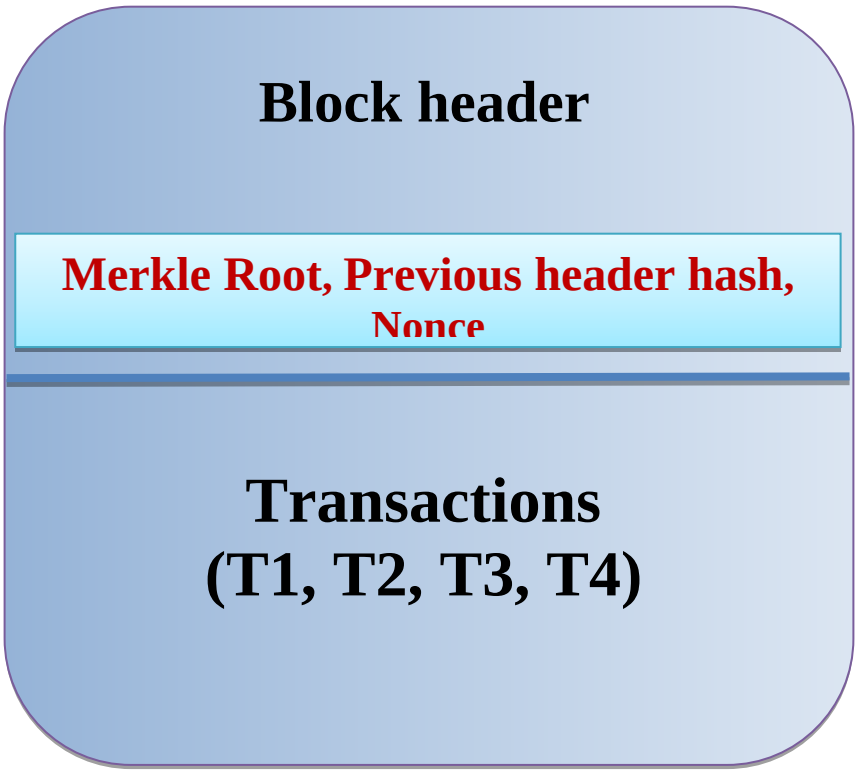

Figure 2: Inclusion of Merkle root in block header for next level data security

F. **Potential of Content Addressing**

In content addressing, all contents are uniquely identified by their cryptographic hash. So, content addressing will ensure tamper-proof data and prevention of double spending. Content addressing is referring to something by what it is instead of where it is. It seems silly to ask for the Universal location of a book because any copy of the book will suffice. The location is irrelevant, so long as the content behind an address is the same. The file is assigned a globally unique address. For Example, International Standard Book Number is your book's unique identifier. A separate ISBN is assigned to each edition[13].

G. **Unique features of B2CFS**

- VCS (Version Controlling System): The VCS ultimately depends on the existence of content addressed to DAG
- SVC (Self-Validation Certificate): SVC is a unique file naming system. The system ensures: File validation and authenticity of sending the node

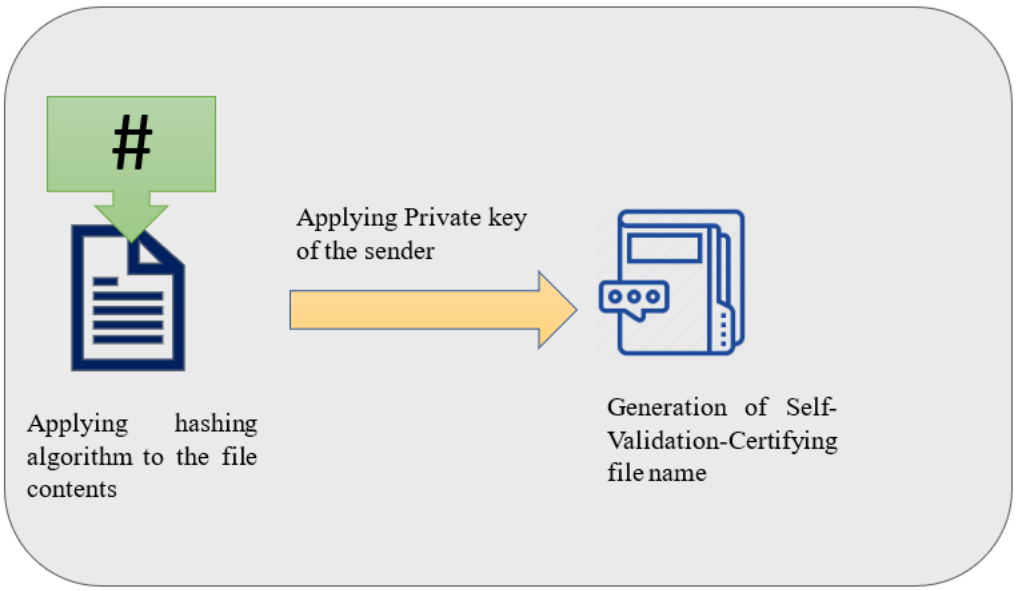

Figure 3: SVC file naming system

The above figure shows the key steps in the generation of a Self-Validation-Certifying file name. According to the system, the hashing algorithm is applied to the file contents. After that, the Private key of the sender node is applied to the hash generated in the first phase.

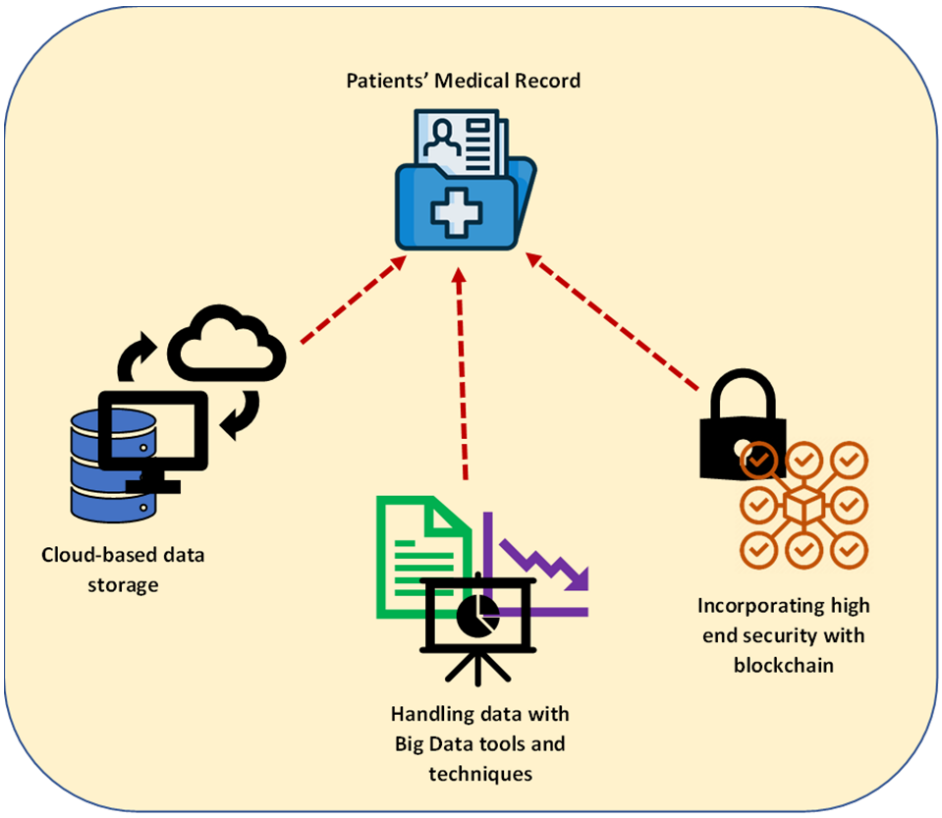

Figure 4: B2CFS Architecture

IV. CONCLUSION

By studying the prominence of all three cutting-edge technologies like Blockchain, Big data, and Cloud Computing researcher has stated that there is future scope in the aforementioned framework. The aforementioned framework will be one of the best solutions for Secured recordkeeping.

REFERENCES

[1] Koshechkin K.A., Klimenko G.S., Ryabkov I.V., & Kozhin P.B. (2018). "Scope for the Application of Blockchain in the Public Healthcare of the Russian Federation", ScienceDirect Procedia Computer Science 126, Pg. No. 1323–1328, International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES2018, 3-5 September 2018, Belgrade, Serbia

[2] Haber, S., & Stornetta, W. S. (1991). How to time-stamp a digital document. Journal of Cryptology, 3(2), 99–111. https://doi.org/10.1007/bf00196791

[3] Badadare, V. L., Shaikh, A. M., & Kulkarni, R. V. (2020, April). BLOCKCHAIN TECHNOLOGY IN HEALTHCARE SECTOR: A REVIEW.

[4] Hassani, H., Huang, X., & Silva, E. (2018). Big-crypto: big data, blockchain, and cryptocurrency. Big Data and Cognitive Computing, 2(4), 1–15. https://doi.org/10.3390/bdcc2040034

[5] Dhanalakshmi, S., & Charles Babu, G. (2019). An examination of big data and blockchain technology. International Journal of Innovative Technology and Exploring Engineering, 8(11), 3118–3122. https://doi.org/10.35940/ijitee.K2497.0981119

[6] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of Big Data, 6(1). https://doi.org/10.1186/s40537-019-0217-0

[7] Atkuri, C. (2022, January 7). Blockchain in Cloud Computing. [X]Cube LABS. https://www.xcubelabs.com/blog/5-ways-in-which-blockchain-is-improving-cloud-computing/

[8] Rovnaya, A., & Altynpara, E. (2023, February 14). Benefits of Blockchain in Cloud Computing. Cleveroad Inc. - Web and App Development Company. https://www.cleveroad.com/blog/blockchain-cloud-computing/

[9] Berisha, B., Mëziu, E. & Shabani, I. Big data analytics in Cloud computing: an overview. *J Cloud Comp* **11**, 24 (2022). https://doi.org/10.1186/s13677-022-00301-w

[10] https://www.fortunebusinessinsights.com/industry-reports/blockchain-market-100072 , Accessed on 1st August 2023

[11] Badadare, V. L., Shaikh, A. M., Kamath, R. S., & Pol, U. R. (n.d.). TRANSFORMING HEALTHCARE RECORD SECURITY SOLUTIONS USING BLOCKCHAIN TECHNOLOGY: A NOVEL APPROACH: Vol. II [English]. Bhumi Publisher.

[12] https://medium.com/techskill-brew/merkle-tree-in-blockchain-part-5-blockchain-basics-4e25b61179a2 , Accessed on 30th July 2023

[13] https://web3.storage/docs/concepts/content-addressing/ , Accessed on 30th July 2023