IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Decentralized cloud storage using blockchain

Rani gade Author 1 | **Sayali kore** Author 2 | **Vaishnavi mundhe** Author 3 | **Nikita pawar** Author 4

Bachelor's of Computer Application in cloud Technology & Information technology
Ajeenkya D Y Patil University Lohegoan , Pune , India

Abstract— One of the best places to store large amounts of data is in the cloud, however storing a single cloud space on a computer is not safe. Blockchain, on the other hand, is a cloud-based data storage solution that guarantees security. Any Internet-connected computer node can join and create peer networks, which increases resource utilization. Blockchain is a distributed peer-to-peer system that is kept immutable by each node in the network maintaining a copy of the blockchain. The IPFS (Inter Planetary File System) protocol is used in the suggested system to encrypt user files, which are then stored across several network peers. Hashes are generated by IPFS.

Keywords -Decentralized Cloud Storage, Blockchain Technology, Ethereum, Smart Contracts, IPFS Protocol, Hybrid Cloud, Encryption, Secure Transactions

I INTRODUCTION

In the 2008 article Bitcoin: a peer-to-peer payment system, an electronic cash system characterised as a record of Bitcoin transaction history, blockchain technology was first mentioned. a structure that brings together data units in chronological order. To guarantee that the distributed ledger cannot be harmed or altered, a cryptographic process is employed. Generally speaking, distributed nodes and consensus methods are used by blockchain technology to generate and update data, while encryption is used to secure data transmission and access control. The blockchain's data structure is used to verify and store data. A new paradigm for distributed computing and infrastructure that uses smart contracts, or automated script code, to programme and manipulate data. These are the features of blockchain. dispersed architectural design. The foundation of blockchain is Users can simultaneously observe that no transaction listed in the superblock has been altered. Blockchain employs string data structures, which offer great traceability and timestamp for every block. In addition, the consensus process and encryption system guard against blockchain manipulation. All things considered, blockchain is a novel kind of distributed computer architecture built on peer-to-peer network communication, smart contracts, cryptography, and consensus algorithms. uploading task scheduling data to the cloud. Ethereum is a cryptocurrency that was created in 2013. It is a blockchain-based distributed public computing platform. Following compilation, the code is converted to operational code, which is subsequently converted to binary code Using the file picker, the user uploads the file. The system verifies the availability of network storage and examines the file size. If sufficient space is available, the file will be uploaded. Next, the system executes step If there is insufficient disc space, users will be asked to try again. The 256-bit AES algorithm will be used to encrypt the submitted file. The user's pockets are used, and a randomly generated salt cost is produced by the encryption secret. The user's data is encrypted using this encryption key and the IV. This protects the privacy of the user's information. Then, using the IPFS protocol, the encrypted record is divided into 64-KB chunks and distributed t to other friends within the community.[1]The goal of this is to present an overview of decentralized file sharing utilizing blockchain technology while examining the many strategies, advantages, and difficulties related to this technology.[2]

As per the Forbes article [1], 2.5 quintillion bytes of data are produced each day. Out of the total data in the world over 90 percent of data was produced in the last 2 years. With such a massive increase in the data, cloud storage is required to store the data. Much of the data currently available through the internet is quite centralized and is stored with a handful of technology companies that have the experience and capital to build massive data centers capable of handling this enormous data. The problem with this approach is the security of data. As this data is stored in a centralized manner, if an attacker can gain access to the server he can easily view and modify the data. Another problem with this approach is the privacy of user data. In many instances, this data is used by third parties for data analysis and marketing purposes. Also, the cost incurred in storing data in centralized servers is more and many times users have to pay for the entire plan which they have selected even if they have used only a fraction of storage portion thus it does not provide flexibility to the user to pay only for what they are using. Another issue is the scalability of the system, it is difficult to scale a centralized storage system to meet the increasing demand. With zero trust two parties can transact in Blockchain.

The development of blockchain technology in recent years has created new opportunities for safe and decentralized networks. The use of blockchain for decentralized file sharing has garnered a lot of attention. Traditional file sharing systems are susceptible to hacking, censorship, and other security problems since they rely on centralized servers. On the other hand, decentralized file sharing makes use of a network of peers, each user having a copy of the file, making it far more secure against failures and assaults. Utilizing blockchain technology improves the system's security and transparency while enabling a trustless and tamper-proof method of file exchange. TThe goal of this is to present an overview of decentralized file sharing utilizing blockchain technology while examining the many strategies, advantages, and difficulties related to this technology. It also shows a few of the current decentralized file sharing platforms and their applications.

"Decentralized Cloud Storage Using Blockchain" serves as a foundational section that contextualizes the research by discussing the significance of secure and efficient data management solutions in cloud storage. It outlines the problem statement, highlighting the vulnerabilities of centralized data storage and the need for enhanced security measures. The research objectives are clearly defined, focusing on the development of a secure and decentralized cloud storage system leveraging blockchain technology. The introduction also sets the scope of the project, delineating the boundaries of the study, and briefly mentions the methodology that will be employed to achieve the project goals. Additionally, it provides a roadmap for the report, giving a glimpse of the structure and chapters that will be elaborated upon in subsequent sections.

II Objective

Our intended framework's primary goal is to mandate blockchain development in the cloud, and its secondary goal is to provide users with a safe cloud environment. cloud-based storage network with the goal of expanding the capabilities of standard cloud storage providers.

III Background and Related Work

Blockchain

The first description of the blockchain technique appeared in the 2008 paper Bitcoin: a Peer-to-peer Electronic Cash system, written by Satoshi Nakamoto, which is described as recording the history of Bitcoin transactions[1]. a block chain is a chained data structure that combines data blocks sequentially in chronological discuss order. The cryptographic method is used to ensure that the distributed account book cannot be tampered and cannot be forged. Broadly speaking, block chain technology uses block chain data structure to verify and store data, uses distributed nodes and consensus algorithms to generate and update data, and uses .using AES algorithm for data security purpose it's a encryption algorithm also additionally, RSA algorithm use for public and private keys pair to encryption and decryption processes . also using ipfs protocol create a hash value using SHA (secure hash algorithm) is a hash function .

IV problem identification

Blockchain is one of the most advanced technologies today. Our goal is to create a system that overcomes the problem of data security using encryption, decryption, and blockchain techniques in the cloud. Today, there are a variety of security issues, including issues with access control, scalability, virtualization, privacy, and big data. Traditional security techniques are no longer sufficient for data and applications in the cloud. Applications and data stored in the cloud have no fixed limitations because cloud computing is scalable and independent of location. When data is exchanged and stored on central servers, difficulties in data processing and authentication often arise. By preventing malicious users, blockchain provides a platform for cloud and data storage, increasing security.

V LITERATURE SURVEY

He Zhu, Yichuan Wang[1]discuss the The block chain technique is used to generate unalterable records, and blocks are generated for each data record, which is verified to ensure the transparency, authenticity and validity of the data.

This method mainly uses a distributed solution that combines block chain technology and traditional cloud server to solve the problem of integrity and security of cloud task.

Nazmun Nahar, Farah Hasin and Kazi Abu Taher[2] It has been established in the blockchain that transaction data can be more secure, protect privacy, and avoid attacks over the decentralized cloud storage network.

This system increases cloud stability around the network. There is also the advantage of longterm storage of large files for business or transaction. Blockchain data stored in a decentralized ledger keep secure but still, there is a challenge of the security issue.

Mrs. Rohini Pise, Dr. Sonali Patil[3] This kind of architecture makes the system more robust and resistant to different security attacks which are performed by unauthorized users who try to steal and disclose the information in data files of user for their benefits.

There is no single party to own the data and control over it. Another advantage of decentralized cloud storage is that these are continuously live networks for access to data any time. The maintenance of this server is also high. Another issue with this data storing technique is that they charge the highest possible price to their customer to use their services.

Tudor Gabriel, Andrei Cornel – Cristian, Madalina Arhip-Calin, Alexandru Zamfirescu[4] Ethereum Swarm is a solution under development and it is not suitable for production purposes but it represents a building block for the Ethereum blockchain vision together with Contracts to form a decentralized logic, Swarm decentralized storage and Whisper decentralized messaging. Another advantage of the Ethereum based storage is the encryption capability at the data file level using SHA256 asymmetric encryption method, the data inside the blockchain network cannot be altered. The Smart Energy Grid is defined as a series of advancements with the purpose of solving specific issues concerning the energy sector.

VI. methodology

One of the best solutions for safely storing all of the user's files is cloud storage. Blockchain is a data security cloud-based storage system. Any Internet-connected computer node can join and create peer networks, which increases resource utilisation. Initially, the user registers for an account on the metamask. Using web3.js, the app retrieves the user's wallet balance and account address from the metamask. Using the file picker, users choose the file to upload. The system determines how many peers are available. We encrypted files using the AES technique and encrypted text using the RSA algorithm. The combination of blockchain and these algorithms makes our hybrid cloud extremely secure. Our primary objective is to save any file on ipfs.

Metamask: Browser extension which acts as a bridge to connect with the ethereum network.

Ethereum network : It is an open-source, public blockchainbased distributed computing platform. Ethereum uses smart contracts where one can add business logic to make decentralized applications as per the business requirements. Peers: These are the users of the system who have pledged to rent their free storage for another user’s to store files.

AES: Advance Encryption Standard (AES) is a symmetrickey algorithm that supports block length of 128 bit and can have a key size of 128, 192, and 256 bits. IPFS protocol:

IPFS is an open-source peer to peer file transfer protocol.

A. Uploading of file

User uploads file using the file picker. The system checks the file size and ensures storage availability in the network. The file is uploaded when enough storage is available. Then system performs step B. Users are notified to try again when enough storage is unavailable.

B. Encryption of file

The uploaded file is encrypted using AES 256 bit algorithm. The encryption key is generated using the user’s wallet address and randomly generated salt value. This encryption key along with an IV is used to encrypt user’s data. This maintains the confidentiality of the user’s data.

C. Storing of file across multiple peers

The encrypted file is then divided into blocks of 64KB and sends to different peers across the network with the help of the IPFS protocol. The proposed system uses a private IPFS network to allow registered peers to store the file in the network. The file block is replicated on multiple peer’s storages for high availability using the IPFS cluster.

D. Storing of file across multiple peers IPFS returns a hash value which indicates the path of the file. The hash value along with metadata is mapped with the user’s wallet address and is stored in the blockchain using a smart contract. Smart contracts are like agreement and are used to eradicate the need for a third party. They control the transaction between nodes or assets between parties under certain conditions. This is lines of code stored on a blockchain network and are automatically executed when predetermined terms and conditions are met. In our proposed system preconditions for the smart contract to execute are: 1) Enough Space is available in the network to store files. 2) The user has sufficient wallet balance to pay the peers. Fig. 2. Smart Contract to store file details The above smart contract stores all the files details in the structure named FileDetails and maps this structure with the user’s address. It consists of two functions, one to add a new file and another to get the details of the uploaded file

E. Paying the peers for file storage

Once the file is stored across peers, total cryptocurrency is calculated and is deducted from the user’s wallet. This cryptocurrency is first transmitted to the smart contract from the user’s wallet. With the smart contract, this amount is distributed to the peers who have stored the user’s file

VII software and hardware requirements

A. Current/ existing System

Blockchain networks have two applications. To guarantee stability, the hash data acquired from cloud collection to the block chain network is distributedly stored and its integrity is safeguarded. Furthermore, every response received from the cloud server and every website visit will be logged in a block series for future analysis or research. The data record will be preserved indefinitely, and in addition, a data block will be created to verify the information. Similar to block chain networks, the cloud server handles data from cloud collections and data access records.

The cloud server must request block data from the block chain network as irreversible proof of data in order to secure the data record.

-disadvantages of existing system

One drawback of the current system is that maintaining and implementing a hybrid cloud environment might be challenging. To guarantee that the public cloud solution being used is compatible with private infrastructure, design and implementation frequently call for the expertise of an experienced cloud architect.

2. Full visibility over the entire environment is required in the hybrid cloud architecture due to the merging of dissimilar environments.

3. The cloud model's lower costs are one of its main benefits. However, the implementation cost of a hybrid cloud solution is more than that of a public cloud.

4. Files and applications utilised in hybrid cloud systems need to work in all settings, including on-premises and private and public clouds.

B. Framework/proposed system

Initially, the user registers for an account on the metamask. Using web3.js, the app retrieves the user's wallet balance and account address from the metamask. Using the file picker, users choose the file to upload. The system determines how many peers are available. Additionally, the uploaded file is encrypted via the AES method, which uses the user's wallet address as a key. A payment dialogue box asks for the user's approval. Following payment confirmation, the user's file is shared among peers using the IPFS protocol. IPFS then provides a hash value that includes the file path. A smart contract is then used to map this path with the user's address, and the data is safely saved in the blockchain. In order to attain high dependability and availability

-advantages of proposed system

1. A hybrid cloud eliminates vendor lock-in by allowing enterprises to select from a variety of management and service models offered by many providers, enabling them to employ the most creative and appropriate solution for their needs.

2. In order to get more scalability, several businesses have shifted to the hybrid cloud. Organisations may optimise performance and scale resources in response to changing needs by utilising a hybrid cloud environment.

3. A company's competitive edge may be determined by its time to market. By optimising IT performance, hybrid clouds enable businesses to provide goods and services to clients more quickly.

4. Businesses must be able to quickly adjust and shift course in this digital age. Businesses can integrate their on-premises infrastructure with their current technology by utilising the hybrid cloud strategy.

C Requirements for Hardware and Software

- **Hardware**

o Processor: i3 or above o RAM: 4 GB or higher

o Hard drive: at least 16 GB

- **Software**

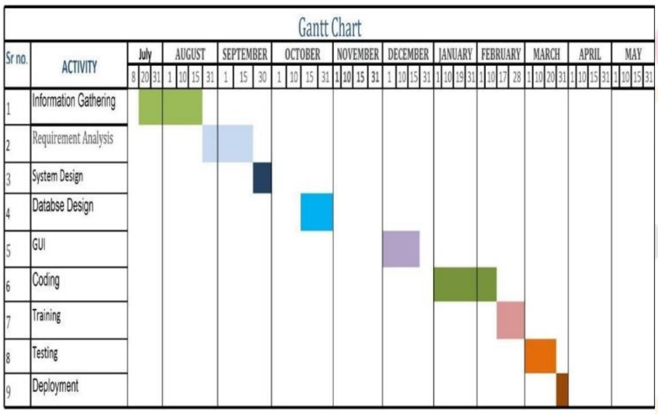
  - Python is the software used in Windows 10, 7, 8, and 10.
  - Anaconda. · Jupyter notebook, flask, spyder.
  - Microsoft SQL Server. ·
  - Ganache

A.Python is an object-oriented, high-level, interpreted programming language with dynamic semantics. Its dynamic typing and dynamic binding, along with its high-level built-in data structures, make it an appealing language for Rapid Application Development and for usage as a scripting or glue language to join existing components. Because of its straightforward, basic syntax, Python emphasises readability, which lowers programme maintenance costs. Python's support f packages and modules promotes code reuse and programme modularity. The interpreter for Python and the comprehensive standard

B.FLASK: A Flask is a web application framework designed with speed and flexibility in mind. Python, which is widely used by data scientists, is built into Flask. Flask handles the project setup and environment required for web applications. releasing the developer to concentrate on their programme instead of worrying about routing, HTTP, datasets, etc. Data scientists can construct straightforward single-page applications with Flask, and this is something they should consider if they wish to produce products for consumers. Python-based Flask is a microweb framework. It is categorised as a microframework since it doesn't need any specific libraries or tools. There is no form validation, database abstraction layer, or other component where third-party libraries are already in place.

C Ganache: Ganache is an Ethereum and Corda private blockchain that promotes app development. Ganache is available for usage throughout the development cycle, giving you the opportunity to create, test, and use your dApp in a top-notch, safe environment. There are two flavours in Ganache: UI and CLI. Software for computing devices called Ganache UI supports Corda and Ethereum technologies. For Ethereum development, the command-line tool ganache-ehl (formerly known as TestRPC) is to be utilised. Would you rather use a command line? Only Ganache's UI preferences will be taken into account in these texts. Please refer to the Ganache CLI Readme for information on command line tools. You can get all Ganache versions for Linux, Mac, and Windows.

VIII Life Cycle of Software Development
The project was completed within a total of six to seven months. The Waterfall model was used to successfully design and construct a model that was affordable.



Phase of requirement collection and analysis:

This stage of the project began when we had organised the work into groups and divided it into smaller components. Key factors to take into account were

1. Clearly define and illustrate each goal.
2. Compile needs and assess them

After taking into account the necessary technical parameters, gather the technical details of the different peripheral components (hardware) that are needed.

3. Examine the coding languages that the project requires.
4. Specify your coding techniques.
5. Examine potential dangers and issues.
6. Identify risk avoidance techniques; if not, identify backup plans for handling these risks.
7. Examine the viability financially.
8. Explain Gantt diagrams

IX Planning and Execution

<u>App validation:</u>

TestingBlack box testing and white box testing are the two main categories of software testing techniques. When creating test cases, a test engineer's perspective is characterised by these two methods.
Unknown-box testingSoftware is tested using a technique known as "black box" testing, which ignores internal implementation.

Equivalency partitioning, boundary value analysis, all-pairs testing, fuzz testing, model-based testing, traceability matrix, exploratory testing, and specification-based testing are examples of black box testing techniques.

Testing according to specifications:The goal of specification-based testing is to evaluate software functioning in accordance with the relevant specifications. As a result, the tester only sees the test object's output after entering data into it. Typically, this level of testing necessitates the provision of comprehensive test cases to

The goal of specification-based testing is to evaluate software functioning in accordance with the relevant specifications. As a result, the tester only sees the test object's output after entering data into it. For this kind of testing, the tester typically has to be given comprehensive test cases. After receiving these, they may easily confirm whether the output value (or behaviour) for a given input "is" or "is not" the same as the expected value stated in the test case.While testing based on specifications is important, it is not enough to prevent all hazards.

Benefits and drawbacks: A black box tester's vision is straightforward: a code must contain defects. The tester has no "bonds" with the code. Bugs are discovered by black box testers using the "Ask and you shall receive" concept.

Testing of integrationAny kind of software testing that compares a programme design to the interfaces between components is called integration testing. Software components can be incorporated all at once ("big bang") or iteratively. The former is typically seen as preferable practice since it makes it possible to localise and fix interface problems more rapidly.

Acceptance examination

One of two things can be meant by acceptance testing:

Before adding a new build to the main testing process, i.e., before integration or regression, a smoke test is used as an acceptance test.

User acceptability testing (UAT) is acceptance testing carried out by the client, frequently on their own hardware in a lab setting.

X future scope/modifications

In the future, a versatile editing algorithm might combine files that are frequently accessible for each user as opposed to those that are infrequently accessible. This will guarantee that the user can always easily access the accessible files when needed. Additionally, depending on how long their system has been operational, each allocated peer can earn an additional 100 credit default from the programme. A small number of successfully approved file access requests have had their credits drawn or added. Peer-to-peer peers will receive the most crucial data storage item.

XI Conclusion

Data security is enhanced by the suggested system, which encrypts and distributes data among several system peers. The operating system encrypts data to guarantee user data secrecy using the AES 256 bit encryption technique. The IPFS protocol is then used to send and store the encrypted data to peers on the network. Our method maximises the utilisation of the storage facility by solving the issues of privacy and security associated with central cloud storage and enabling peers to rent their idle storage and earn cryptocurrency-returns.

The proposed system improves data security by encoding and disseminating data to multiple peers in the system.

The operating system uses the AES 256 bit encryption algorithm to encrypt data that ensures the confidentiality of user data.

The encrypted data is then transmitted and stored to peers on the network using the IPFS protocol.

Our system not only solves the privacy and security of central cloud storage but also provides peers to rent their unused storage and receive cryptocurrency returns, thus maximizing the use of the storage facility.

In conclusion, the project "Decentralized Cloud Storage Using Blockchain" represents a significant advancement in addressing the challenges of data security and integrity in cloud storage systems. By leveraging blockchain technology, the project has successfully developed a secure and decentralized storage solution that enhances data protection and transparency. The research has contributed valuable insights into the potential of blockchain in revolutionizing cloud storage practices, offering a more resilient and efficient approach to data management. Moving forward, the findings and framework developed in this project pave the way for further exploration and implementation of decentralized cloud storage solutions, marking a crucial step towards a more secure and reliable data storage environment in the digital era.

REFERENCES

[1] He Zhu , Yichuan Wang, Xinhong Hei, Wenjiang Ji, Li Zhang " A Blockchain-based Decentralized Cloud Resource Scheduling Architecture " International Conference on Networking and Network Applications, 2018.

[2] Nazmun Nahar, Farah Hasin and Kazi Abu Taher " Application of Blockchain for the Security of Decentralized Cloud Computing " International Conference on Information and Communication Technology, 2021.

[3] Mrs. Rohini Pise , Dr. Sonali Patil " Enhancing Security of Data in Cloud Storage using Decentralized Blockchain " Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, 2021.

[4] Meet Shah, Mohammedhasan Shaikh , Vishwajeet Mishra Grinal Tuscano "Decentralized Cloud Storage Using Blockchain" h International Conference on Trends in Electronics and Informatics (ICOEI 2020.

[5] Yan Zhu1 , Chunli Lv1 , Zichuan Zeng1 , Jingfu Wang1 , Bei Pei2 "Blockchain-based Decentralized Storage Scheme"

[6] Tudor Gabriel, Andrei Cornel – Cristian, Madalina Arhip-Calin, Alexandru Zamfirescu " Cloud Storage. A comparison between centralized solutions versus decentralized cloud storage solute,ions using Blockchain technology " Department of Power Engineering Systems, 2019..