Project Proposal:
BLOCS
Blockchain-based Ledger for Open Cloud Storage

Prakhar Singhal [2022111025]
Harpreet Singh [2022101048]

September 29, 2024

Abstract

This project proposes a decentralized cloud storage system built on a blockchain with Proof-of-Storage (PoS) consensus. The system allows users to contribute storage space from their devices and, in return, access distributed storage on the network. The proposed system ensures data privacy, integrity, and redundancy through encryption and sharding, while a blockchain ledger tracks all storage transactions. This proposal details the system's architecture, storage mechanisms, incentive model, and file retrieval methods.

Contents

1 Introduction

Cloud storage has become a fundamental service for individuals and enterprises. However, centralized cloud providers pose risks of single points of failure, privacy concerns, and data monopoly. This project introduces a decentralized peer-to-peer (P2P) storage network using blockchain technology. Nodes in the network contribute unused storage space, while users can store files in an encrypted, distributed manner. Blockchain records each storage transaction, ensuring transparency and immutability. The system operates on a Proof-of-Storage (PoS) consensus algorithm, enabling nodes to prove their storage contribution without exposing the content they hold.

2 System Overview

The proposed system comprises the following key components:

- **Blockchain Layer**: A distributed ledger recording storage transactions and utilizing Proof-of-Storage for consensus.
- **Storage Layer**: A distributed storage layer that fragments, encrypts, and distributes file pieces across multiple nodes.
- **Incentive Layer**: A token-based economy where nodes earn tokens for providing storage and users pay tokens to store their files.
- **File Management**: Mechanisms for file upload, retrieval, and validation of stored data without revealing its content.

3 Blockchain Layer: Proof-of-Storage Consensus

The blockchain is the backbone of the proposed system, providing a transparent and immutable ledger for all storage operations. The consensus mechanism is based on Proof-of-Storage (PoS), where nodes must prove that they are indeed storing the data they claim to hold. The PoS protocol can be broken down into:

- **Storage Registration**: Nodes announce the amount of space they are willing to contribute.
- **Storage Proofs**: Periodic challenges are issued to nodes, requiring them to prove that they hold specific pieces of data.
- **Rewards and Penalties**: Nodes that successfully prove their storage are rewarded with tokens. Nodes that fail to provide proof are penalized or slashed.

Figure 1: Blockchain Layer and Proof-of-Storage Workflow

3.1 Data Validation

The PoS protocol employs cryptographic techniques like **Zero-Knowledge Proofs (ZKPs)** to allow nodes to validate storage without revealing the data itself. The system also supports periodic challenge-response protocols where nodes are required to prove they hold specific file pieces.

4 Decentralized Storage Layer

4.1 File Sharding and Encryption

When a file is uploaded to the system, it is:

1. Encrypted using a symmetric encryption algorithm to ensure data confidentiality.
2. Fragmented into multiple smaller pieces (shards).
3. Distributed across various nodes, ensuring that no single node holds enough information to reconstruct the entire file.

This method provides security, redundancy, and scalability.

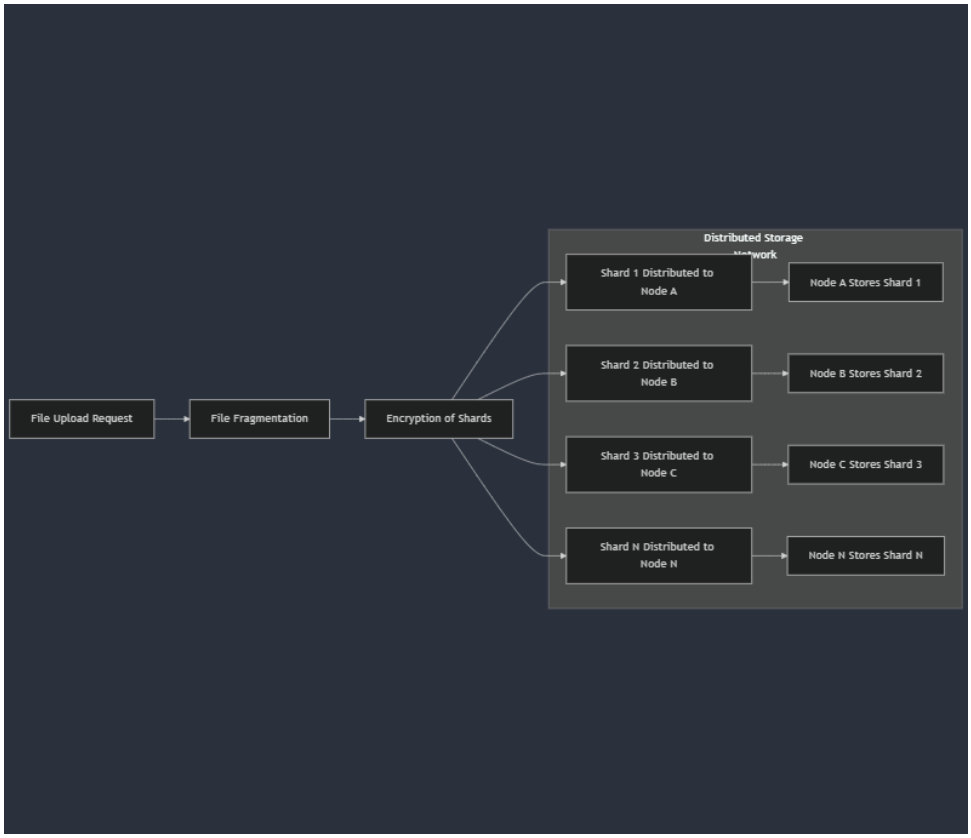

Figure 2: File Fragmentation and Distributed Storage

4.2 Redundancy and Availability

To guarantee data availability, the system introduces redundancy. Techniques such as **Erasure Coding** or replication are used to distribute file shards across multiple nodes, ensuring data is retrievable even if some nodes go offline.

4.3 File Retrieval

When a user requests a stored file, the blockchain is queried for the metadata and shard locations. The user retrieves and reassembles the file by downloading encrypted shards from various nodes and decrypting them using the encryption key.

5 Incentive Layer: Token Economy

The system incentivizes node operators and storage providers by implementing a **token-based economy**. The core aspects include:

- **Earning Tokens**: Nodes that provide storage receive tokens as rewards for storing and maintaining data integrity.
- **Paying for Storage**: Users pay tokens to upload and store data.
- **Smart Contracts**: The storage system uses smart contracts to automate payments and rewards, ensuring that no central authority controls the economy.

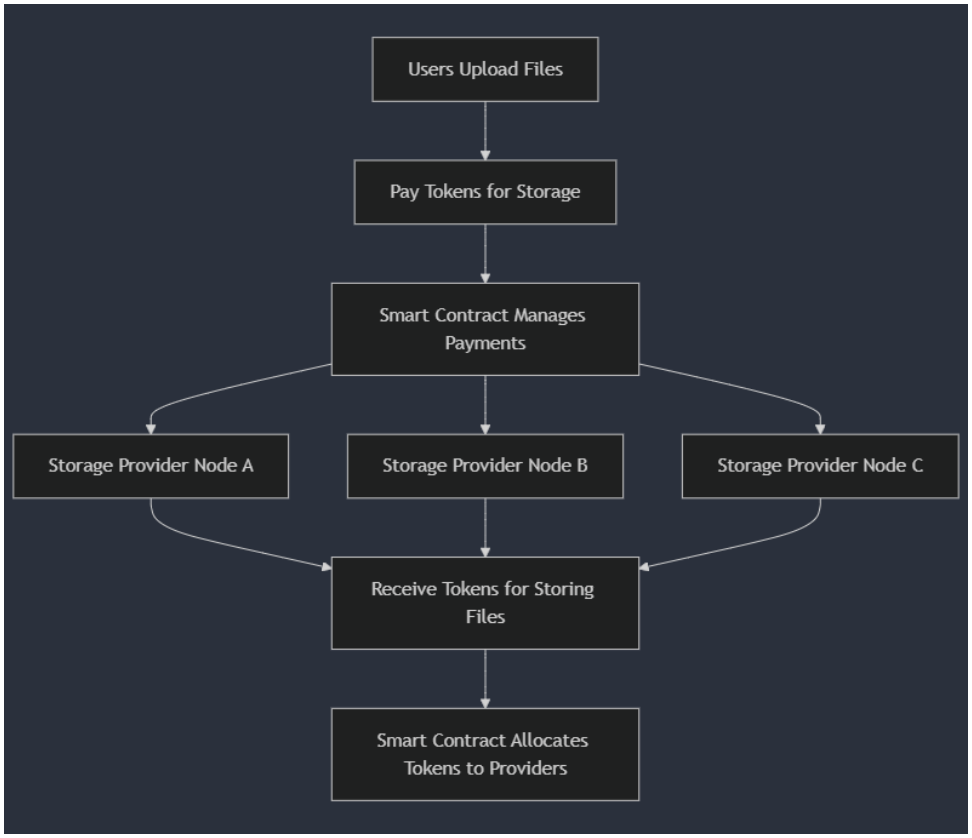

Figure 3: Token-Based Economy for Storage and Retrieval

5.1 Token Distribution and Penalty

Tokens are distributed based on the amount of space provided and the reliability of the node. Nodes that fail to provide storage proofs or go offline are penalized through token slashing mechanisms.

6 File Management and Metadata Storage

6.1 File Upload

File uploads involve splitting the file, encrypting the pieces, and distributing the shards across multiple nodes. Metadata, such as the file hash, ownership information, and shard locations, is stored on the blockchain.

6.2 File Tracking and Ownership

Each file stored in the network is linked to its owner's public key, ensuring clear ownership rights. This information is immutable and publicly verifiable through the blockchain ledger.

7 Security and Privacy Considerations

7.1 End-to-End Encryption

All files are encrypted on the client side, ensuring that only the uploader can access and decrypt the file. This provides a secure means for data to be stored without risking exposure of sensitive information.

7.2 Zero-Knowledge Proofs (ZKPs)

To ensure data privacy while allowing validation, Zero-Knowledge Proofs (ZKPs) are used for verifying the storage of data by nodes. These proofs allow storage providers to demonstrate they are holding a file without revealing its content.

8 System Architecture

The proposed decentralized storage system can be summarized in the following architectural diagram:

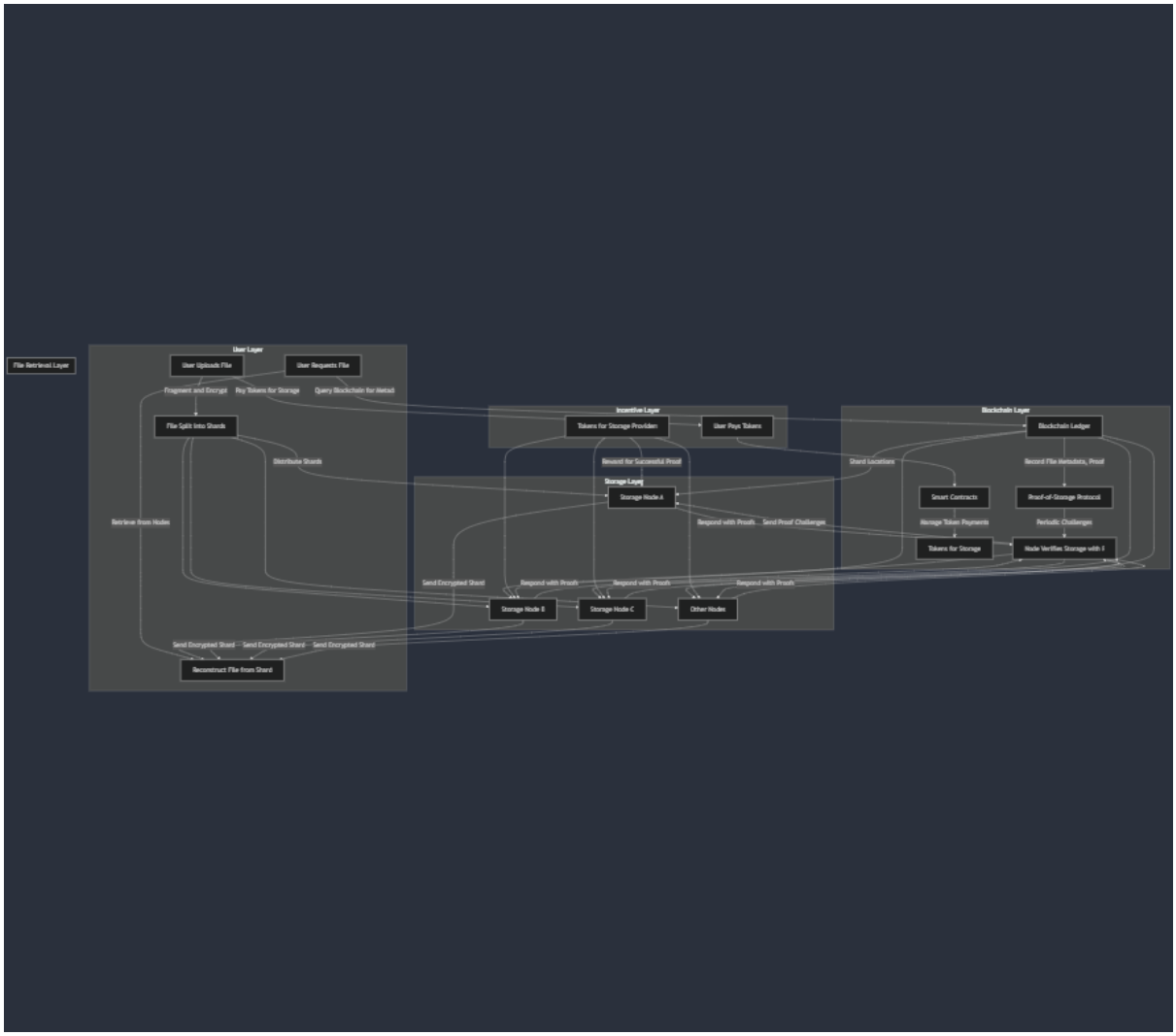

Figure 4: System Architecture Overview

9 Advantages of the Proposed System

- **Decentralization**: Removes the reliance on centralized storage providers, ensuring no single point of failure.
- **Security**: Data is encrypted and sharded, providing high levels of security and privacy.
- **Transparency**: Blockchain ledger records all storage transactions and operations, ensuring trust and auditability.
- **Incentives**: The token-based reward system motivates nodes to participate and contribute storage.

10 Challenges and Future Work

- **Scalability**: Off-chain scaling solutions may be required to handle a large number of transactions.
- **Data Availability**: Ensuring that data is always retrievable despite the decentralized nature of the network.
- **Governance**: Future work could explore decentralized governance models (e.g., DAOs) for system updates and management.

11 Conclusion

This proposal presents a decentralized storage system using blockchain and Proof-of-Storage. It ensures data privacy, immutability, and redundancy without requiring a central authority. The token-based incentive model motivates nodes to participate and provide storage while allowing users to safely store files in a distributed environment. Future work will focus on scalability and governance to further enhance the system.