CESS

Cumulus Encrypted Storage System

An Infrastructure of Decentralized Cloud Network Facilities

https://cess.cloud

Abstract

From stone-carved tablets to writing paper; from hardcopy printing to electronic printing of digital records, and from the birth of creation to data virtual space… Today, civilization keeps pace with the rapid progression of science and technology. But now it must also reflect upon the ceaseless pursuit of ever more convenient, efficient and safe data storage. Humans have a natural preoccupation with recording data, yet technological advances in the pursuit of data storage security remains questionable.

Civilization is now entering the digitized era. On the information superhighway that is today's Internet, huge amounts of data have achieved rapid transmission, retrieval, storage, mining and analysis. To use an analogy, if typewriting is the seeder of human civilization, then big data is the new energy, new technology and new organization method for the future. Big data will ultimately hail in a new era of human social development.

Society is being increasingly impacted by the digitized world. In this digitized world our data must be protected. Data security and ownership should belong entirely to the data's creator: from personal data recorded during everyday life, through to big data that requires high storage and computational requirements - all must be kept private. The value of large amounts of personal data has not yet been fully exploited and benefits the data creator. Humanity is building a new civilized order based on data.

Blockchain is considered to be the basis of the next generation of Internet - "Value Internet". The emergence of blockchain technology will promote the accelerated implementation of Web3.0 Internet. Based on blockchain and other technologies, a new digital world will be created in the network world and virtual space, which may contain greater wealth than the physical world.

The Decentralized Distributed Cloud Storage Infrastructure set up by CESS is dedicated to protecting data security from leaks, loss, tampering and theft, protecting individual privacy and rights, and keeping a firm grip on the right to share and trade personal data.

CESS Decentralized Cloud Storage ecosystem solves the problem of data storage security that has plagued mankind since inception of the worldwide web. It is also an open, transparent, co-building, sharing and co-prosperous business system operation platform based on blockchain for strong storage security requirements.

CESS is committed to promoting data and information connectivity and building a more open, fair and secure distributed network world. It will redefine storage systems, which will also have the most profound impact on the future development of an interconnected digital world. This is the meaning of CESS and a powerful driving force for the development of Web3.0!

Cumulus Encrypted Storage System has ushered in an era of de-centralized distributed cloud storage.

Catalogue:

1. Project overview.. 1

2. Managing accelerated growth..3

2.1 Web2.0 Sword of Damocles...3

2.2 Dilemmas of data storage..3

2.3 GDPR and Data Storage.. 5

3. New solutions based on blockchain...8

4. Application scenarios...11

4.1 Distributed Network Disk..11

4.2 NFT Storage...12

4.3 Distributed Storage.. 13

5. Technical Implementation... 15

5.1 Architecture Design.. 15

5.2 Blockchain Layer Design..17

5.3 Distributed Storage Resource Layer.. 28

5.4 Key technology 1: Data tampering prevention mechanism...32

5.5 Key technology 2: cross chain mechanism...34

6. Unique technology...36

6.1 Cross-Chain data authentication and communication technology..............................36

6.2 Data tamper-proofing technology...37

6.3 Multi-policy Storage Capability of Proof.. 39

6.4 Data Rights Protection Mechanism Based on Smart Contracts....................................40

7. Security Mechanisms..42

7.1 Data Security..42

7.2 Consensus Security...42

7.3 Multi-copy Attacks.. 43

8. Economic Model...44

8.1 Roles and Functions..44

8.2 Storage Capacity Model...44

8.3 Storage Award... 44

8.4 Currency Allocation Model.. 46

8.5 Benefit model for each role..46

9. Decentralized transactions and storage mining...47

9.1 Storage Markets: Verifiable and Trusted Markets..47

9.2 Storage mining: commercial implementation of de-centralized storage............................ 47

9.3 Storage Asset Markers: Improving Resource Integration in the Economy........................ 48

10. Future Prospects..49

11. Reference... 50

1. Project overview

With the development of new technologies such as computing technology, virtual reality and artificial intelligence, human society has gradually stepped into the digital era, accompanied by the explosive growth of data in cyberspace. According to the world economic forum, the amount of data in the world reached estimated 44 à (ZB) in 2020 alone. Among them, according to incomplete statistics, Google search alone has 4 billion daily visits, and the monthly number of active users of social media Facebook is even more than the total population of China, generating 4 megabits of data every day, including 10 billion messages, 350 million photos and 100 million hours of video browsing. There is no doubt that storing such big data is significant. But, how to save such massive data proves difficult. In the current situation of increasing data and even accelerating to the exponential level, cloud storage technology is widely accepted and applied in order to overcome the shortcomings of traditional independent data storage media in terms of reliability and easy loss. With the help of cloud storage technology, people can access their data anytime and anywhere, without worrying about whether the storage space is enough, whether the data will be lost, and whether it is inconvenient to carry. However, the over centralization of cloud storage also faces the problems of data loss and privacy leakage. Based on this, we developed a distributed cloud storage infrastructure based on blockchain technology - Cumulus Encrypted Storage System (CESS), which has the storage efficiency and experience of traditional cloud storage, but effectively resolves the security issues of current data storage by utilizing our decentralization model.

CESS are the trailblazers of decentralized cloud storage infrastructure built upon the foundation of blockchain and cloud storage. Utilizing blockchain technology, CESS makes effective use of online idle storage resources to establish an infinite distributed storage network of low capacity data nodes. Simultaneously, with the help of virtualization technology and cloud computing technology, it implements effective management of these resources to reprocess then redistribute data and provide users with a secure and efficient data storage service. CESS has proven data storage reliability. All datum will exist in the form of three copies (CESS can provide additional bespoke copies for an additional fee), simultaneously, CESS will use the Redundant Error Correction Coding to code Redundant Coding in order to further improve the data reliability, subsequently encrypting and shredding the data, then distributing all over the CESS network. In the unlikely event of damage to datum, loss of data files will not occur. In addition, based on the characteristics of blockchain, CESS provides a trading platform template for efficiency, transparency and equality for all data owners, who in turn may independently manage their own data on the CESS network, including data sharing and data trading, so as to facilitate users to mine the intrinsic value of data and yield its potential benefits.

2. Managing accelerated growth

2.1 Web2.0 Sword of Damocles

Currently, the Internet is only in the Web2.0 stage. With the continuous efforts of many Internet giants like Google, Facebook, Amazon and Twitter, Web2.0 has made the Internet easier to access. Nearly two-thirds of the world's population has access to the Internet, which has a profound impact on human society.

In the Web2.0 era, everyone interacts with tens of thousands of individuals, businesses, and organizations in ways that were previously unimaginable, producing a vast amount of binary data consisting of 0 and 1. According to IDC, the amount of data generated globally will reach 33 ZB in 2018 and 175 ZB by 2025.The global cloud storage market grew from 30.7 billion in 2017 to 88.91 billion in 2020, with an average annual growth rate of more than 20%. IBM researchers suggest that 90% of the world's data today has been generated over the past two years; that data growth will accelerate further with the advent of new devices and technologies, and that trillion-dollar storage market is expected.

However, with the increasing size and importance of data in recent years, the security of Web2.0 centralized cloud storage, known as security, has been questioned. One of the most famous examples is that Facebook's 50 million user data has been compromised. Data security and sovereignty are the sword of Damocles hanging over Web2.0.

2.2 Dilemmas of data storage

In 2016, American scholars published *Platform Revolution: A Business Model that Changed the World* clearly showing that the essence of Web2.0 is platform economy and declaring that "platforms are eating the whole world". Platforms create value by facilitating transactions/interactions between different user groups. As monopoly platforms continue to seek rents, the proportion of income distribution obtained is increasing. The whole society is working for the platform, and the platform has obtained most of the economic value-added. Internet companies monopolize the market by controlling data on a large number of service providers and users. Data is considered the core assets of the enterprise. The technology giants of Web2.0 have enormous power over all the data they control. The rules governing the use of data and data are completely under the control of the enterprise, and ordinary users have no power to participate. The tech giants control the rules of the game.

Currently, Web2.0 data is completely controlled by the enterprise or central organization, which has the following problems:

Data is leaky. External attacks make data vulnerable to theft. Especially in many enterprises, internal data is stored in clear text, once stolen, all information is equivalent to full disclosure.

Data is easily lost. Two thousand years ago, the Alexander Library was burned down. The fire destroyed thousands of valuable documents in our history. Everyone thinks it is a human tragedy. However, this happens every day in Web2.0. Accidental loss of business operations, by hackers, or shutdown of business services can cause data loss.

The data will be reviewed. With centralized servers, central agencies can easily block access to them. For example, access to Wikipedia has been banned in Turkey for two years.

Data can be tampered with. Enterprises have supreme rights over their internal databases. In theory, engineers, algorithmists, and administrators can modify any data, such as deleting records of criminality, even when using so-called pure incremental databases.

Data will be sold in packages. At present, the black-market products sold by network data have shown the trend of internationalization, corporatization, intellectualization and anonymization. The illegal black and gray market gangs began to steal, transfer, integrate and trade data by utilizing tools such as dark nets and Telegram. Some black and gray products make use of a company's corporate cover to engage in illegal acts of data trading, and there are black and gray market workers who clean and integrate data obtained from multiple channels to provide services to the outside automatically.

Data islands. In most industrial/manufacturing enterprises, most of the existing information-based application systems are built independently according to the traditional IT system vertical structure, with low openness, relatively closed, difficult to integrate, forming multiple chimney-style application systems, as well as multiple data islands. At the same time, these systems generally follow the traditional application development mode, resulting in large and complete functions, tight coupling of internal modules, low re-usability, heavy development workload and update difficulties. In this situation, we are facing great challenges in breaking through the data and value chain lines that cross business, departmental and even enterprise boundaries, realizing information and process connectivity, and developing and iteratively updating numerous business applications to meet the ever-changing business and innovation needs.

2.3 GDPR and Data Storage

At present, countries have deeply recognized the problem of data storage in Web2.0 and have formulated laws and regulations related to data security. In April 2016, the European Union formally passed the General Data Protection Regulation (GDPR) and entered into force in May 2018. The so-called "strictest in history" legislation incorporates the Privacy Protection Directive, the Electronic Communications Privacy Protection Directive and the EU Civil Rights Directive. It stipulates that current or future business models and business organizations that will cross national boundaries should follow the principle of personal data protection and strictly control acts or products, including records, preservation, downloads, organizations, which may lead to the disclosure of personal privacy, and that data subjects have the right to require business organizations to correct, clean up and stop processing their personal data. This plays an important role in protecting users' private ownership in the current high-speed development of the big data era.

However, legal protection alone is not enough for data storage security, and infrastructure based on new technology innovation is also needed. The current Internet Web2.0 is not perfect but defective, so minor technical patches cannot be solved and need to be updated.

The call for Web3.0 is getting stronger and stronger.

So, what is Web3.0?

Web3.0 is the design and vision for the next generation of the Internet.

If internet Web1.0 and Web2.0 bring about the explosion of free flow of information traffic, Internet Web3.0 will bring about the explosion of free flow of information value.

In 2014, Dr. Gavin Wood proposed a revolutionary Web3.0 vision, one of wish is that Web3.0 will enable users to access the Internet with their protected identity and data. Web3.0 will launch a new global digital economy system, create new business models and markets, break platform monopolies, promote broad, bottom-up innovation, build a more open, decentralized and distributed network world where data creators can protect their data and protect their privacy and rights. Individuals have a strong grasp of the right to trade and share personal data, which will make a huge difference to future production relations.

Whether it is the Internet, or the blockchain, artificial intelligence, the Internet of Things, all the storage, calculation, transmission, all are still around data. Decentralized distributed storage based on blockchain, with the leading advantage of data certainty, makes data storage security possible and enables cross-platform, cross-collaboration and cross-format interoperability of data. The technical features of de-centralized storage, including data privacy non-tampering, privatization, capitalization, and trusted underlying infrastructure, future digital economy, are "scalability" of distributed storage for future applications. At the same time, distributed storage can provide a back-up aid for the underlying infrastructure, and it will inevitably spawn many applications. Distributed storage networks can better drive the arrival of Web3.0 and are one of the underlying technical infrastructure of Web3.0.

As a de-centralized cloud storage infrastructure, CESS hugely benefits some blockchain technology, such as distributed storage, non-tampering, traceability, multi-party maintenance, cross-validation, to effectively define data ownership, track and monitor data flow, and share data benefits rationally.

CESS creates a distributed storage ecosystem with account systems, smart contracts, and underlying distributed storage infrastructure to build a large-scale distributed storage network. It can launch a variety of trading and consensus mechanisms according to different application needs and create specific schemes for the construction of CESS storage ecosystem. At present, AI, Internet of Things, 5G all need a stable and secure CESS de-centralized cloud storage system. Digital creation, live broadcasting, video, music, novel, comic, e-commerce and other applications running on the Internet can be presented in the CESS ecosystem.

CESS provides a powerful and scalable storage infrastructure for massive growth of data, built on a sound and stable foundation.

3. New solutions based on blockchain

The rise of blockchain technology has revolutionized the path of Internet development. This novel and ingenious technology allows data to be distributed throughout the ledger without any central authority, but the basic requirement is that participants validate the data, making it possible for data to become more open and secure. The "free market" for data storage no longer seems remote.

Decentralized cloud storage based on blockchain technology will make full use of the advantages of blockchain technology to meet the actual needs of storing large amounts of data. Decentralized cloud storage, as its name implies, is the process of cutting data into small pieces and encrypting it and spreading it across multiple network nodes. Providers of de-centralized cloud storage are only responsible for managing and maintaining a decentralized blockchain network rather than controlling its storage data. Storage technology is open source, and no company has full control over the data in the blockchain network. Smart contracts automate transactions and are responsible for issuing instructions to store data that cannot be deleted and changed once uploaded up in a blockchain network.

For this reason, blockchain-based de-centralized cloud storage is more secure than centralized networks. All data is encrypted and copied to ensure redundancy, and each user has his own private key to control his data, which greatly ensures that the data uploader can control his own information. In addition, the node that stores the file is responsible for only a small piece of the file's content, which also means it is pointless to attack the storage node to obtain file information.

The idea of de-centralized cloud storage allows ordinary people to achieve their data storage needs by renting cheap hard disk space in peer-to-peer networks. To liven up the use scenarios for such emerging cloud storage, most projects start with two levels of supply-demand: storage mining (supply), where users add idle storage space to a distributed network, allowing storage space providers to receive storage demand, and using Token as an incentive mechanism. In terms of storage services (requirements), customers use blockchain networks to store data that is sent to many different miners around the world, But most of them are still in the experimental stage.

CESS combines the concept of de-centralized cloud storage to build a new ecosystem for de-centralized cloud storage and to target commercial applications to meet the actual needs of commercial applications for large amounts of data storage. The following goals will be achieved:

(1)Low-Cost storage: High storage costs pose a major challenge for Internet companies in choosing business data storage cycles and storage quality. CESS's well-designed incentives provide attractive economic returns to miners and encourage them to contribute to the storage network. At the same time, it punishes offending participants. Therefore, based on CESS storage network, a large, high-quality storage market can be established. This enables CESS to take advantage of a large number of unused bandwidth and storage resources and provide powerful storage services at a lower cost. To make storage systems easy to use, CESS has created a distributed cloud storage platform that provides a wide range of storage types of products, as well as perfect APIs and SDKs, making it easy for developers to dock and providing users with a better product experience.

(2)Privacy, security and stability: Data security and privacy protection are the primary requirements of CESS distributed file system design. By using data fragmentation and encryption algorithms, a single storage node does not store all the data of the data owner, but is distributed to multiple nodes, and the miners get only part of the information in the process of packaging and storing the data. At the same time, the user's data can only be retrieved by the person who owns the user's unique private key, and the system allows users to easily share their data with others. CESS uses an efficient distributed hash algorithm to maintain the integrity and reliability of its stored content. An available Storage Service Certification Algorithm (PoAs) has been developed to provide valid validation of available resources provided by nodes or miners. The PoSt algorithm is optimized to ensure that a node or miner actually stores the specified data within a specified time. The replication proof algorithm (PoRep) is used to ensure that a node or miner replicates the user's data. The traffic proof algorithm (PoF) guarantees data that nodes interact with their traffic.

(3)**Data validation**: data ownership has always been one of the great challenges for content application. The traditional means of validation use the mode of submitting ownership certificates and expert reviews, but lack of technical confidence, and there are potential tampering and other uncontrollable factors. To solve these problems, a solution of blockchain + distributed storage network + content identification is proposed. System content ownership certification claims: certification confirms ownership. Of course, the system also introduces a complaint mechanism. When a user considers that the content in the platform is infringed, the system can remove the content or transfer the ownership of the data after authentication.

The process of confirming data rights is as follow**s**: users submit content, content authentication, ownership decision, result upload. The core of the process of confirming the rights of content is the determination of ownership. Based on this, the system introduces the content determination node, which is similar to the authentication node, distinguishes the single authentication node mode, introduces a large number of authentication nodes. All authentication nodes determine the content uploaded by users. When more than 50% of nodes determine that the content is owned by the current user, that is, they decide that ownership belongs to the user. Upload user-owned content features to the block chain.

- Users use the system to upload content to the feature gateway.
- The signature gateway determines whether the system has stored the content or not, and then randomly extracts the file content signatures of the corresponding number of nodes based on the current number of online nodes and distributes them to the authentication nodes.
- The Authentication node obtains random file content characteristics and compares content feature databases.
- The Authentication node compares the databases and votes on the ownership of the content based on the result of the decision.
- Smart contracts determine the ownership of data confirmation based on voting results. file ownership and complete the content attribution on blockchain.

The feature database contains samples of content text or audio.

(4)Data rights protection: In this era of sharing economy, whether the rights and interests of data owners can be protected is a question of comparative relations among all users. CESS, as a de-centralized content storage and distribution network, provides a set of mechanisms to protect data owners' rights and interests from infringement. First, with the help of blockchain network, CESS protects every piece of data 24 hours uninterrupted, ensuring that every use of data is recorded, providing a complete, reliable and verifiable evidence service for data revenue distribution. CESS then ensures that data owners can receive real-time benefits from data sharing through smart contracts. The CESS system provides system contracts for two economic models by default for users to choose from. One is the fixed income model: each time the data is used, the owner's fees are fixed and the benefits are released immediately. A model of equity return: periodic statistics users' earnings, which are proportionally extracted from a portion of the earnings, and then proportionally distributed to each data owner according to the number and frequency of data used. This earnings cycle is issued.

(5)Build a prosperous application ecosystem: CESS provides a set of storage APIs similar to cloud services such as Amazon Cloud, Apple Cloud making it easier to migrate data from these services. CESS provides additional cloud storage products to better support distributed applications, including object storage, audio-video transcoding, and file access management. CESS has a comprehensive plan to develop a complete and sustainable ecosystem to facilitate the development of applications and services, and to encourage developers to participate in the development of CESS ecological applications.

4. Application scenarios

4.1 Distributed Network Disk

CESS can upload distributed disks and build a de-centralized cloud storage system based on CESS distributed infrastructure.

CESS disks do not require cloud servers, effectively avoiding dependency on backbone and centralized servers. User data can be stored in multiple storage nodes, Users download resources are no longer subject to network disk service providers, and data transfer speed has been greatly improved. Blockchain-based encryption algorithm, the stored data can be encrypted, ensuring the privacy of the storage, without worrying about data loss and central server shutdown. The CESS network supports unlimited expansion, so there is no need to worry about disk expansion. The capacity can be dynamically expanded according to the actual needs, breaking the storage restrictions of traditional network disks, so that users can enjoy the new generation of cloud storage and enjoy the pleasure of scientific and technological progress.

4.2 NFT Storage

Over the past year, NFT has become one of the most popular areas in the cryptographic world.

First, it attracts great attention from the world of the arts, luxury brands, celebrities and other public figures. NFT's secure storage is the core base of NFT's artistic and commercial value, and CESS provides a secure and reliable storage for NFT. Users only need to upload NFT files to CESS and the system will hash them and assign addresses. By combining knowledge structure, subject and semantic features with vector space modeling technology, NFT feature information can be matched automatically. Others can request files with their private keys to form a good environment for free flow of NFT. For example, in NFT transactions, American artists created NFT artworks and sold them to buyers from the United Kingdom. Copyright transfer and Timestamp generation were systematically completed on an open, transparent and traceable CESS chain. By mastering the private key of CESS, you can master the exclusive rights and interests of data asset sharing and distribution transactions, eliminating counterfeiting, counterfeiting and piracy. At the same time, transactions can be completed in an instant.

CESS can not only trade NFT digital assets, but also store NFT digital products.

It can also develop a variety of applications on CESS to act as an ecological incubator for NFT, reallocate benefits through smart contracts, and truly reward NFT producers and distributors. At the same time, the threshold for traditional platforms to access the blockchain is constantly lowered. Users do not need to have a deep understanding of the on-chain technology and can publish their own applications by calling friendly APIs. CESS has created a platform for developers to work together to process and use NFT data in a variety of ways based on CESS system. It enables more common users to use CESS system to complete various functions such as games, literary creation, painting and so on. It gives everyone the opportunity to have their own NFT, and to conduct NFT transactions and communications.

4.3 Distributed Storage

CESS is the pioneer of decentralized cloud storage system, and its advent will trigger a distributed storage revolution. CESS uses blockchain technology to design a multi-strategy, proof of storage-capability mechanism by using storage proof algorithm. It can make full use of idle bandwidth and storage resources to provide more powerful and efficient storage services than traditional cloud storage at a lower cost. It also validates data and the protection rights, providing better data security and privacy services, and further creates a de-centralized, incentive system, which is both flexible and scalable. The blockchain cloud storage platform on the chain can be verified and trusted, so as to establish a free trading data storage market. Specifically, CESS can bring the following changes to the data industry chain in terms of data storage:

Reorganizing the Data Production Chain:

CESS's distributed storage can make the production process non-linear and networked, that is, all production elements can be networked and multipoint configured. The mode of production is fully synergized, completing products in the human-machine-human collaboration, and realizing “crowd-production”. Its blockchain technology provides such technical conditions as "network synergy" and provides a non-linear production platform.

Reorganizing the Data Flow Chain:

CESS can eliminate or reorganize old intermediary channels based on stored transactions and flows, a de-intermediation feature that enables data products to meet and record with any user at any time and space. New modes will appear in the circulation of data products and services, such that non-linear circulation, monopoly channels and platform status will be more subtle. Smart contracts between producers, products and users make it possible for trustworthy dissemination between products and service points, which will reduce the cost of value circulation and improve the efficiency and effectiveness of value.

Reorganizing the Data Consumption Chain:

Instead of Corporations or Data Centers dominating data arena, the sovereignty may now be changed. The true owner of the data may now filter, integrate, optimize, match, and / or participate in managing his or her own data.

5. Technical Implementation

5.1 Architecture Design

CESS is a high-speed, secure, scalable and decentralized cloud storage system. It can handle tens of thousands of transactions per second through parallel technology. Through Data slicing technology, it can achieve the secure storage of massive data, and it has the functions of Data confirmation and Data rights protection, which provides powerful data service ability. It provides DAPP with unlimited scalable storage capacity and perfect Data rights protection capability.

As shown in the figure 1, CESS adopts a layered and loosely coupled design method, which is divided into Blockchain service layer, distributed storage resource layer, Distributed content distribution layer and Application layer. Among them, Blockchain service layer provides blockchain service of the whole CESS network, including encouraging idle storage resources, computing resources to join CESS network to provides data transaction, data confirmation and other services for the application layer. the Distributed storage resource layer uses virtualization technology to realize the integration and pooling of storage resources. The infrastructure consists of storage capacity miners and storage scheduling miners. The distributed content distribution layer uses content caching technology to realize the fast push of stored data, which is composed of data index miner and data distribution miner. By API node of Application layer, CESS can realize data storage service, blockchain service to support enterprise level SDK, data storage network disk and computational intelligence applications, etc.

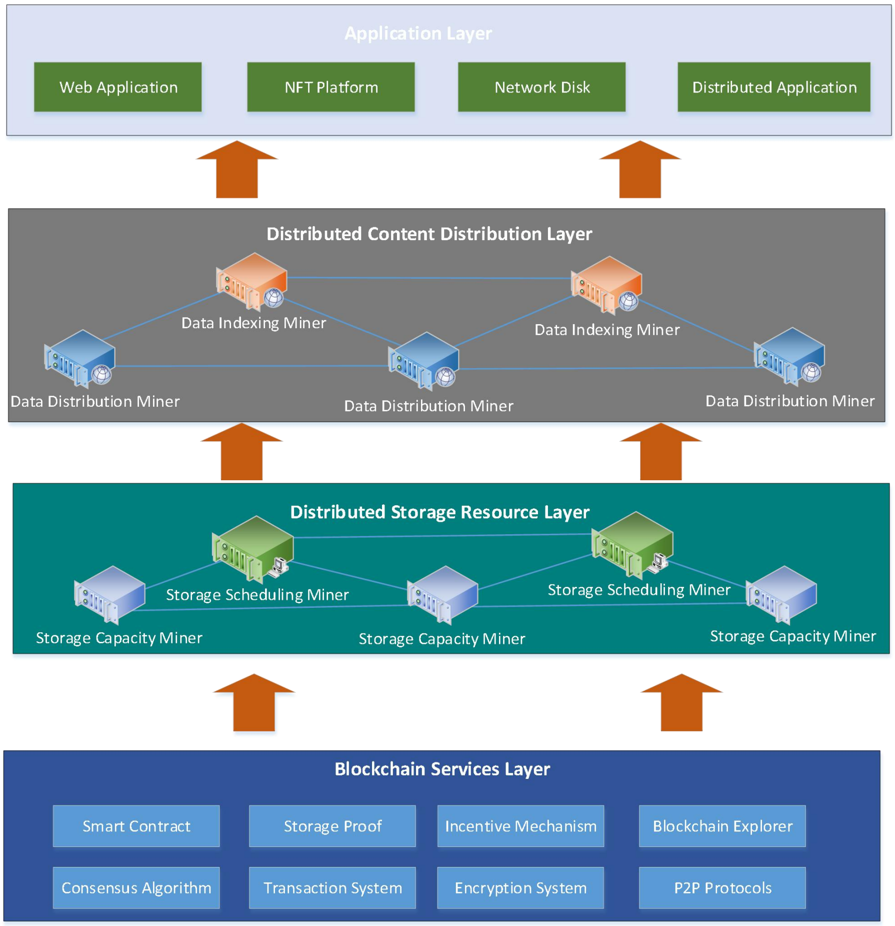


Figure 1 Design schematic diagram

High performance of blockchain service: In addition to encouraging idle computing resources and storage resources to join the distributed storage network and distributed content buffer network, it can also provide efficient blockchain services. CESS will make use of asynchronous Byzantine consensus algorithm technology to provide high TPS access capability, as well as Turing's complete smart contract and cross-chain transaction capability.

Distributed storage resource layer: Using the incentive mechanism of blockchain to effectively manage the storage capacity of servers, desktops, laptops and other storage capacity devices, providing massive data storage capacity for the world is the most critical hardware infrastructure layer of the whole system. This part is composed of storage scheduling nodes and storage capacity miners. Storage scheduling node stores data metadata, provides fast data index, and storage capacity provides data storage space. Users' data stored on the nodes of distributed storage resource layer is usually stored on the storage capacity in the form of encrypted slices, so it will not cause user data leakage.

Distributed content distribution layer: The use of decentralized cloud storage technology will lead to more decentralized data than the traditional cloud storage data center, which leads to the problem of slow data access by users. How can CESS solve this problem? We use the design idea of IPFS for reference and introduce the content buffering technology to provide efficient data storage services for the world. This layer network is an efficient content distribution network, which is composed of data distribution miner and data index miner. The data distribution miner is responsible for buffering data, and the data index node is responsible for efficiently querying data.

Application layer services: provide user-friendly and convenient data storage and blockchain services, such as network disk, cloud disk, etc.

5.2 Blockchain Layer Design

CESS uses a multi-tiered architecture, is scalable and robust, and provides a friendly complete API interface and SDK for developers to connect.

As shown in Figure 5, the overall architecture is divided into six layers: infrastructure layer, data layer, network layer, consensus layer, incentive layer and application layer. The infrastructure layer will develop dispersion hardware equipment including service, network and storage resources for CESS network. The data layer, which supports scalable data storage, uses some algorithmic techniques to optimize the use and utilization of resources and the security of data; Network layer for node connection, data transfer. Provides load balancing and P2P network protocols and algorithms; Consensus layer, using consensus algorithm to achieve rapid consensus on transactions, with 100,000 of TPS processing capacity; The incentive layer uses a variety of storage algorithms to achieve fair income distribution through smart contracts, thus forming positive feedback and motivating the entire CESS technology to move forward. Interface layer, which provides rich API interface and perfect SDK, makes it easy for developers to dock. Application layer, supporting DAPP or APP applications developed by third-party developers.

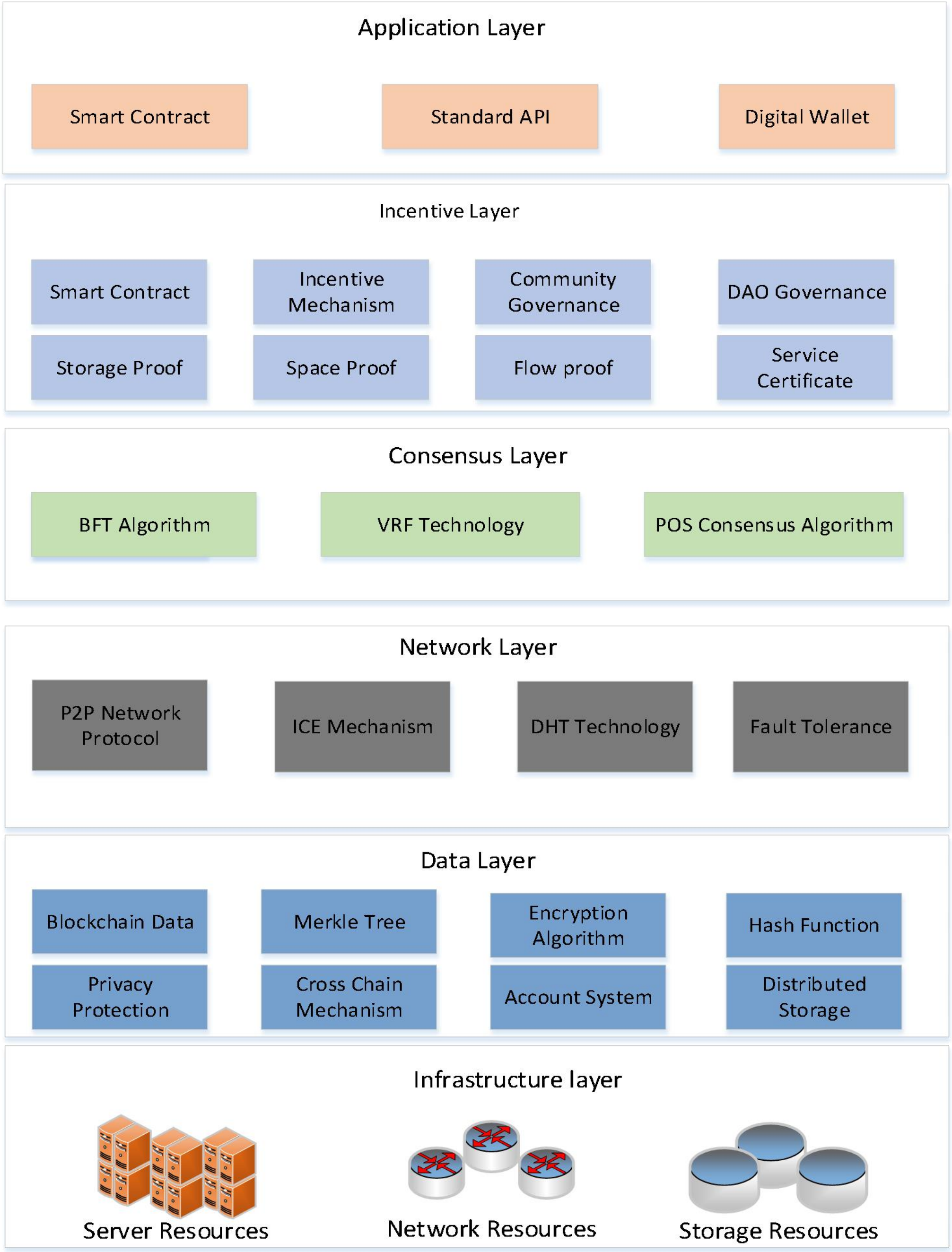


Figure 2 Overall Technology Stack of CESS

5.2.1 Infrastructure Layer

In order to support the upper layer of application stably to access storage resources, CESS needs to construct stable and reliable infrastructure facilities. According to Figure 2, the CESS system needs to solicit three types of resources from all over the world, such as Servers type, Network type and Storage type. Server-type resources focus on computing performance and will carry the computing and task scheduling tasks for CESS network. Network-type resources will provide network bandwidth support for CESS network. In order to support users' global undifferentiated data access, the CESS system will use nodes with network bandwidth to construct distributed content distribution network to accelerate data access. Storage resources are the key part of the whole CESS system, and the soul of the entire CESS system. It will request a large part of the team's operational strength, attract the capacity nodes with storage capacity to join the network, provide a stable and reliable data storage infrastructure for the CESS system, and lay a solid foundation for unified scheduling and management of CESS platform resources.

5.2.2 Data Layer

The data layer contains block data, data stored in the P2P network, and so on. To ensure the security and integrity of the data, encryption algorithms will be used for data transmission, storage and verification, such as digital signatures, hash algorithms, Merkle trees, and so on.

Block data: Chain data, which records transaction records over the entire public chain network. Some nodes need to save block data and run the whole node to ensure the security and stability of the public chain.

Distributed storage: Stores data on separate devices. It provides efficient, robust and load balanced file access. As shown in Figure 3 below:

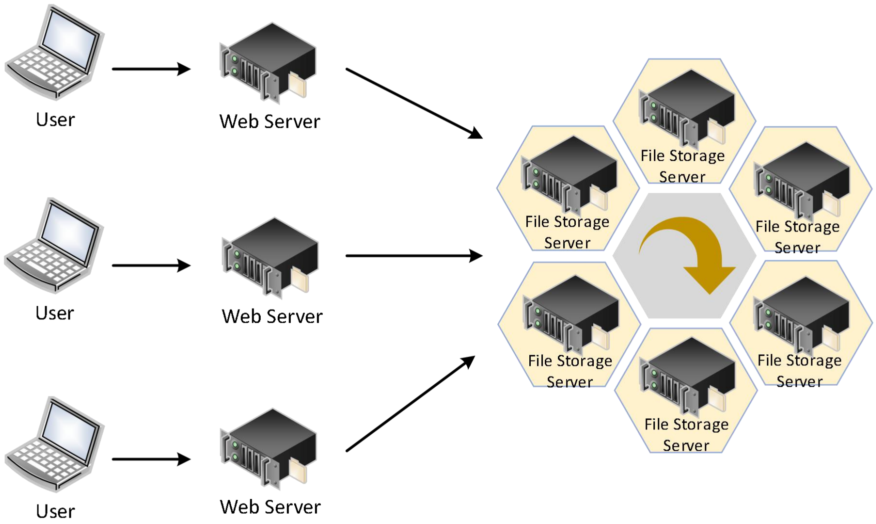


Figure 3 Distributed Storage System

Digital signature: A digital signature (also known as a public key digital signature) is a digital string that can only be forged by the sender of the information. It is also effective proof of the authenticity of the information sent by the sender. It is similar to a physical signature written on paper, but it is implemented using public key encryption technology to identify digital information. A set of digital signatures usually defines two complementary operations, one for signatures and the other for verification. We use digital signatures for system authentication, data integrity verification, etc.

Hash algorithm: A one-way cryptography, that is an irreversible mapping from clear text to cipher text, only the encryption process, no decryption process. A hash function can change any length of input to get a fixed length output, but different inputs have different outputs. This one-way characteristic of the hash function and the fixed length of the output data allow it to generate messages or data. Common hash algorithms are MD5, SHA-1, SHA256, etc. We use a hash algorithm to uniquely identify the data and ensure that it is not tampered with.

Asymmetric encryption: An encryption algorithm that uses a different key for encryption and decryption, also known as public-private key encryption. The public key is a pair of the private key. If the data is encrypted with the public key, only the corresponding private key can be de-crypted. Because encryption and decryption use two different keys, this algorithm is called an asymmetric encryption algorithm. Commonly used are RSA, ECC, etc. Digital signatures are an application of asymmetric encryption.

Merkle Tree: A Merkel tree (also known as a hash tree) is a tree that stores hash values. The leaves of the Merkle tree are hash values for data blocks, such as files or collections of files. A non-leaf node is a hash of its corresponding child node concatenation string. Get the Merkle tree root of the file from a trusted source before downloading data on the P2P network. Once you get the root, you can get the Merkle tree from other untrusted sources. Check the received Merkle Tree through trusted roots. If the Merkle Tree is corrupt or false, get another Merkle Tree from another source until you get a Merkle Tree that matches the trusted tree root. We can download and immediately verify a branch of Merkle Tree. Because files can be divided into small blocks, if a piece of data is corrupted, simply download it again.

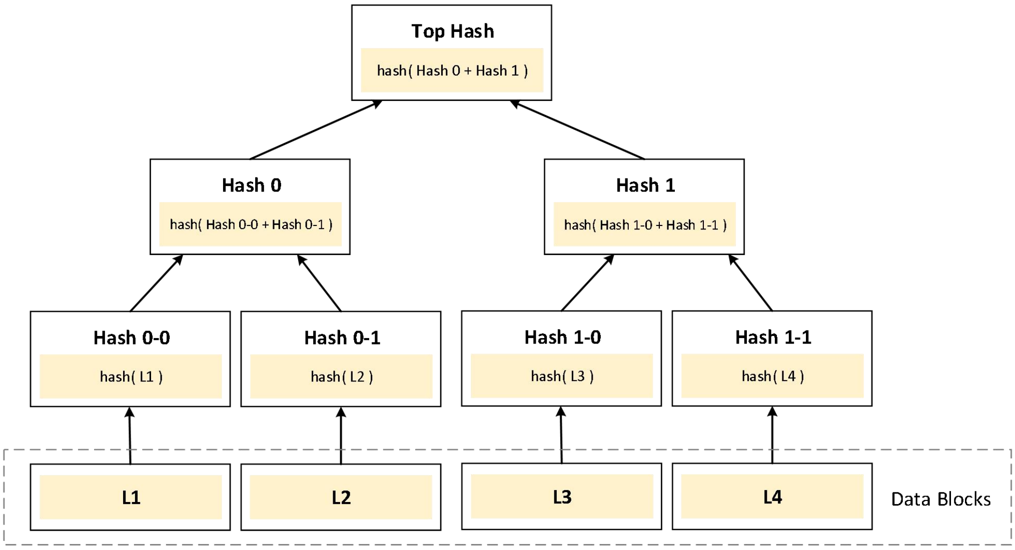


Figure 4 Merkel tree data structure

5.2.3 Network Layer

To ensure efficient access to data in the network, a DHT-based P2P storage network will be built.

P2P network: A peer-to-peer computer network is a distributed application architecture that distributes tasks and workloads among peers. It is a form of networking or network formed by the peer-to-peer computing model in the application layer. In a P2P network environment, multiple computers connected to each other are in the same position. Each computer has the same functions, and has no master-slave. A computer can act as a server, set up shared resources for use by other computers in the network, and can also act as a workstation. Generally, the whole network does not depend on a dedicated centralized server. There are no dedicated workstations. Each computer in the network can act as both a requestor for network services and respond to requests from other computers to provide resources, services and content.

DHT: A distributed hash table is a distributed storage method. Without the need for a server, each client is responsible for a small range of routing and for storing a small portion of the data, thus enabling the addressing and storage of the entire DHT network. Users who connect to a DHT network are called nodes, and there is a routing record between the nodes, so as long as they are connected to any node already in the DHT network, the client can find more nodes to connect to the network. DHT technology is to enable any machine in the network to perform part of the server's functions, so that users' data are no longer dependent on the server.

ICE: ICE is an object-oriented middleware platform for communication between nodes. ICE provides an RPC protocol that can use either TCP/IP or UDP as the underlying transport mechanism. Nodes do not need to know their implementation. ICE also allows SSL to be used as a transport mechanism to encrypt all communication between nodes.

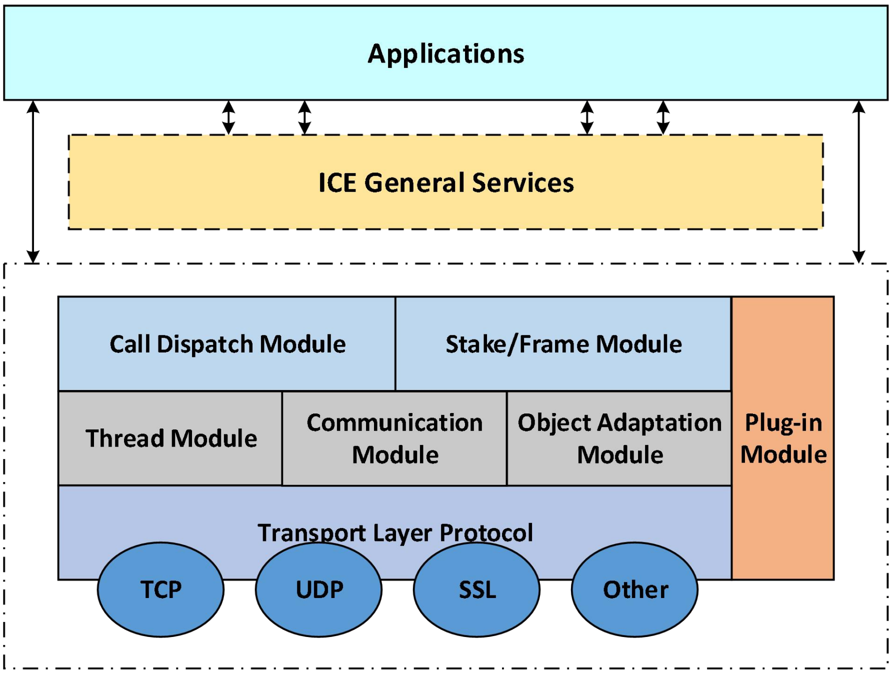


Figure 5 ICE framework

5.2.4 Consensus level

In order to ensure the transactions and activities on the blockchain network can reach a consensus quickly, CESS will adopt a superior consensus mechanism based on Byzantine fault tolerance, which further improves the performance and scalability of the system. Theoretically, any node can be one of consensus node. There is no limitation to the of amount of consensus nodes. Any node can access into the consensus system as stand-by consensus nodes in CESS chain. In order to improve reliability, CESS will use credit ranking module to verify the nodes and select the qualified node to be stand-by consensus nodes. Subsequently, within each time-line, CESS selects next 11 Consensus nodes as rotating nodes by Verifiable Random Function under the current time-line. Rotating nodes will have responsibility for on-chain packing, transaction, block out to attain fairness, transparency and security by random consensus selection mechanism. Here are the Modules mentioned

Credit rating: the consensus mechanism adopts the health score evaluation model. The module evaluates the health score of the consensus behavior of distributed energy nodes in the platform, the health score of the nodes showing honest & loyalty behavior, which provides the basis for the election of the consensus node committee when changing views, so as to improve the reliability of the consensus node and ensure the authenticity of the data on the link.

Verifiable random function: using the block node election protocol based on verifiable random function, CESS improves the way of blockchain generation nodes generate the block from the traditional PBFT residual election method. To increases the randomness and unpredictability of the node election will improve the anti-attack ability of the master node and ensures the security of the CESS platform based on blockchain.

Consensus fragmentation: the adoption of consensus fragmentation protocol helps to reduce the overall load from CESS system, reduce the expense of communication overhead in the network transaction process, reduce the delay and improve the transaction throughput. Compared with the expansion scheme under the chain, the fragmentation scheme ensures the transaction on the chain and gives consideration to the decentralization to a certain extent.

Dynamic node: constructed a dynamic consensus node protocol, which includes new node join protocol and active exit protocol. The join and exit mechanism of system consensus node is a supplement to the consensus algorithm process. It realizes the function of dynamically adding and deleting nodes in the scenario of chain network cluster without downtime. CESS improves the robustness of CESS and further expands the practical application of CESS.

5.2.5 Incentive layer

As a distributed file system, CESS mainly resources system are storage and network resources. In the CESS network, there will be two types of mining nodes: Content Storage Node (CSN) and Content Delivery Node (CDN). CSN node storage is responsible for file storage, CDN node is responsible for file propagation. Therefore, users can provide two types of resources to join the CESS network, and the CESS system will reward the node with CESS tokens according to the contribution generated by the nodes.

In order to encourage users to join and remain with the CESS network, it is necessary to design an incentive mechanism for the CESS network, providing rewards according to the contributions generated by the nodes.

- How to partner with the CESS network

For users, rewards are similar to "crypto mining". The reward provided by CESS is based on the contribution of users to the distributed network and the contribution of proof (COP) by algorithm, thus determining the contribution from CESS. The Contribution Proof Algorithm is a comprehensive consensus algorithm, which mainly considers the factors of miners' storage capacity, network bandwidth and machine configuration, subsequently calculating a comprehensive score, andyielding a reward from CESS based on score.

- How to earn a CESS Token

In order to promote the network development of CESS, and in addition to the incentive system at the technical level, CESS introduces the main node incentive mechanism, and users who meet the incentive mechanism will be issued CESS token.

1）Storage mining: Miners can join the distributed storage network and distributed content network to earn Tokens. Stakeholders earn a corresponding incentive pro-rata to bandwidth and storage capacity.

2）Consensus mining: any nodes who qualified by computing resources can automatically be consensus node after staking a certain amount of CESS, consensus nodes can earn the block rewards by on duty after being randomly selected.

3）Community contribution: To further promote the CESS network, developers, community members and partners can submit their proposals and receive community votes. When a quorum is established, corresponding rewards will be issued from the blockchain system.

4）Token governance will adopt the Decentralized Autonomous Organization (DAO) to achieve on-chain governance capability. CESS community transparently operates the Community Development Fund (CDF) through the votes.

- Token Distribution

Because CESS is a distributed, decentralized network storage system. All operations related to incentive are implemented based on smart contracts. It includes the user's time and space proof, the chain operation of authentication results and reward distribution. This ensures the fairness and transparency of the CESS incentive mechanism.

5.2.6 Interface Layer

One of the core design goals of CESS is to provide programmable distributed storage, also known as Distributed Storage Services (DSaaS), providing a user-friendly environment.

- SDK

CESS provides APIs on a variety of platforms, such as iOS, Android, Mac, and Windows.

- Web API

Storage applications

Help developers develop Web-based applications.

- json-rpc interface

Allows DApps to invoke functions on CESS nodes to easily integrate the CESS storage system.

- Application Sandbox

CESS can support running a large number of applications concurrently in its storage network. Developers can configure files in their applications.

1) Each application has its own encryption key. Application developers can specify how to encrypt data objects in their applications:

2) Encrypt using only the application's encryption key.

3) Encryption using the encryption keys of the application and the user.

CESS provides two categories of API. The system-level API provides developers with greater flexibility and control over the various functions of CESS. The contract-level API provides developers with a variety of capabilities for storing purchase queries.

- System-level API

1) get: Gets the resource object for the specified hash value.

2) cat: Gets the resource object data stream for the specified hash value.

3) delete: Delete a resource with a specified hash value.

4) add: Increase resources.

5) push: Increase and backup resources.

6) Callback: Get the URL of the download resource.

- Contractual level API

1) Get_Account: Get information about the account in the contract.

2) Get_Table_Rows: Get motivation and store relevant information.

5.2.7 Application Layer

The API for CESS is designed to enable third-party developers to build a variety of applications, including the following:

- Private data storage

Because of its de-centralization, CESS storage networks are well suited for private network storage applications.

Personal data is partitioned, encrypted, and stored on different nodes to ensure privacy is protected. At the same time, access to data is restricted by the user's private key, which makes personal data more secure.

- Enterprise Data Storage

CESS provides high-performance services for enterprise data storage with significant low cost.

- DApps

Storing application data on a blockchain is expensive for de-centralized applications. Data from smart contracts or other applications can be stored on CESS storage nodes using the CESS API, which can result in significant cost savings.

- Media Applications

CESS provides low-cost bandwidth resources, which can effectively reduce the cost of content distribution. CESS is also equipped with specially optimized scheduling and transmission algorithms for smooth data transmission, enabling and maintaining a high-quality user experience for media applications.

- Data Exchange

File assets can be traded over the CESS network. CESS can provide methods to match sellers and buyers and handle transactions safely and reliably without requiring an intermediary. Common applications, such as the application marketplace and content platforms, can benefit from CESS.

- Database

CESS can be used as an enterprise database to store large amounts of historical data, replacing traditional local data storage or expensive cloud storage. In addition to enterprise data, CESS can also be used to store public databases.

CESS will also provide the necessary support for other types of storage requirements. In addition, CESS will soon open-source code. At that time, application enthusiasts and developers will be able to participate in the development of CESS and add support for more applications.

5.3 Distributed Storage Resource Layer

Unlike existing IPFS, and other projects, CESS is designed to build a block-chain-based distributed cloud storage system. The focus is on providing unified and efficient distributed storage services to the outside world by effectively managing distributed resources using virtualization technology. CESS is powerful because it has efficient global data storage capabilities and enables users to access data indifferently through distributed identity information. In terms of implementation, CESS will build two types of infrastructure: distributed content buffer and distributed cloud storage. The distributed content buffer network will deliver data to the nearest content buffer node according to the user's geographic location to speed up access. Distributed cloud storage networks are primarily designed to provide massive, reliable, and scalable cloud storage services.

- Data Storage Process

The process of storing data to CESS network by users will go through several stages, such as production, upload, processing, storage, distribution and destruction. In the production stage, users can implant applications through restful API, SDK and other means to upload data; In the storage phase, based on CESS network resources, intelligent services for pictures, videos and documents can be built to support users to process data online. During the distribution phase, over 10T of network bandwidth can be achieved through a content distribution network. In addition, CESS supports users to delete data online.

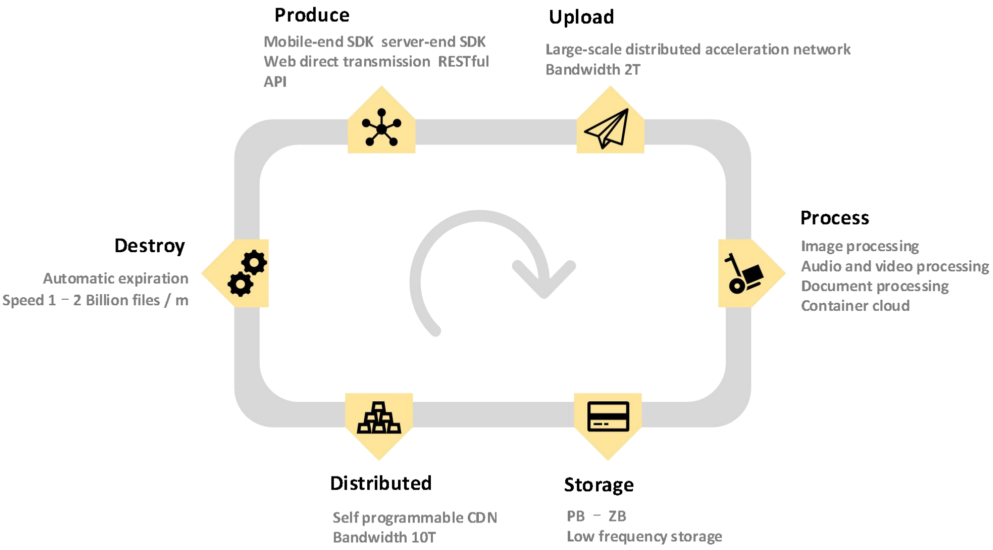


Figure6 Data Storage Process

- Distributed Content Cache Network

In order to achieve efficient file access, the system effectively combines the advantages of both CDN and P2P technologies, forming a content distribution network technology, which effectively reduces the number of proxy servers required by the system, increases the capacity of the system, reduces the overall cost, and uses CDN technology to transfer media content to the client's autonomous domain. Enhances the quality of media access for customers, and improves P2P network performance in a smaller autonomous system. The presence of a high-performance cache proxy server also avoids the "seed" problem in pure P2P networks.

At the same time, on the application side, the stored content in the application will be published on the publishing source node first, and the download service will be continuously provided if the source node is not offline. However, as the number of users downloaded from the same source node increases, the bandwidth of that node will be exhausted and the download speed per user will be reduced. With the design of a content distribution network, a large number of tenant nodes in the network begin to save and provide downloads of the same content. As a result, users can download content from multiple nodes, which greatly improves the user experience.

The overall design of the Distributed Content Distribution Network Layer is perfectly combined with block chain technology. Storage nodes form CDNs with proxy nodes in each region. Proxy nodes form a relatively independent P2P network with the following storage nodes without public network IP. Node contribution awards are issued through smart contracts, forming an autonomous network for development, as shown in Figure 7:

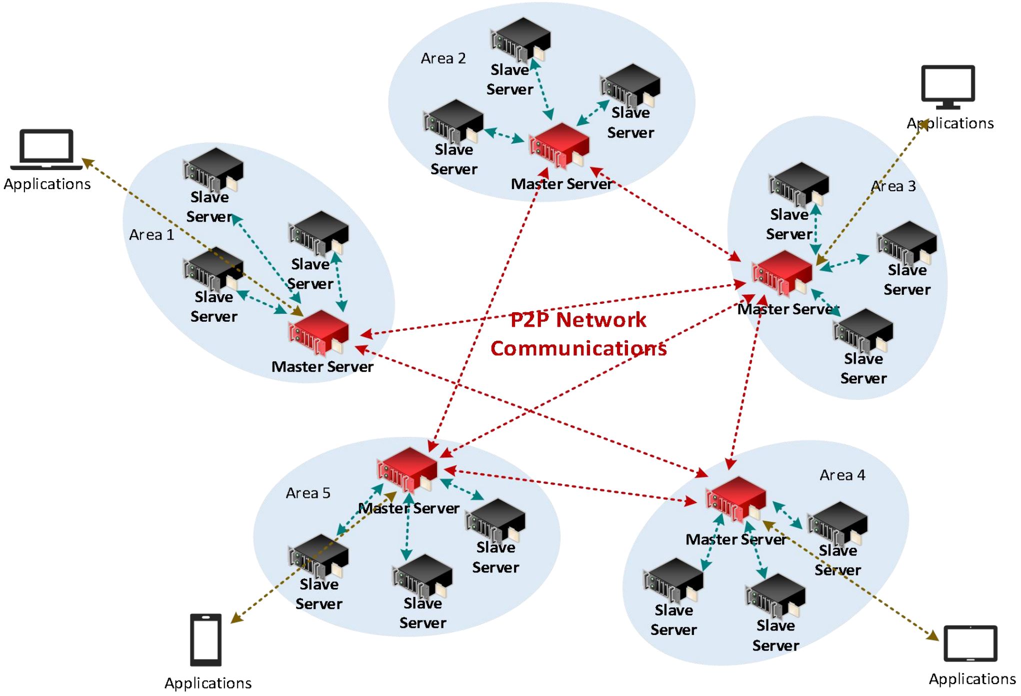


Figure7 Distributed Content Cache Network

- Distributed Cloud Storage Network

To meet different storage needs, we will design and implement a polymorphic data storage access interface to provide storage services in the form of APIs for a variety of applications. As shown in the figure 8 below, on top of the unified distributed object storage engine, the polymorphic data access service provides object storage, block storage and file system storage for the upper application in a standard API way, providing a comprehensive and friendly data storage service support for the top application.

CESS will provide an improved and reliable object storage service. The upper application calls the object storage service interface. The object storage module automatically completes the mapping of the user object storage space to the lower unified distributed object storage space. User data is stored in the distributed object storage engine as object data.

CESS will provide the block device storage service. The upper application calls the block device service interface. The block storage module automatically completes the mapping of the user's block device operation, data read and write operation to the unified distributed object storage space at the bottom. The user's data on the block device will eventually be stored in the distributed object storage engine as object data, supporting snapshot, cloning and other functions.

For generic file systems, the POSIX file system module provides a POSIX-compliant file system interface, supports both kernel file system and user space file system (FUSE) modes, and calls the POSIX file system interface by upper applications. The POSIX file system module and the POSIX file system metadata manager (responsible for mapping and transforming POSIX file system space to object storage space) jointly complete the mapping of user's POSIX file operations to the underlying unified distributed object storage space. User's data in the POSIX file system is ultimately stored in the distributed object storage engine as object data.

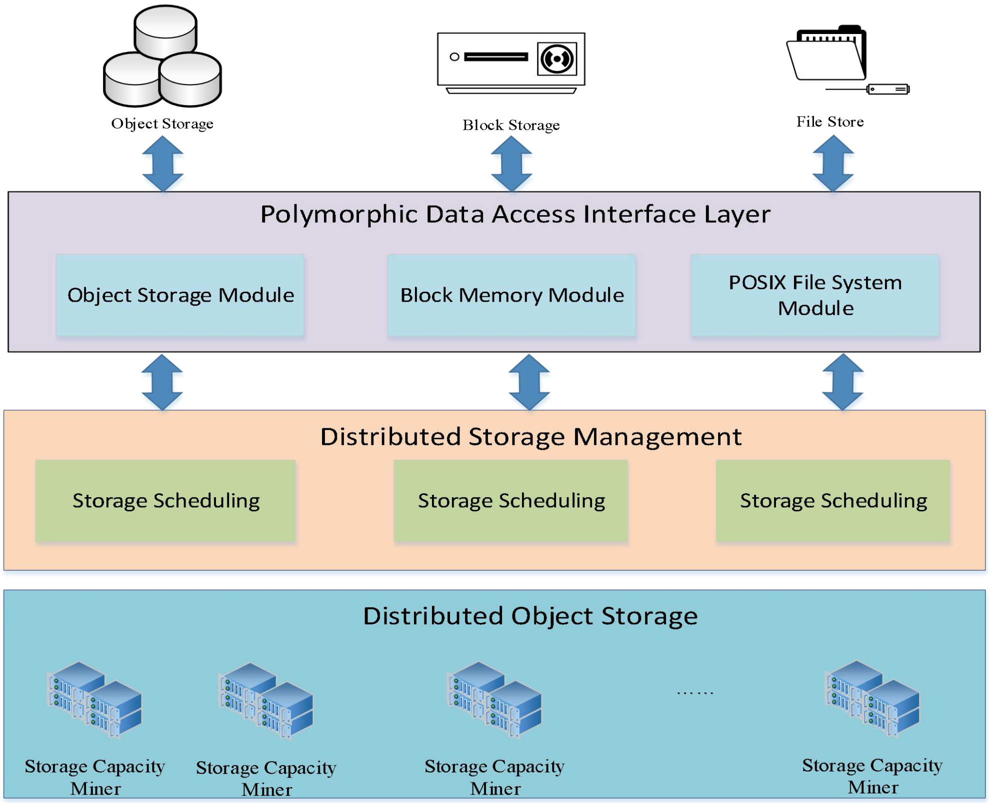


Figure 8 Distributed Storage Network Architecture Design

5.4 Key technology 1: Data tampering prevention mechanism

To be a distributed cloud storage system, the security and stability of user data integrity that is tamper-proof is the foundation of the system. CESS designed and implemented a Data tamper-proof mechanism that fuses blockchain technology and Anti-collision hash function. It cooperated with the PoRep algorithm and the space-time proof algorithm (PoSt) to provide the support for trusted storage and data safe transfer. Subsequently, blockchain chain data traceability and tamper-proof characteristics strengthen the authority of digital fingerprints on the chain, to solve the distributed storage cloud platform digital fingerprints which are easy to lose, difficult to monitor, and afraid of losing packets.

Because the stored data in CESS exists in the form of slices, there is no readable and usable independent data file. Therefore, for an efficient and unified process, if the slice data capacity is small, it is not necessary to set the corresponding data fingerprint scheme for different data types. That is to say, CESS can use unified anti-collision hash function to process individual slice data separately and generate corresponding data fingerprint. If the slice data is relatively large, we need to choose corresponding fingerprint generation algorithm according to different data types. There are text data, image data, audio data and video data on CESS platform. The basic process of fingerprint extraction algorithm includes text collection, text preprocessing, vector space identification, latent semantic space construction and fingerprint generation. Image data mainly use color histogram-based feature extraction algorithm, through calculating the histogram of YUV component of video frame separately, quantization linear fusion feature extraction. Audio data are extracted with an algorithm based on entropy to enhance robustness against noise and distortion. Video data will be extracted by color, texture, shape and brightness of video to achieve strong robustness and discrimination of video fingerprint features.

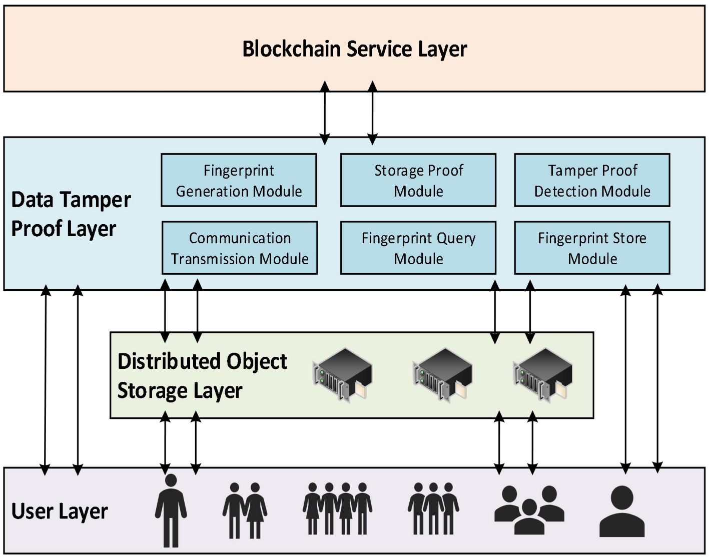


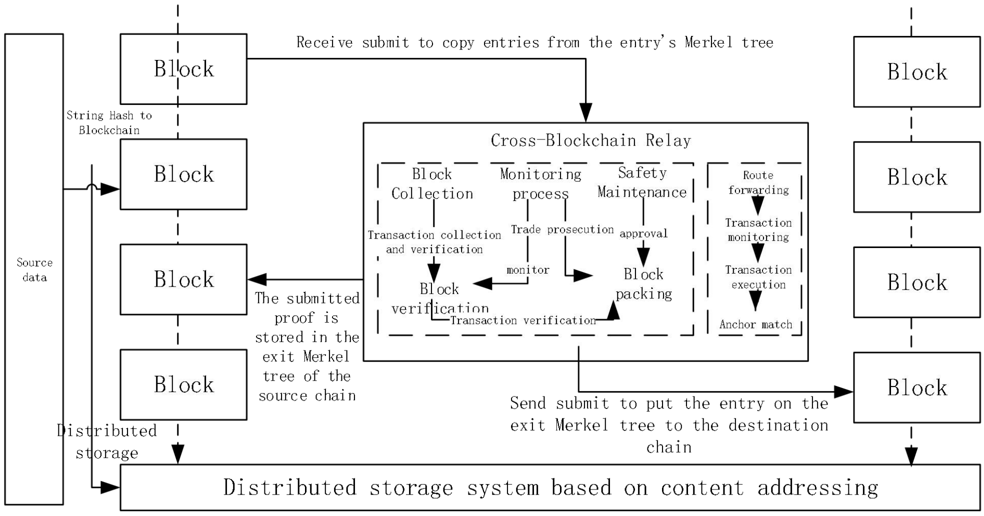
Figure 9 overall architecture of data tamper proof mechanism

Large data files in the data transmission process because of various unavoidable factors which can affect the data consistency. In addition, the reliable transmission technology of communication layer, including erasure code mechanism and retransmission mechanism, etc, is required to complete reliable transmission of data to avoid source data fingerprint mismatch. Furthermore, in order to solve the single-point failure problem, CESS uses multi-copy storage mode, which may cause redundant storage and repeated calls. Therefore, the fingerprint data query function is introduced into the data tamper-proof mechanism, and no further operations are carried out for the fingerprint data on the chain. The overall architecture is shown in Figure 9. The data tamper-proof layer serves the user layer and the distributed object storage layer. It includes the modules of fingerprint generation, storage proof, anti-tamper detection, communication transmission, fingerprint query and fingerprint up-chain, and provides data pre-processing and anti-tampering detection mechanism.

5.5 Key technology 2: cross chain mechanism

Following single chain, multi-chain coexistence and even multi-chain collaboration are the inevitable way forward. trend after single chain. However, due to the different technical structure of private and public blockchains, all chains are disconnected and run independently. The isolation of the network hinders the cooperative operation between different blockchains and greatly limits the space for the development of blockchains. Cross-chain technology is an important link to provide services for communication inter-blockchain between each other, which is the key to the realization of value network and is the bridge between the blockchain to expand outward and connect with each other. Based on the research of cross-chain technology, CESS system implements an integrated cross-chain mechanism that supports cross-chain interaction between isomorphic and heterogeneous chains, supports asset flow, information exchange and application collaboration between different blockchain platforms. It is similar to a bridge between different public chains, achieves data transmission between different blockchain networks, and greatly reduces transmission costs.

In CESS systems, cross-chain relay and parallel chain are the infrastructure to achieve cross-chain interaction between isomorphic and heterogeneous chains, respectively. The security mechanism, consensus algorithm, network topology and block generation validation logic among homogeneous chains are consistent, and the cross-chain interaction between them is relatively simple. Generally speaking, entry and exit based on cross-chain relay can achieve transaction forwarding and data exchange between different chains. However, the cross-chain interaction of heterogeneous chains is relatively complex. The validation of messages between heterogeneous chains and the cross-chain interaction without trust can be achieved by the parallel chain-based state verification method and synchronization consensus. The cross-chain interaction mechanism of the CESS system is shown in figure 10.



Source data
Block
String Hash to Blockchain
Block
Block
Block
Distributed storage
Receive submit to copy entries from the entry's Merkel tree
Cross-Blockchain Relay
Block Collection
Monitoring process
Safety Maintenance
Transaction collection and verification
Trade prosecution
approval
monitor
Block verification
Block packing
Transaction verification
Route forwarding
Transaction monitoring
Transaction execution
Anchor match
The submitted proof is stored in the exit Merkel tree of the source chain
Send submit to put the entry on the exit Merkel tree to the destination chain
Block
Block
Block
Block
Distributed storage system based on content addressing

Figure 10 CESS Cross Chain Mechanism

Safe connections and free transactions between homogeneous blockchain can be achieved through routing forwarding, transaction monitoring, transaction execution and anchoring and matching functions of cross-chain relay. The mutual recognition and trust of heterogeneous block chains can be achieved through the protection of parallel chain transaction verification, transaction verification, transaction collection, transaction reporting and other mechanisms. The cross-chain mechanism of CESS system focuses on the basic cross-chain protocol standard. It achieves standardization in many aspects, such as input and output caliber of cross-chain messages, authenticity certification of cross-chain messages, unified format of cross-chain messages, validity certification of messages and cross-chain execution result certification, and provides a solid technical support for cross-chain security sharing.

6. Unique technology

6.1 Cross-Chain data authentication and communication technology

CESS proposes a cross-chain data authentication and communication technology to support the implementation of a complete blockchain cross-chain interoperability process, not only to solve the underlying basic data mutual recognition problem, but at the same time to build the inter-chain communication path. In the cross-chain process, identity protocol supports the self-description of blockchain, including the identity of its own chain, the type of chain, data model, data content, proof on the chain, verification root and so on. Through identity protocol, cross-chain interaction can establish mutual recognition foundation. Without changing the original data semantics, the transformation protocol can transform the data format, the data proof and the trust root, and provide a unified cross-chain data format and cross-chain proof selection for multi-party. Finally, the data protocol coordinates the mutual recognition of each chain data, which helps to complete the data analysis and authentication requirements. The communication protocol supports the smart contracts on the chain to reliably send cross-chain messages to each other for information communication. Data authentication and communication architecture for cross-chain interoperability is shown in Figure 11.

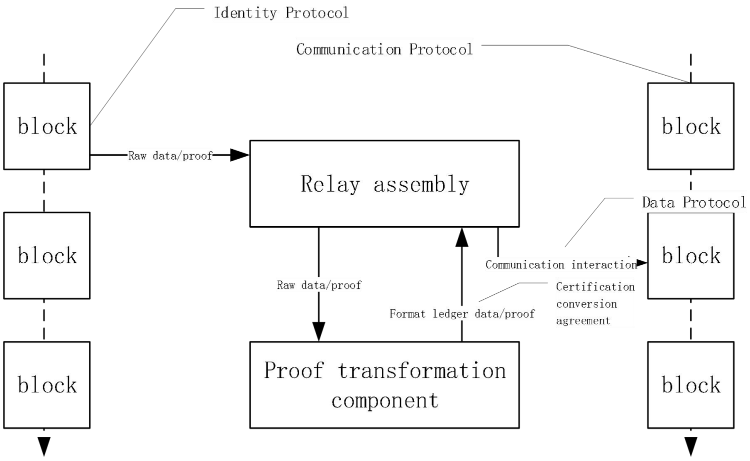


Figure 11 Cross-chain data authentication and communication architecture

Based on cross-chain data authentication and communication technology, in the cross-chain system of CESS, the blockchain cross-chain interoperability process is as follows: each blockchain first obtains a unique chain identity based on the identity protocol, which is used to represent the sender or receiver of cross-chain information. The ledger transmitting chain generates normalized and lightweight self-describing data packets after the relay component and the proof transforming component, and sends the ledger data to the corresponding receiving chain according to the message format and message flow defined by the communication protocol. A relay component in the receiving chain extracts the ledger and proof to the system on the chain, which performs the ledger verification and related business execution operations.

6.2 Data tamper-proofing technology

In order to maintain the benign operating state of the system, data owners regularly check their data integrity, or data users call data verification services before using data which are inseparable from the data tamper-proof detection mechanism. Therefore, we propose a data tamper-proof technology that combines blockchain technology and anti-collision hash function to achieve efficient data anti-tamper function detection and improve data storage security from multiple angles.

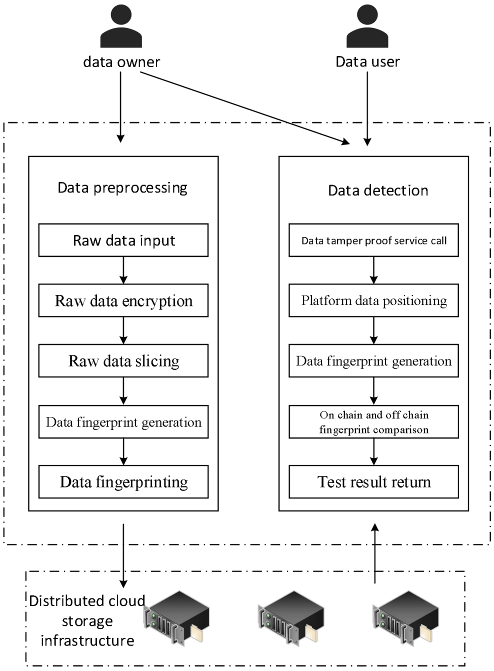

data owner
Data user
Data preprocessing
Raw data input
Raw data encryption
Raw data slicing
Data fingerprint generation
Data fingerprinting
Data detection
Data tamper proof service call
Platform data positioning
Data fingerprint generation
On chain and off chain fingerprint comparison
Test result return
Distributed cloud storage infrastructure

Figure 12 Data tamper-proofing detection service

The operation before CESS data is stored and up-chained can be called data preprocessing. In order to ensure the privacy of stored data, data preprocessing includes data encryption and slicing operation. Each distributed cloud storage infrastructure only saves encrypted partial slicing data. Therefore, the integrity of the original data file can be proved as long as the slice data integrity is guaranteed. Specifically, before transmitting the slice data to the distributed cloud storage infrastructure, CESS processes the slice through a collision-resistant hash function, and passes the generated data fingerprint through an on-chain data retrieval service to check whether the fingerprint information already exists. If it already exists, it means that the fingerprint information has been recorded on the chain and the operation is abandoned. Otherwise, fingerprint information is on-chain through smart contracts. The sliced data is then sent to the associated distributed cloud storage infrastructure store.

Data tamper-proof detection is widely used. Data owners have the right to check the integrity and tamper-proofing of their stored data in real time. Data users should also check the integrity of the data to be used when using third-party data. According to the spatio-temporal proof algorithm, every distributed cloud storage infrastructure needs to regularly send the integrity proof of its stored data in order to ensure its credibility and revenue are not damaged. All these cases will inevitably use data tamper-proof detection technology. Data tamper-proof detection services include platform data location, data fingerprint generation, chain up chain down comparison fingerprint and test result return, which provides CESS with high efficiency and high accuracy detection capability.

6.3 Multi-policy Storage Capability of Proof

All we need to focus and try to solve is to ensure the security of system especially the integrity which is most important for information system. In a de-centralized storage system such as CESS, the de-centralized attribute provides more privacy for data storage, reduces storage costs, and improves transmission speed. However, the data in the system is also facing security and integrity issues. We hope that CESS is a benign and self-governing ecosystem, so we will use non-centralized algorithms to ensure the normal operation of the system, while ensuring the fairness and impartiality of reward and punishment mechanisms in the system. The main storage proof algorithms used by CESS are PoRep, PoSt, PoF, and PoAs.

First, when a user stores a file to the CESS network, the system uses the PoRep algorithm to ensure the consistency and integrity of the user's file. The PoRep algorithm is an interactive proof algorithm that stores the node to provide a storage certificate to the verification node, proving that the user's data has been copied and stored on the dedicated physical storage device of the storage node. The algorithm has the ability to defend and retaliate against Sybil attack and Exogenous attack.

The system not only validates the file when it is uploaded, but also uses the space-time proof algorithm (PoSt) to ensure the integrity and recoverability of the file during the contract period for file storage. Spatio-temporal proof algorithm (PoSt) is an algorithm that measures and calculates the time and space of data storage stored in a network. Spatio-temporal proof can be understood as continuous replication proof. Storage nodes must continuously generate certificates, and periodically submit certificates. If the storage nodes do not submit certificates in time during the submission cycle, the nodes will not be able to get rewards for the period, and the credit score of the nodes in the system will be reduced. In addition to the replication and space-time proof algorithms, the CESS system will analyze and statistics the traffic of nodes to determine the contribution of nodes to network transmission and sharing. Therefore, the CESS system will use the traffic proof algorithm (PoF) to measure and calculate the traffic contributed by nodes in the network.The traffic proof algorithm (PoF) uses the mechanism that nodes certify each other. Each node needs to periodically submit traffic data (data records of the interaction between this node and other nodes) to the validation node, which will be used for statistical analysis to determine the contribution of the node to system traffic.

CESS is a decentralized distributed network with strong autonomy of nodes. Considering the fluctuation of network environment, the system will use the Storage Available Authentication Algorithm (PoAs) to ensure that nodes have sufficient storage capacity and stability. On the one hand, CESS uses periodic monitoring nodes to judge the node status, and on the other hand, CESS uses the Storage Available Authentication Algorithm (PoAs) to determine the node status. The system evaluates the storage capacity in conjunction with the node's credit score. When a user stores a file, the system chooses a node with a higher rating to store the data.

In addition to the above proven algorithms, CESS system uses other algorithms and technologies to maintain the security and stability of the whole system.

6.4 Data Rights Protection Mechanism Based on Smart Contracts

CESS, as a content sharing platform, is committed to return the ownership of users’ data to users, so that users can enjoy the value of their own data. CESS protects the privacy of users' data. CESS implements an automated, fair and transparent data rights protection system based on blockchains and smart contracts, ranging from rights confirmation, rights tracking to rights maintenance.

CESS will always prepare two smart contracts for each data benefit model for users to choose from. When publishing data, users only need to set the value according to their own wishes. The system generates the rights and interests' attributes of the data according to the user's settings. The equity attributes of the data mainly include income model, whitelist, blacklist and so on. These strategies are published with the data as equity attributes of the data. Each time the data is accessed, a smart contract in the data's equity properties is enforced to ensure that it is executed according to a strategy developed by the data owner.

When other users need to access the data, they have to apply for data rights first. The system will have more rights and interests' attributes of the data. Check and verification if the data can be accessed by the applicant account and If the check is passed, the system will execute the fee according to the income model in the data rights and interests' attributes and return the data to the applicant.

In order to promote the liquidity of the data, when the users request access to the data, they can also set the value of the benefit model contract themselves. The system will follow the whitelist strategy to be set by the data owner or wait for the owner to decide. If passed, the income model of the applicant and the applicant will be recorded in the whitelist attribute of the data. As shown in the figure 12 below, CESS's data rights protection system, with the help of blockchain traceability, scans random transactions and has a dedicated evidence library module. The evidence library provides a set of interfaces that allow users to view access records of data at any time to provide open and accurate evidence for resolving data rights disputes.

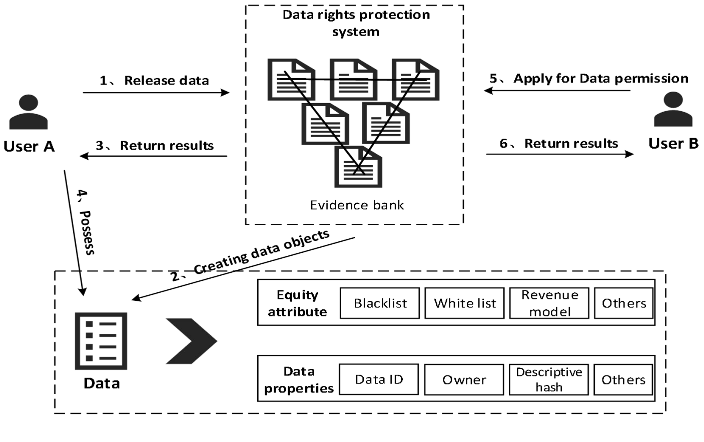


Figure 13 Data Rights Protection System Use Case Diagram

7. Security Mechanisms

CESS is a distributed storage system based on blockchain. In addition to building a stable storage infrastructure network, it also needs to maintain user data security. Therefore, CESS takes strict security measures to ensure the integrity and reliability of its stored data.

7.1 Data Security

In order to provide users with highly reliable data storage services, CESS protects user data security from three dimensions: data availability, data integrity and data privacy. Firstly, the client software strips the data, uses the erase code mechanism to encode the data, which enables the data sub-layer to have some error correction ability, and avoids storing the unusable elements of the data caused by hacker attacks, physical device failures, etc. Second, the data integrity verification mechanism and the space-time verification mechanism are used to ensure that the user data of the storage node is complete and available during the contract period. Third, for user data privacy issues, we use efficient encryption algorithms to encrypt user data to ensure that the data stored on the storage node is not clear text storage, effectively preventing data leakage.

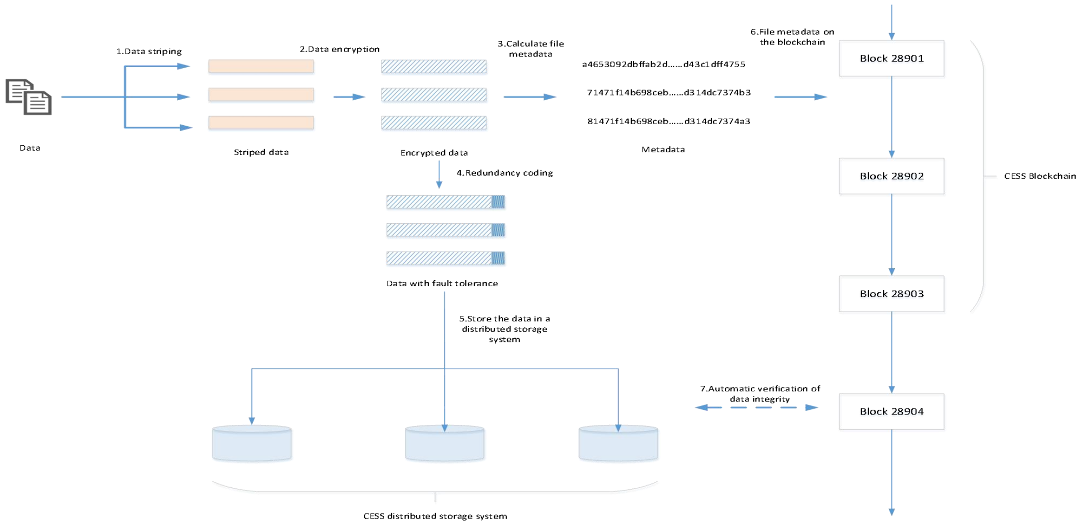

1.Data striping
2.Data encryption
3.Calculate file metadata
6.File metadata on the blockchain
a4653092dbffab2d......d43c1dff4755
71471f14b698ceb......d314dc7374b3
81471f14b698ceb......d314dc7374a3
Data
Striped data
Encrypted data
Metadata
Block 28901
Block 28902
Block 28903
Block 28904
CESS Blockchain
4.Redundancy coding
Data with fault tolerance
5.Store the data in a distributed storage system
7.Automatic verification of data integrity
CESS distributed storage system

Figure 14 Data storage process

7.2 Consensus Security

Blockchain, as a de-centralized distributed public database, are maintained jointly by distributed nodes using cryptographic protocols. Byzantine attacks are attacks in which an attacker controls a number of authorized nodes in a communications network and arbitrarily interferes with or destroys the network, thus preventing a consensus among the blockchain nodes. The CESS platform may maliciously purge the data in the chain during the process of controlling the data in the chain.

To this end, CESS platform builds a hybrid and efficient consensus module. The PBFT algorithm is an effective fault-tolerant consensus algorithm that can accommodate up to one-third of malicious nodes in the network. When there are f Byzantine nodes (malicious nodes) in the system, the entire network must have 3f+1 replica node to ensure that the entire network can make correct judgments. This effectively prevents malicious behavior on the chain.

7.3 Multi-copy Attacks

Any storage miner needs to promptly verify with the verification server that he or she does store the distributed data, but there may be a way to temporarily generate the data, provide proof and receive storage rewards. These attackers are rewarded with unfair storage. Furthermore, attackers provide storage data and users cannot access the original data correctly, which results in a storage system reliability error.

The replication proof algorithm of CESS generates a Merkle tree synchronously when multiple storage copies are created for a single data. When the storage miners need to prove that they have a copy of the data stored, they must provide the correct path to the Merkle tree to ensure that the miners cannot store copies of the data in their own way, thus resulting in unavailable data and no additional storage rewards for the miners.

8. Economic Model

8.1 Roles and Functions

CESS network operations require a variety of roles to participate in operational maintenance including storage-miner.

Storage miners: Includes content miners, content distribution miners, responsible for storage network construction, storage and mining, and participating in storage market transactions.

Consensus node: Responsible for validation of storage capacity model for storage miners, and produce block of CESS.

8.2 Storage Capacity Model

In order to ensure the reasonable profits of the storage minor and to truly achieve the fairness and benefits of the profits of the storage miners, the CESS network has built a complete storage capacity model evaluation system, and scores scientifically according to each miner's storage capacity model. The system calls this score the “credit value” and rewards it according to the interval where the credit value is located. Indicators that affect the size of credit values are as follows:

- Basic Operation Indicators

  1. Machine Configuration
  2. How long the machine is online
  3. Up and down bandwidth of the machine

- Contribution Indicators for Resources

  1. Size of storage file
  2. Node network contribution traffic size

These are the existing indicators to build the model. Subsequently, according to the operation of the project, indicators will be continuously enriched to improve the storage capacity model, and through artificial intelligence machine learning, self-adjustment and improvement.

8.3 Storage Award

1) Evaluation of Basic Situation Operating Indicators

Verify with the PoAs algorithm and rate the available storage service of the node.

2) Assessment of Resource Contribution Indicators

Storage file size: The system uses the PoRep and PoSt proof algorithms.

The contribution traffic of a node network is mainly the traffic between the nodes. The traffic proof algorithm (PoF) is used.

The storage nodes of the CESS network and the rewards of the storage miners are allocated reasonably through their contributions, and the contribution value R is determined by their cumulative storage certificates as follows:

$$R = \left(\alpha f(PoSt) + \beta f(PoRep) + rf(PoF) + df(PoAs)\right) * x$$

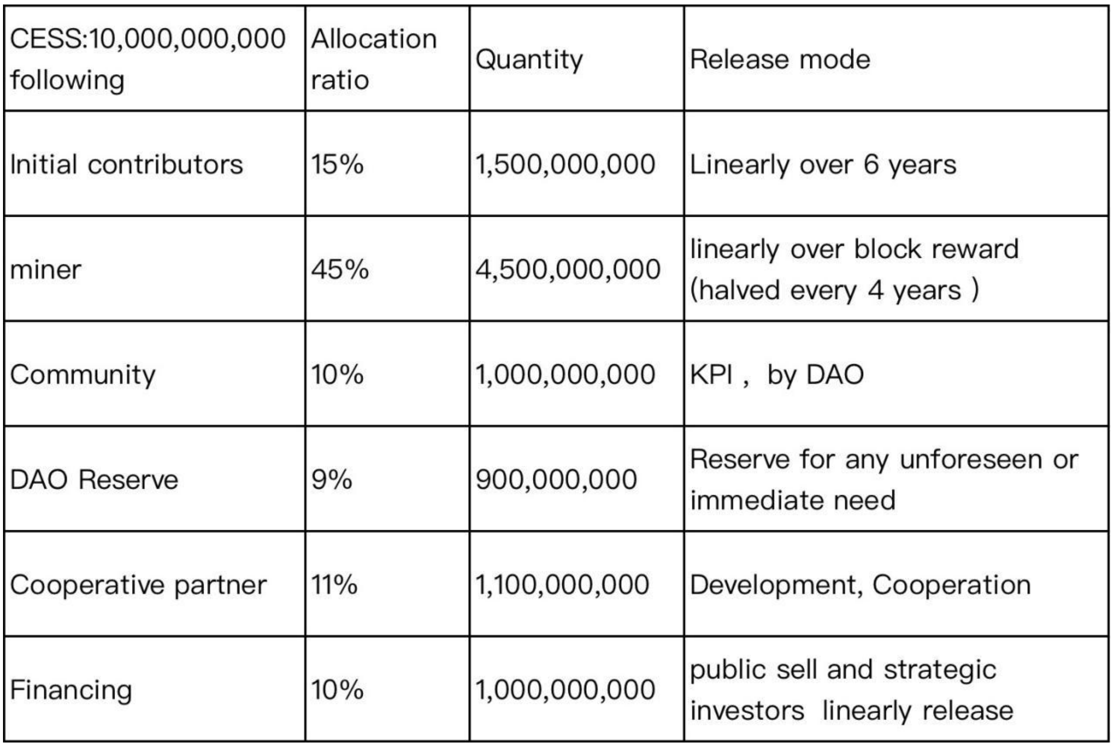
x is the mortgage parameter, if the cumulative contribution of the two miners is equal, the miners who serve longer in the future will be assigned a higher mortgage parameter and thus a higher mining reward. The default values for each proven scale factor are shown in Table 1 below:

Scale factor	describe	Default value
α	Spatio-temporal proof algorithm	40%
β	Replication proof algorithm	20%
r	Traffic proof algorithm	30%
d	Available Storage Service Certification Algorithm	10%

Table 1 Default scale factor values for each demonstration

During the early deployment of CESS, it is expected that the amount of data stored and storage transactions will be low. In order to encourage new miners or nodes to join the network and provide available storage services, the PoAs scale factor will be increased during this period. As storage increases, the scale factor for PoAs decreases, while the scale factor for other storage certificates increases.

8.4 Currency Allocation Model



CESS:10,000,000,000 following	Allocation ratio	Quantity	Release mode
Initial contributors	15%	1,500,000,000	Linearly over 6 years
miner	45%	4,500,000,000	linearly over block reward (halved every 4 years)
Community	10%	1,000,000,000	KPI ， by DAO
DAO Reserve	9%	900,000,000	Reserve for any unforeseen or immediate need
Cooperative partner	11%	1,100,000,000	Development, Cooperation
Financing	10%	1,000,000,000	public sell and strategic investors linearly release

Table 2 Default tokens allocation mechanism

8.5 Benefit model for each role

At the first stage of CESS network developing, there are two roles:

Storage of mining: Storage of mining proceeds + Pledge of mining proceeds + Payment of transaction fees to user accounts.

Consensus miners: Verify node commissions, generate block, obtaining block rewards for deployment.

9. Decentralized transactions and storage mining

9.1 Storage Markets: Verifiable and Trusted Markets

From commercial storage, there is an industry chain of "storage supplier to application to household". CESS will improve this industry chain and create an open trading market for clients and consumers. In the CESS economy, the storage market is a verifiable and trusted trading market where customers (buyers) can purchase low-cost storage space directly from the storage miners (sellers) to store data.

The system designs a storage market agreement based on the following requirements:

- Order Chain Up

Order price is open and transparent. Customers can make their own order according to the situation of the whole web order, and submit the order to the chain. Only the successful order can be accepted by the network, and cannot be modified after the chain is successful.

- Contributors invest resources

In order to maintain the stability of the storage market and prevent the storage miners from not providing services or providing services over time, the storage miners must mortgage a certain proportion of the amount of storage in the storage verification pool. Customers will pay the order fees in the verification pool first. Only after the verification is passed, the order fees will be entered into the storage miners' account. So as to ensure the smoothness and security of transactions.

- Self-organizing processing

Storage miners need to submit orders multiple times within the agreed time, to verify their storage activity to the verification miners, and confirming that the storage miners returned the results in an accurate and timely manner.

9.2 Storage mining: commercial implementation of de-centralized storage

CESS de-centralized storage is based on the CESS consensus. Storage Mines Data holding certificates for storage data must be submitted to the validated miners before they can be rewarded for validation.

The commercial implementation of de-centralized storage requires the mining to store valid data not valid random data, the storage miners need to be qualified and pledge a portion of the CESS award as order compensation. As the order duration decreases, the storage miners need to accept storage orders in the storage transaction market for storage requirements. After the system has validated the validation, the validation is successful. Storage miners receive storage mining rewards and fees paid by each transaction.

9.3 Storage Asset Markers: Improving Resource Integration in the Economy

Decentralization is not equivalent to de-centralization, especially in scenarios such as enterprise-level resource allocation, where the matching of individual storage miners and storage demand households in the trading market is obviously inefficient and uneconomic. The emergence of storage asset marketers solves the problem of efficiency. They can provide storage resources centrally through the scale of marketers and trade them according to the storage demand. The existence of marketers will greatly improve the integration of resources in the economy and the efficiency of the industrial chain.

10. Future Prospects

With the rapid development of distributed storage technology, hot spots have emerged. The development of cloud storage industry and blockchain industry always has a variety of new growth points and epochal revolution. Just as the Apple phone redefines the phone, CESS redefines the blockchain storage system.

Don Tapscott, the father of the digital economy and a world-renowned New economist, once said, "In the coming decades, it may not be social media, nor big data, nor robotics science that will have a huge impact. If we connect not only information but also value, so that large, globally distributed accounts can run on tens of millions of computers, and everyone has access, we can store, move, trade and manage any asset, from money to music, without going through a strong middleman. The technology that enables these implementations is the technological basis of digital currencies such as bitcoin and other crypto currencies. Blockchain might not be a current buzz-word, but It will be the future of the Internet and support a bright future to every transaction, every society and every person."

Based on the Decentralized Cloud Storage Infrastructure of Blockchain, CESS is committed to promoting data and value interconnection, building a more open, fair and secure network environment, and promoting the development of Web3.0. CESS will redefine the storage system, which will have the most profound impact on the future of the interconnected digital world.

People expect their data to remain their data, because “my data = my assets!”

This is a core value of Cumulus Encrypted Storage System (CESS) and is most relevant to our pioneering journey towards Web3.0.

11. Reference

1.Bitcoin: A peer-to-peer electronic cash system. Satoshi Nakamoto, 2008.

2.Proof of space from stacked expanders. Ling Ren, 2016.

3.Practical byzantine fault tolerance, M Castro, B Liskov - OSDI, 1999.

4.Practical Byzantine fault tolerance and proactive recovery, Miguel Castro, Barbara Liskov, 2002.

5.Secure and efficient proof of storage with deduplication, Qingji Zheng,Shouhuai Xu, 2012.

6.Comparative analysis of blockchain consensus algorithms, LM Bach, B Mihaljevic, M Zagar, 2018.

7.From blockchain consensus back to Byzantine consensus, Vincent Gramoli, 2020.

8.Proof of luck: An efficient blockchain consensus protocol, Mitar Milutinovic, 2016.

9.An innovative IPFS-based storage model for blockchain, Qiuhong Zheng, 2017.

10.Ethereum: A secure decentralised generalised transaction ledger[J]. Wood G.，2014.

11.Data object store and server for a cloud storage environment, including data deduplication and data management across multiple cloud storage sites, 2012.

12.Cassandra: a decentralized structured storage system, Lakshman A, 2010.

13.Enabling public auditability and data dynamics for storage security in cloud computing, Wang Q, 2010.

14.Blockchain and Distributed Ledgers - Mathematics, Technology, and Economics, Alexander Lipton, 2021.

15.Blockchain and Smart Contracts - Design Thinking and Programming for FinTech, Swee-Won Lo, 2021.

16.Protecting personal sensitive data security in the cloud with blockchain, Zhen Yang, 2021.

17.A Survey on Blockchain Technology: Evolution, Architecture and Security, Muhammad Nasir Mumtaz Bhutta, 2021.