Cumulus Encrypted Storage System (CESS): An Infrastructure of Decentralized Cloud Data Network

CESS(Cumulus Encrypted Storage System)

An Infrastructure of Decentralized Cloud Data Network

White Paper V0.9

CESS Lab

April 2024

Abstract

With vast volumes of data being stored in the cloud there are concerns over centralization, the fundamental issue around ensuring the ownership and privacy of the data and the majority of the data remains dormant for the most part. CESS aims to mitigate these issues by introducing numerous decentralized protocols and infrastructures. The key elements of CESS are, the Decentralized Cloud Data Network Infrastructure, Decentralized Cloud Storage System, Decentralized Online Data Sharing Platform, and Cumulus Gap.

The Decentralized Cloud Data Network Infrastructure is a peer-to-peer network, taking advantage of the blockchain framework and eliminating the need for intermediary governance. The system capitalizes on the token system and network's idle resources to enhance data protection and privileges.

The Decentralized Cloud Storage System comprises an account system, smart contract and a trusted decentralized cloud data network to establish an extensive distributed storage network.

Supports a variety of transactions and consensus mechanisms, and creates solutions for development of storage ecosystems that is well suited for commercial cloud storage.

The Decentralized Online Data Sharing Platform is designed for developers, creators, and consumers where everyone can build and evolve together. The platform accommodates a diverse range of content, including literary works, paintings, music, videos, and media. The involved community creates a token economy, establishing an innovative business model surrounding CESS which has inexhaustible scalability in future applications, encompassing data privacy, security, stability, rights confirmation, and rights protection.

Cumulus Gap, a Byzantine-Resistant Circuit with Privacy and Data Sovereignty from the CESS network. It allows participants in each CESS node to collaboratively train a shared model without sharing their own original data. Cumulus Gap will utilize a smart contract to delegate the task of local models to the computing nodes (it can be GPUs, GPU clusters, or even another Decentralized GPU Computation Web3 DePINs) in the CESS network.

1. Project Overview

As technologies like big data and machine learning are rapidly progressing, we are uncovering the value of digital assets, often referred to as the "Digital Gold." The substantial increase of data in the cyberspace demands innovative solutions for secure data storage and efficient sharing. The challenges we face involve ensuring secure storage, facilitating efficient sharing, and enabling trading while safeguarding the rights of data owners. The current solutions are complex and raise concerns.

1.1 Decentralized Distributed Cloud Storage

Cumulus Encrypted Storage System (CESS) is dedicated to develop a new global decentralized cloud storage online data sharing platform - a network infrastructure that is transparent, efficient, and provides equal opportunity to the global community. The CESS data sharing protocol enables:

a) data interoperability in the manner of cross-platform, cross-collaboration, and cross-format,

b) tracing and monitoring data trading market, and

c) fair and transparent data profit rewards for network participation.

CESS will adopt a phased approach to implement the above goals.

CESS protocol utilizes the idle resources provided by the participating nodes and Incentivize these nodes via the token economy. The nodes contribute data storage, computational resources , and network bandwidth. These resources are organized and managed by CESS protocol, exposing public interface to allow clients to access cloud data storage services

securely, and efficiently. The protocol further enables interconnection of network nodes, to build a large decentralized cloud storage system that supports 100's of Petabytes of storage and scale on demand.

1.2 Decentralized Cloud Data Network Infrastructure

In the present-day digital landscape, super platforms wield significant dominance over the internet. In 2016, a book named "Platform Revolution, How Networked Markets are Transforming the Economy" by American scholars Sangeet Paul Choudary, Marshall.W.Van Alstyne, and Geoffrey G.Parker, indicated that the essence of Web 2.0 is platform economy, and how platforms are reshaping the world. By obtaining and controlling the data from both service providers and consumers, these Web 2.0 giants gain the majority of market shares and profits. The data has become the core of these Web 2.0 enterprises, central to their profits. However, consumers are excluded from participating in the digital economy.

To empower consumers with control over their data, a global economy backed by decentralized, open and, transparent network world is necessary. The participants of the blockchain will be Incentivized for the resources they provide to the network. The decentralized distributed storage offers the advantages of data security and data rights protection, and will establish a robust groundwork for a forthcoming business model characterized by data-driven approaches and a bottom-up structure.

1.3 Decentralized Cloud Storage and Data Sharing Platform

We propose a **Multi-format Data Rights Confirmation Mechanism (MDRC),** which provides data owners with data rights protection and is capable of processing numerous data types. The digital assets will be uploaded, shared and traded within a protected market and the values of the digital assets will continually reveal new possibilities and insights.

2. Blockchain Based Cloud Storage Solution

A blockchain based decentralized cloud storage offers more security, integrity and scalability than traditional centralized storage networks. In CESS, all user data files are encrypted, replicated and sharded to ensure security and redundancy, and users are given unique private keys to access their private data. In addition, storage nodes only store segments of data files, greatly protecting networks from data leakages.

Storage nodes are incentivized to contribute their unused storage and bandwidth to the network. Clients pay to store or retrieve shared data. All user transactions are recorded and

secured by CESS blockchain, and the integrity of stored data is guaranteed by CESS storage proof schemes.

2.1 Incentive Model

The purpose of designing a CESS network incentive model is to encourage miners to provide honest and quality storage service, and therefore to maintain entire network stability. Our model mathematically quantifies contributions of network participating nodes, and fairly allocates rewards. It is also very important to make an overall CESS token allocation plan among initial contributors, miners, and CESS partners. Both miner incentive model and CESS token allocation plans are detailed in Chapter 7.

2.2 Consensus Mechanism

Various consensus algorithms exist today and the most common ones are Proof of Work (PoW) represented by BTC and ETH (ETH recently switched to PoS), Proof of Stake (PoS) represented by Cosmos, and Delegated Proof of Stake (DPoS) represented by EOS. While most algorithms are proven to be reliable and robust, significant challenges still remain, such as low Transaction Per Second (TPS) rate, expensive transaction fees, lack of average miner incentives, and security issues.

CESS proposes a block authoring protocol, namely the Random Rotational Selection(R^2S) consensus mechanism, designed to optimize and address the aforementioned issues. R^2S is complimented by a deterministic finality mechanism for blockchain, named GRANDPA, implemented in Substrate framework. Together, they form the consensus protocol. The goal is to provide a data sharing platform with a consensus mechanism that has low gas fees, high transaction throughput(10,000+ TPS), and fair incentives to all participating consensus miners. Section 4.2.4 describes our approach.

2.3 CESS Client-Platform Interactions

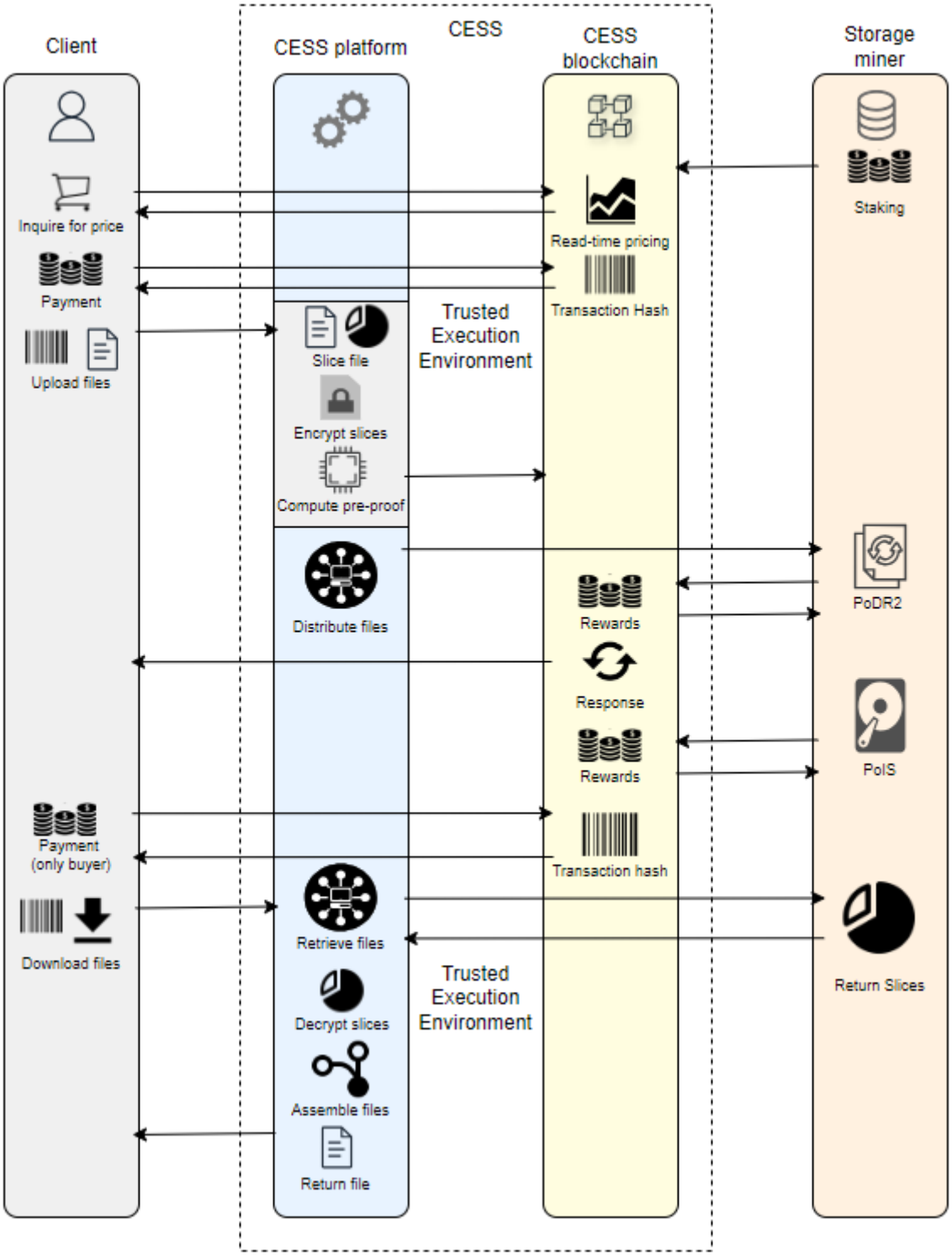

Figure 1. Client-Platform Interaction

A typical CESS client and platform interaction is shown above: first, a data storage client queries the CESS chain to get current storage price. The client then places an order for data files via on-chain smart contract. Once the payment is made and order is approved, the client then uploads the dat afile using the API. The data file is not directly uploaded to storage nodes, instead it is uploaded to a CESS storage scheduling node. The scheduling nodes are the ones with a secure hardware environment(Trusted Execution Environment or TEE) and where the data file will be processed, encrypted, and sharded(as described in section 2.3). Finally, the scheduling node distributes data segments to the storage nodes.

The storage nodes(storage miners) are Incentivized/slashed by the chain based on a set of protocols that enforces the storage nodes to provide the Proof of Idle Space and proof(PoIS) and Data Reduplication and Recovery ($PoDR^2$).

2.4 On-Chain Data Rights Confirmation

In the absence of robust enforcement of data rights protection, a market for data exchange runs the risk of becoming a prime target for cyber piracy, ultimately compromising its ability to deliver exclusive content to end users. With this in mind, CESS adopts an innovative on-chain Multi-format Data Rights Confirmation Mechanism (MDRC), generating a unique data certificate ID by extracting a data fingerprint from each data file. Subsequently, it employs a comparison of certificate IDs to identify and prevent potential violations of data rights, ensuring the safeguarding of high-quality content. Section 5.4 discusses the details of MDRC mechanism.

3. Application Scenarios

3.1 Distributed Network Drive

CESS offers Distributed Network Drive/disks service to end users. Compared to traditional Network Drive service providers, the network disk service has significant advantages in security, ownership protection, cost, and capacity.

The storage in CESS is not provided by cloud servers, effectively avoiding dependencies of centralized services. Instead, the data is stored in multiple independent storage nodes, no longer restricted by the centralized restriction of download/upload bandwidth. Blockchain-based cryptographic algorithms to encrypt data ensures the privacy of data, without data loss or central server outages. CESS also allows storage nodes to join the contribute their idle space dynamically, allowing the network storage to expand without limits.

3.2 NFT Storage and Trading Platform

Recently NFTs have attracted significant enthusiasm from artists, auction houses, art collectors, celebrities, as well as long term holders and societal elites and those who wish to be involved in this kind of investment and safeguard their portfolio against inflation. Decentralized and secured storage of NFT provenance and trading data underpins consumer confidence on the respective NFT trading platform.

NFT developers and owners only need to upload NFT Files and CESS will verify and confirm owners' data rights using the Multi-format Data Rights Confirmation Mechanism(MDRC), and then distribute the data files to storage nodes. Essential structural, subject, and semantic features are automatically reflected in the vector space of CESS. This process ensures accurate

indexing and mapping, ultimately enhancing both public exploration and secure private retrieval of NFTs within the system.

3.3 Distributed Enterprise Storage Service

The network of CESS storage nodes is built on blockchain technology and multiple storage proof mechanisms such as Proof of Data Reduplication and Recovery(PoDR^2) and Proof of Idle Space (PoIS). The system makes use of idle bandwidth and storage resources to provide powerful and more effective storage services than traditional cloud storage at a low cost. CESS brings the following changes to the data storage industry:-

3.3.1 Evolution of Data Production Chain:

CESS distributed storage can make the data production process non-linear and network oriented. That is, all production elements can be network configured. Data production becomes fully synergized, generating products in a manner of human-machine-human collaboration (so called "crowd-production").

3.3.2 Evolution of Data Flow Chain:

CESS makes it possible to eliminate traditional intermediary data transaction channels. This elimination enables all kinds of users to access data products anytime and anywhere. New data flow chain modes for data products and services, such as non-linear circulation. Traditional monopoly data channels and platforms will be replaced. Smart contracts between producers, products, and users make it possible for trustworthy dissemination between products and service points. As a result, transaction costs will be greatly reduced.

3.3.3 Evolution of Data Consumption Chain:

Instead of giant corporations or data centers dominating the data arena, the sovereignty of data is not shifting towards their rightful owners. The owners can now store, integrate, optimize, match and participate in managing their digital assets.

3.4 Data Rights Protection

From clients' point of view, CESS is a decentralized and user-managed data content sharing platform. Our mission is to give data ownership back to users, to encourage users to explore the values of their digital assets, and at the same time protect users' rights. With this in mind, CESS has implemented an on-chain smart contract based data sharing platform that is self-executable, fair and transparent. It also covers the entire life cycle of data rights confirmation, data rights tracking, and data rights protection.

CESS offers two types of smart contracts to users with different client-profit models. When users upload data files, they get to choose model values. CESS generates data file attributes based on

user inputs. The data attributes include client-profit model type, whitelist, blacklist, and so on. Data attributes are published together with user data. Whenever a data file is retrieved, by clients of the data provider, the underlying smart contract is executed according to the program set by the data file owner. Based on data file attributes, the system checks if data buyers have permission to retrieve the files. If permission checks are passed, the system will issue charges to buyers based on the relevant client-profit model, and then start data downloading.

CESS data users can also set their data file attributes by themselves. On the CESS platform, all data file retrieval records are recorded on blockchain and hence are backward traceable. The CESS data rights protection mechanism maintains a recording module to allow users to view their data file retrieval records, providing strong evidence for user data rights protection.

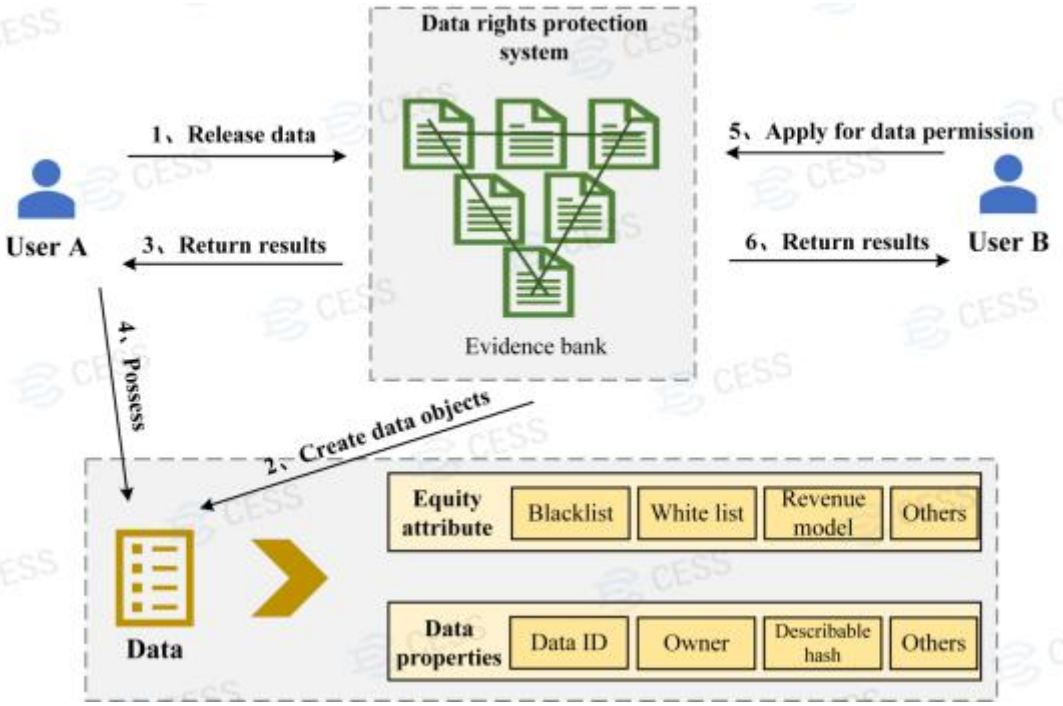

Figure 2. Data Right Confirmation and Protection

4. Technical Implementations

4.1 Overall System Architecture

CESS adopts layered and loosely coupled system architecture, which is divided into blockchain service layer, distributed storage resource layer, distributed content delivery layer and application layer.

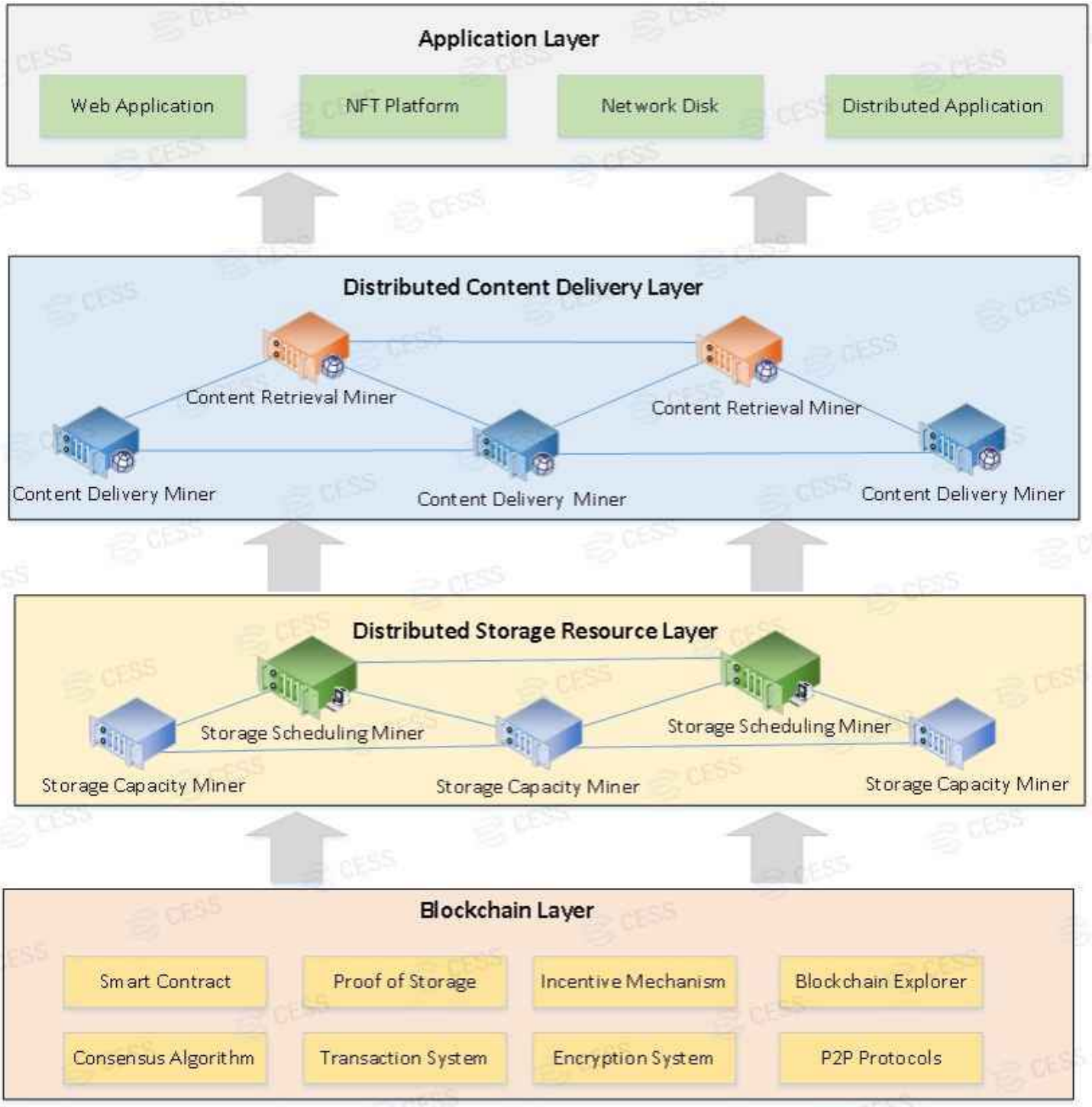

Figure 3. Layered System Architecture

The **Blockchain layer** provides blockchain service for the entire CESS network, including encouraging unused storage resources and computational resources to join the CESS network to provide data storage, data rights 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 delivery layer** uses content caching technology to achieve fast delivery of stored data, which is composed of data index miners and data delivery miners.

The **application layer** provides API/SDK tools to support data storage service, blockchain service, network drive service, enterprise level SDK, AI applications, etc.

4.2 Blockchain Layer

The blockchain layer is further divided into six layers: infrastructure layer, data layer, network layer, consensus layer, incentive layer and application layer. The infrastructure layer consists of hardware equipment including servers, network hardware and storage hardware. The data layer supports scalable data storage and provides various data processing algorithms. The network layer enables communication, load balancing, and data transfer between nodes across the network. The consensus layer provides sets of protocols which work together to find consensus among the nodes. The incentive layer is to achieve fair income distribution through smart contracts. The application layer supports DApps or Apps developed by third-party developers.

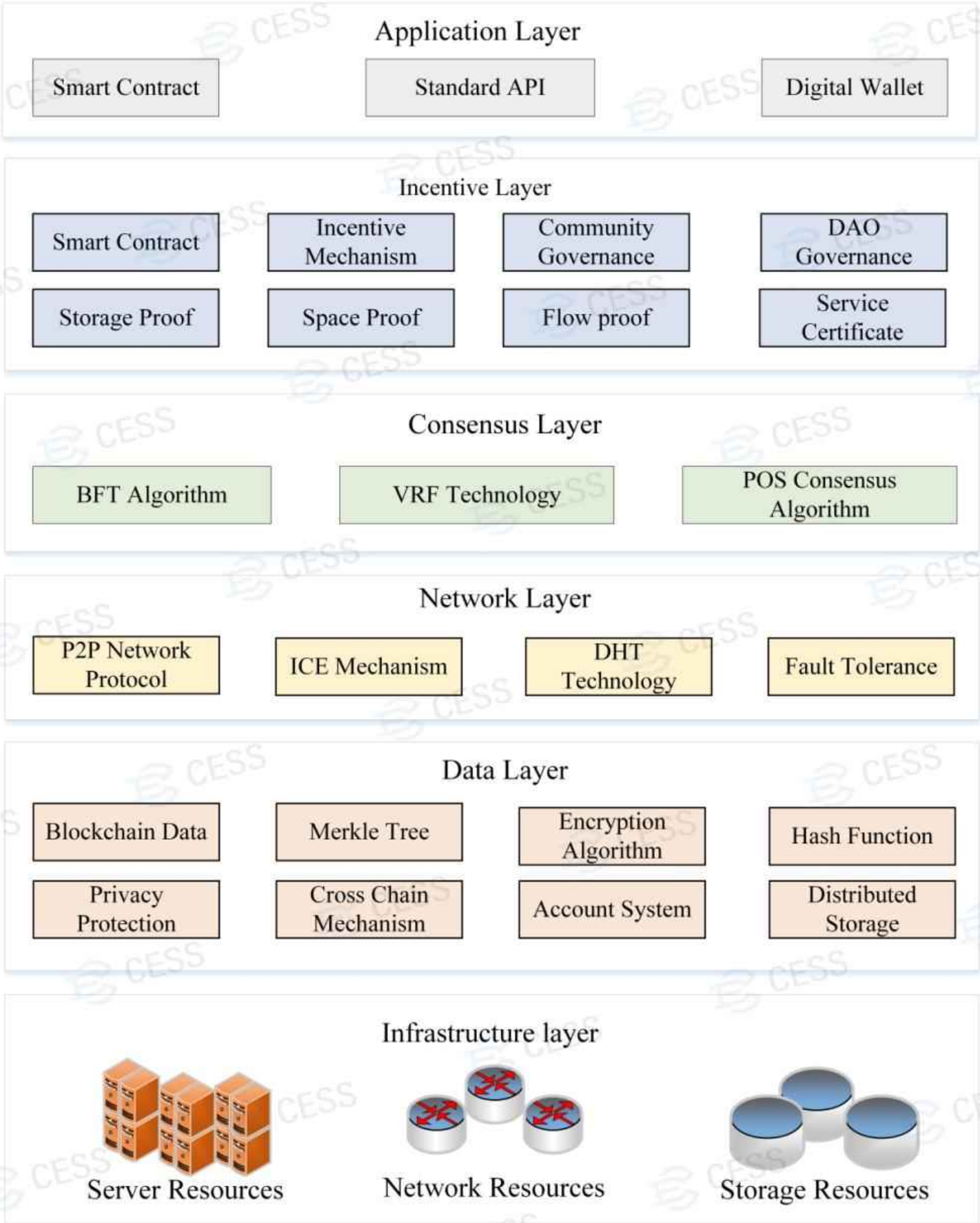

Figure 4. CESS Blockchain Architecture

4.2.1 Infrastructure Layer

As shown in Figure 4, CESS provides three types of global resources: Server Resources, Network Resources, and Storage Resources. Server resources focus on computing performance and will

carry out computation and task scheduling. Network resources provide network bandwidth and communication channels. Storage resources are the kernel part of the CESS system, providing stable and reliable storage infrastructure.

4.2.2 Data Layer

The data layer stores blockchain data, as well as client data files. To ensure the security and integrity of user data during data transmission, storage, and verification, various industry ready tools are used, such as digital signatures, hashing algorithms, merkle tree, etc.

4.2.3 Network Layer

CESS uses a peer-to-peer (P2P) network based on Distributed Hash Table (DHT) system that provides lookup and storage schemes. Network of nodes communicating to each other using the P2P protocol. The data and task are distributed over the direct network connection over P2P between the nodes. The routing is recorded between the nodes, so as long as they are connected to any other node already within the DHT network, the client can find more nodes to connect to.

4.2.4 Consensus Layer

To ensure that transactions and activities on the blockchain network can reach consensus quickly, CESS uses Random Rotational Selection(R^2S) block authoring. R^2S combined with the GHOST-based Recursive ANcestor Deriving Prefix Agreement (GRANDPA) from substrate, forms a consensus mechanism through which each node agrees upon the state of the chain at a given timestamp.

4.2.5 Incentive Layer

In CESS, there are two types of mining nodes: Content Storage Node(CSN) and Content Delivery Node(CDN). The CSN node is responsible for file storage while the CDN node is responsible for file delivery. The miners are Incentivized based on how much storage, computation and bandwidth they provide to the network.

4.2.6 Application Layer

Application layer provides API's for third-party developers allowing them to build on top of the CESS network. The developers can build a variety of DApps or Apps.

4.3 Distributed Storage Resource Layer

The CESS storage system is designed to work closely with blockchain systems, allowing it to take advantage of blockchain's robust security and distributed nature. CESS infrastructure introduced two types of nodes: Distributed Content Buffer and Distributed Cloud Storage. The distributed content buffer network will deliver data to the nearest content buffer node according

to the geographic location to optimize the access of data. The distributed cloud storage network is designed to provide massive, reliable, and scalable cloud storage services.

4.3.1 Data Storage Process

The process of storing data to the CESS network by users will go through several stages, such as production, upload, processing, storage, delivery 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 delivery phase, over 10T of network bandwidth can be achieved through a content delivery network. In addition, CESS supports users to delete data online and the CESS blockchain will keep track of all data operations.

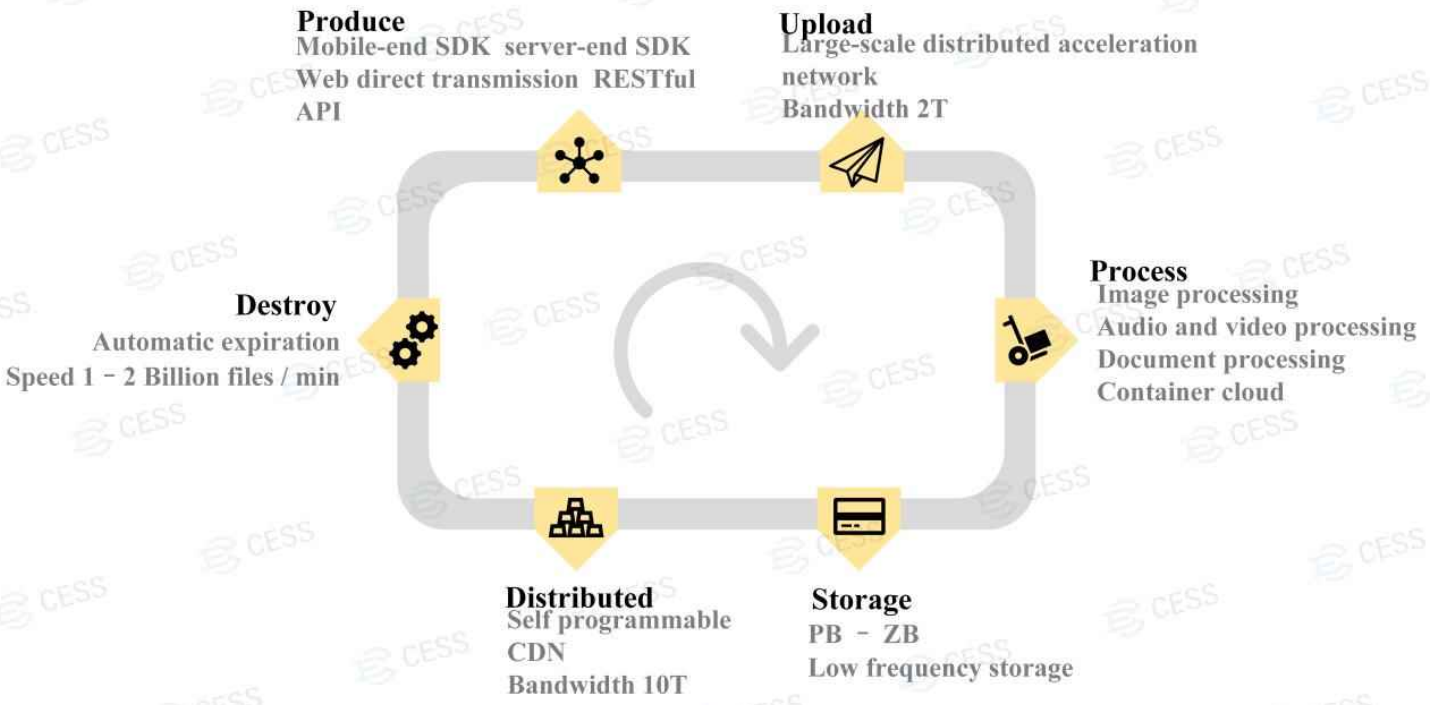

Figure 5. Data Storage Process

4.3.2 Distributed Content Delivery Network

In order to achieve efficient file access, the system effectively combines the advantages of both CDN and P2P technologies. By forming a content delivery network layer, it 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. It also 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 user downloads 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 delivery 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 Delivery Network Layer is perfectly combined with blockchain 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 6:

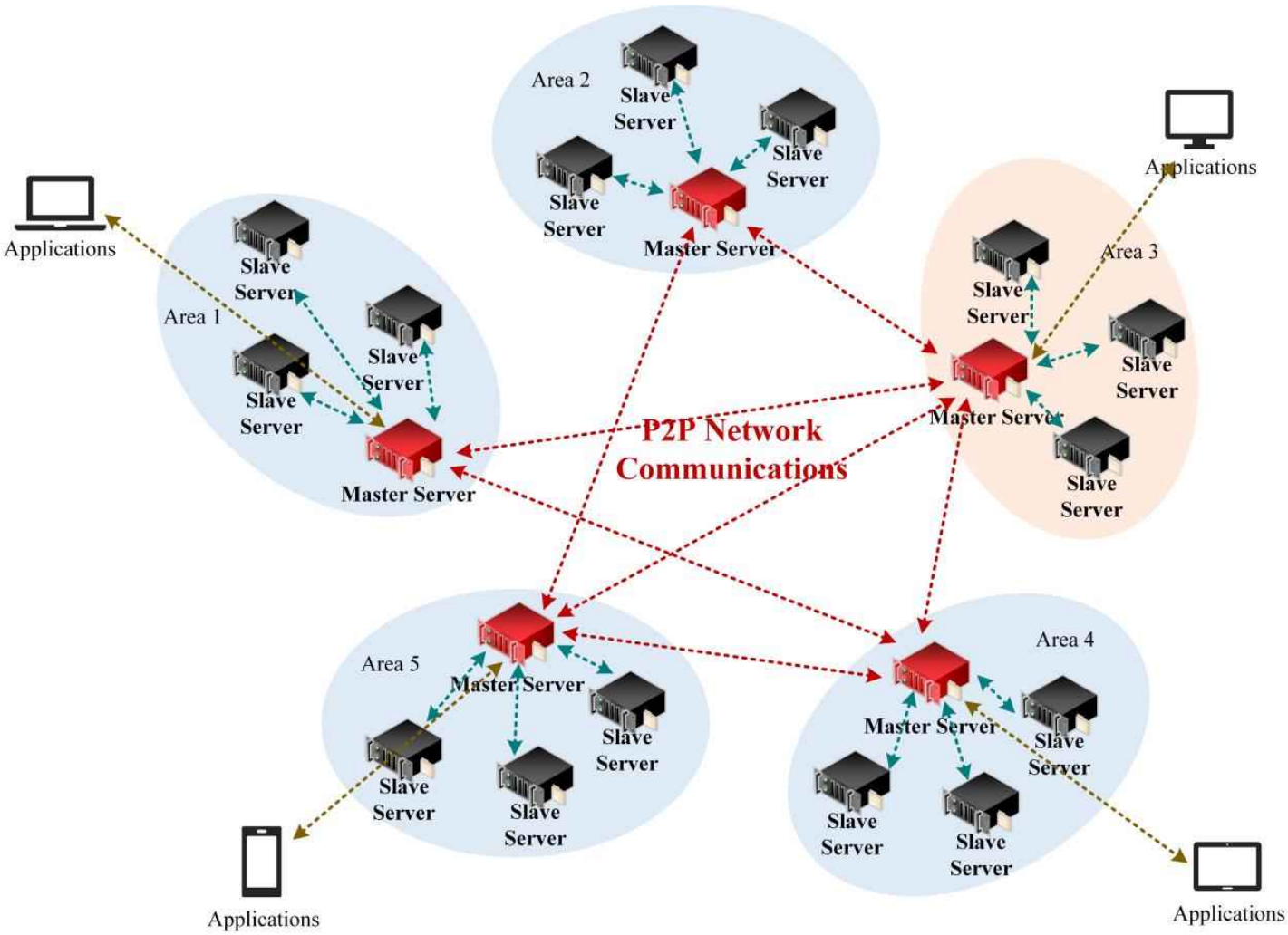

Figure 6. Distributed Content Cache Network

4.3.3 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 Figure 7 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.

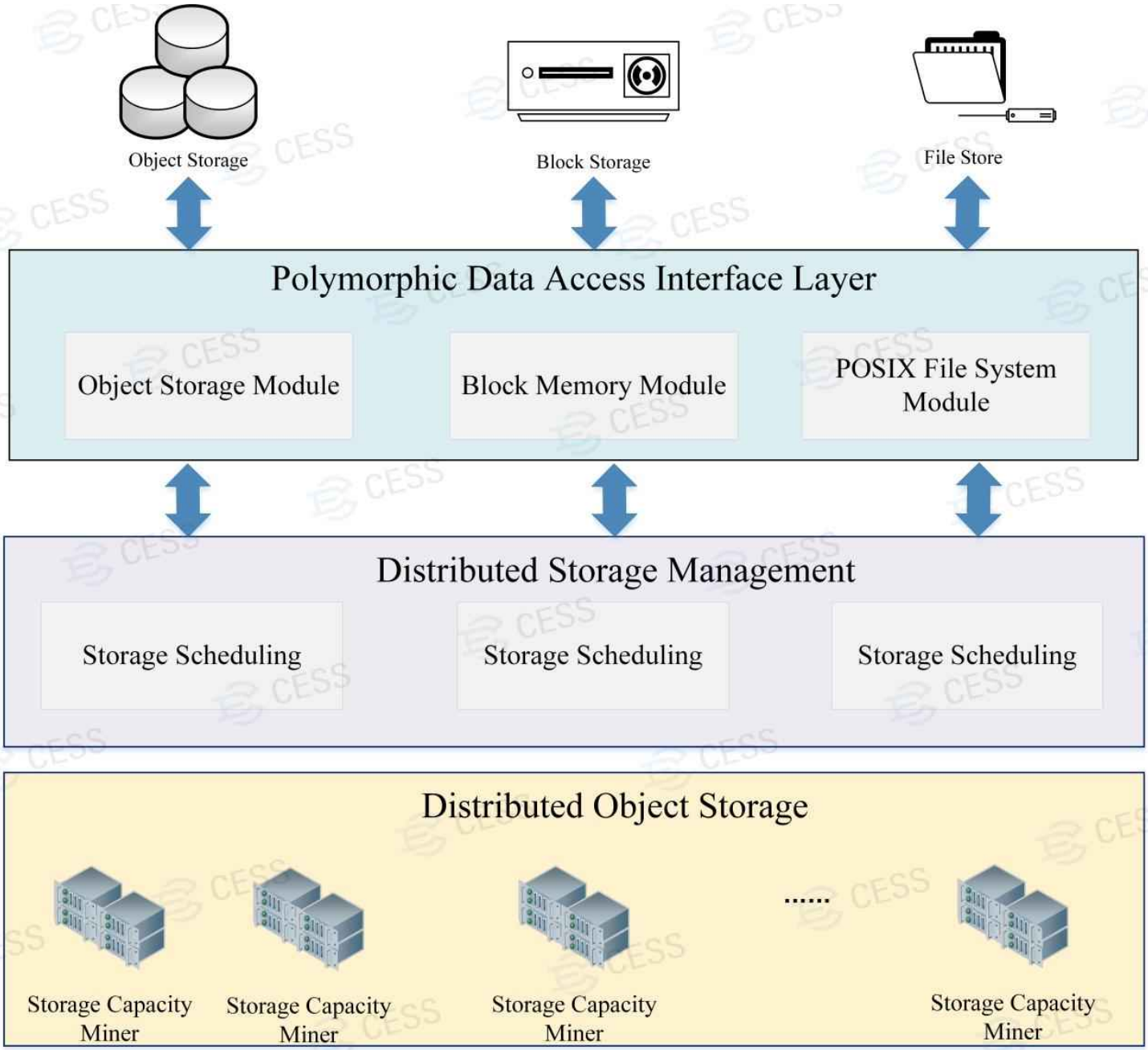

Figure 7. Distributed Storage Network Architecture Design

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 users' 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. Users' data in the POSIX file system is ultimately stored in the distributed object storage engine as object data.

5. Key Technologies

5.1 Random Rotational Selection (R^2S)

CESS adopts an innovative Random Rotational Selection (R^2S) mechanism to implement block packaging and other on-chain transactions. The R^2S mechanism allows all users who wish to become node operators to freely join candidate nodes, but within each time window (such as every 3600 blocks), only 11 formal rotation nodes are selected to participate in block production. Candidate nodes that have not participated in block production can also provide proof of their work ability by participating in data preprocessing and other processes, and thus participate in the next round of formal rotation node selection. During this process, the network will score the reputation of each node. When a node exhibits behavior that harms the overall interests of the network during its work, its score will be reduced. When the score is below a certain baseline, the node will not be able to participate in the competition for candidate nodes.

During the rotation process, if there are individual formal nodes that intentionally act maliciously or fail to meet network requirements, resulting in being forced offline, the network will randomly select nodes from the candidate nodes for supplementation until the completion of this round of time window. Therefore, for consensus node operators, even if they are only candidate nodes, they need to maintain continuous contributions to the network, which can greatly increase the possibility of obtaining benefits.

5.1.1 Verifiable Random Function

CESS adopts the model of "randomly selecting several consensus nodes from candidate consensus nodes, and then collaboratively packaging and trading blocks through consensus algorithms", which enhances the security of the blockchain while increasing the randomness and unpredictability of node election. Each candidate consensus node has a public and private key pair. When selecting the corresponding consensus node in each time window, each candidate consensus node calculates the hash random output through the following formula.

$$R = VRF\_Hash(Sk, Seed)$$
$$P = VRF\_Proof(Sk, Seed)$$

Among them, Sk it is the private key of the node, $Seed$ which is a field information in a block on the CESS chain and cannot be predicted in advance. R It is a hash random output, P and a hash

proof. Through the above steps, the verifier can easily verify that the two values are indeed generated by the node that owns the value.

$$R = VRF\_P2H(P)$$
$$VRF\_Verify(Pk, Seed, P)$$

Among them Pk is the public key of the verified node. Through the above algorithm, the 11 nodes with the smallest hash random output can be selected as the consensus nodes. If more than 11 nodes are selected, they will be screened according to the credibility score.

5.1.2 Admission and Exit

Although CESS does not set overly strict prerequisites for the entry of nodes, it still needs to meet the basic operating indicators and resource contribution indicators of network operation conditions, and pledge a certain amount of CESS tokens to participate and prevent nodes from doing evil. After the node completes the pledge of CESS tokens, it can participate in the above process. When the node wants to exit, the network will determine whether to fully refund the collateral tokens based on the node's score during operation. In theory, as long as the node works normally and there are no problems such as long-term disconnection or intentional evil, the network will fully refund the collateral tokens. This admission mechanism can prevent witch attacks and improve consensus security.

5.1.3 Election and Block Production Process

Compared with the Polkadot consensus mechanism, R^{2}S focuses more on the process of node election and block generation. The following is the overall process:

1. Nodes register as consensus nodes through staking, and the current staking amount is 3 million.
2. In each era, validators will rotate. The rotation rule is score ranking. The 11 nodes with the highest scores are selected as validators for the era.
3. The final score is determined by the reputation score and random score. The final score is equal to (reputation score * 80%) + (random score * 20%).
4. Please refer to the next section for credit score calculation, and the random score is determined by VRF.
5. The selected validators generate blocks in order.
6. The current way of confirming blocks is the same as GRANDPA.
7. The last epoch of each era begins the election of validators for the next era.

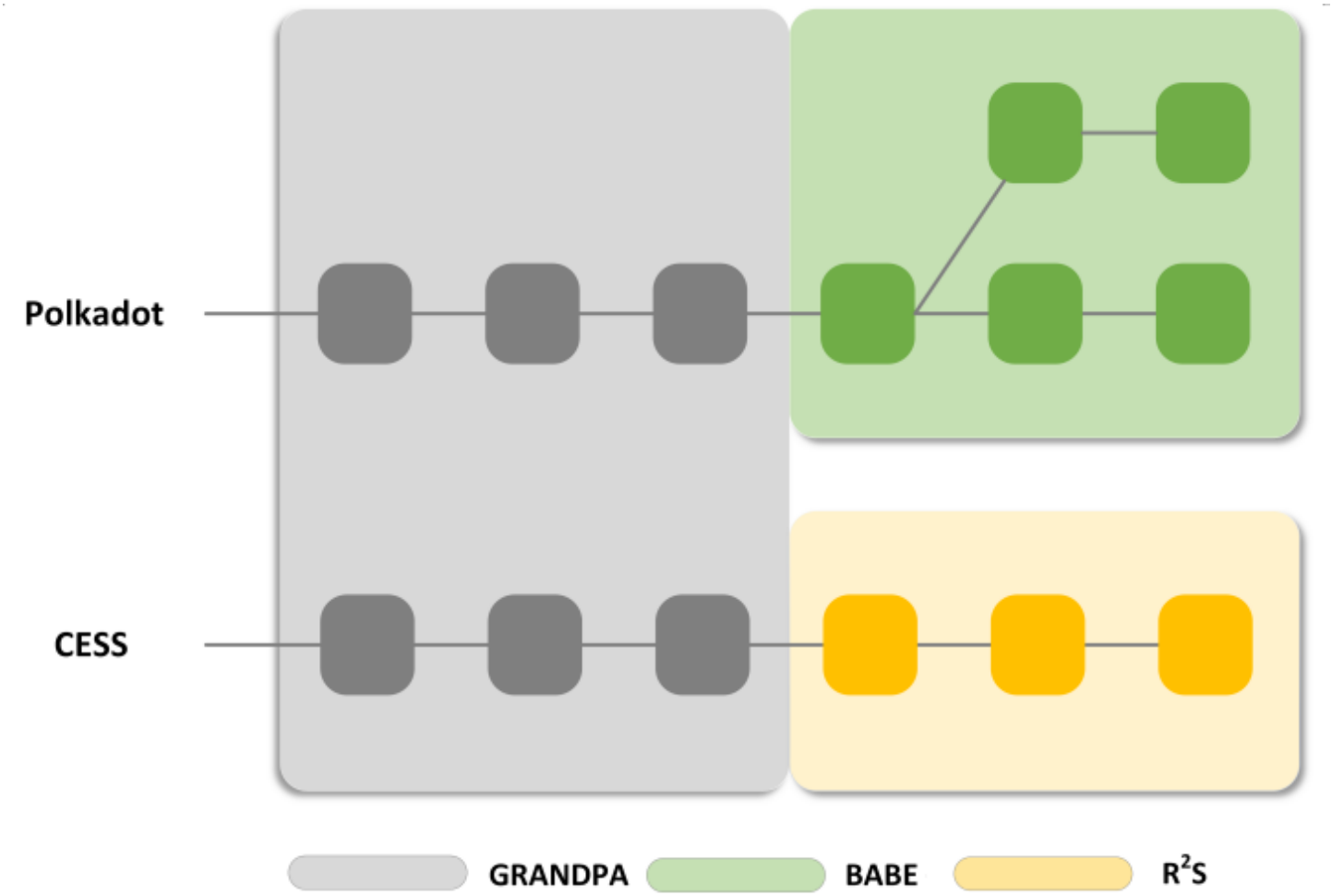

Figure 8. R²S Block Production

Consensus nodes joining the CESS network not only need to maintain network status, but also need to undertake tasks such as storing data preprocessing. In order to encourage consensus nodes to do more work, CESS adopts a reputation mechanism. The reputation score of each consensus node is directly determined by the workload of the validator, including the following items:

1. Total number of bytes for processing service files.
2. Verify the total number of bytes of service files and idle space during random challenges.
3. Total number of bytes to authenticate or replace idle space.

5.1.4 Advantages of R²S

Avoid Monopoly and Centralization

CESS achieves consensus and storage isolation through R²S, which ensures a more decentralized storage of block history and prevents excessive centralization of large miners, which is detrimental to the overall development of the network.

Improve Consensus Efficiency

CESS selects 11 nodes in each window through R²S to participate in block generation and verification, and the block generation process is carried out by these 11 rotating nodes in turn,

achieving efficient consensus while ensuring decentralization.

Implement On-Chain Transaction Processing

Due to the efficiency of on-chain transaction processing, CESS can achieve on-chain metadata, which can directly implement on-chain data storage addressing and other aspects, and ensure the authenticity of data through the blockchain mechanism.

5.2 Multiple Data Storage Proof Scheme

The landscape of decentralized storage networks, exemplified by projects like Storj and Filecoin, underscores the feasibility of leveraging blockchain technology to construct such networks. Users increasingly tend to contribute their idle storage resources to decentralized storage networks, seeking corresponding benefits. However, ensuring data integrity in such networks is a pressing concern, primarily due to potential cheating behaviors among participants.

Cheating behaviors, notably storage space fraud and outsourcing attacks, pose significant challenges. These behaviors involve miners providing falsified storage space or colluding to subvert data reliability by storing multiple copies of data on seemingly independent miners. Various mechanisms have been proposed to address these challenges, including proof of storage, proof of replication, and proof of space-time. While effective in theory and practice, certain mechanisms may encounter efficiency bottlenecks, particularly in data retrieval.

In response to this, CESS has introduced two innovative techniques to enhance its storage services: Proof of Idle Space (PoIS) and Proof of Data Reduplication and Recovery (PoDR^2). PoIS validates the storage space offered by the storage miners, which does not include the user's data; hence called idle space (aka. **idle segment**). On the other hand, PoDR^2 is used to verify the user's data (aka. **service segment**) stored by storage miners.

5.2.1 Proof of Idle Space (PoIS)

As there's no assurance that every node joining the storage network is trustworthy, CESS cannot directly read the idle space of nodes like ordinary computer disk management. One straightforward and efficient approach is to utilize specifically generated random data to occupy these vacant areas. The true idle space size of each node can be determined by the size of the filled data. In addition, some guarantee mechanisms similar to storage proof are needed to continuously save these random data in storage nodes to provide stable and available storage space. When storing user files, replacing the same large idle data can achieve the transformation of space from idle to service.

CESS introduces the Proof of Idle Space (PoIS) mechanism to implement processes such as authentication, verification, and replacement of idle space for storage nodes. Similar to the proof of service data storage mechanism, proof of idle space also needs to check the integrity of idle data through random challenges and verification processes. However, unlike the fact that

service data is provided by users, idle data (idle files) are generated by storage nodes in a certain way, which also makes the proof of idle space algorithm and storage proof algorithm have significant differences in implementation. Currently, there has been a lot of research on proof of space and it is widely used in various distributed storage systems or blockchain consensus protocols. Most existing space proof algorithms manage large or whole storage spaces, such as Filecoin's replication proof, Chia's spatiotemporal proof, etc. However, because the CESS system needs to meet dynamic operations such as user data insertion and deletion, the replacement of large space will be very slow. Therefore, the CESS idle space proof mechanism has made certain improvements to better adapt to the scene of dynamic changes in space.

5.2.1.1 Accumulator

An accumulator is a fixed-length byte sequence (or digest) obtained through a series of element "accumulation operations". It is commonly used to prove whether an element is in a set, but its unordered and flattened characteristics make it easier to dynamically add or delete elements. "Accumulation operation" refers to embedding elements from a set into an accumulator through certain cryptographic calculations.

When there are many elements, the speed of calculating the accumulator and element evidence will decrease significantly. The proof of idle space uses a three-layer multi-level accumulator to improve the calculation efficiency. In the multi-level accumulator, the upper accumulator element is composed of a group of sub-accumulators. When updating an element in the sub-accumulator, only its parent accumulator needs to be recalculated layer by layer, as well as the evidence of its sibling accumulator, without updating other elements. As shown in the figure 9, it is a two-level accumulator. When an element in sub-acc1 is updated, only sub-acc1 and ACC need to be recalculated, and then the evidence of sub-acc2... sub-accN needs to be updated, avoiding most element updates.

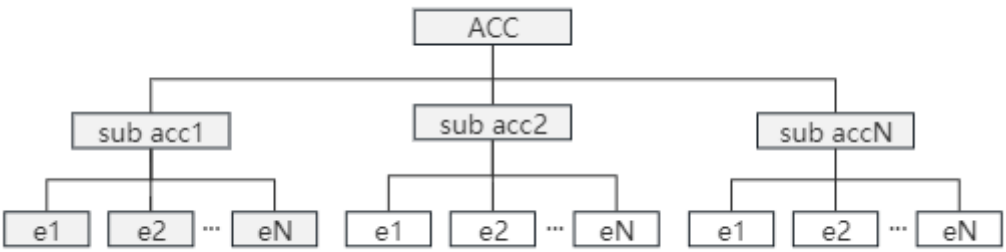

Figure 9. Accumulator and Sub-Accumulator

5.2.1.2 Idle File Generation

When idle files are generated by a prover, three conditions need to be met to ensure security.

1. Generated idle files cannot be compressed to prevent provers from authenticating more space at a lower cost.

2. Idle files cannot be temporarily generated during random challenges to prevent generation attacks and spatiotemporal attacks.
3. Unable to use one idle file to authenticate multiple spaces, and unable to use someone else's idle file to authenticate one's own space, to prevent witch attacks and external attacks. You can generate idle files that meet the above conditions by executing a stone-laying game on a stacked binary expander.

A stacked Bipartite Expander is composed of several sets of bipartite graphs stacked together. Bipartite graphs are a special type of Directed Acyclic Graph (DAG). The vertex set V of the graph can be divided into two non-intersecting subsets, and the vertices connected by each edge in the graph belong to these two subsets respectively, and the vertices in the two subsets are not adjacent. As shown in figure 10, there is an example of a stacked bipartite expander, with a total of K + 1 layers, N = 4 vertices per layer, and D = 2 degrees per vertex. Generally, vertices without predecessors are called source points (such as V0 layer), and nodes without successors are called sink points (such as Vk layer).

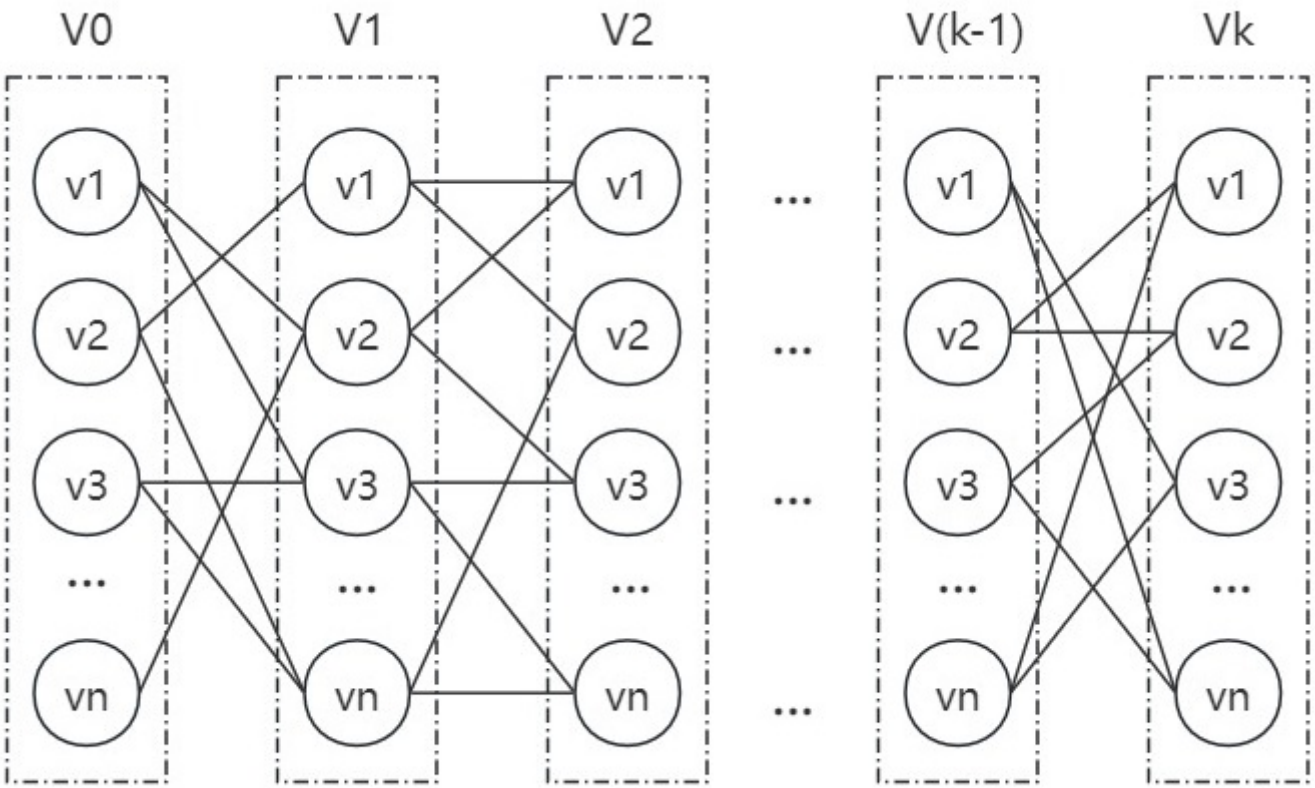

Figure 10. Idle File Generation

One of the key features of PoIS is its dynamic proof of space mechanism, which allows miners to continuously manage their stored space. This mechanism involves a two-layer accumulator structure, where the first layer accumulates all idle files, and the second layer proves the space of individual idle segments. Miners can add or delete idle segments as needed, and they must respond to challenges from verifiers to prove the integrity of their stored space.

5.2.2 Proof of Data Reduplication and Recovery (PoDR²)

CESS **PoDR²** guarantees data availability at all times. This is achieved using **Erasure Coding** (EC) where the data is broken into fragments, then they are expanded and encoded with redundant data pieces and stored across different storage miners. Erasure coding adds redundancy to the CESS network that can tolerate system failures. In addition, CESS PoDR² also implements **Proof of Data Possession** (PDP) to prevent the cheating behaviors discussed above.

When a user uploads a file to the CESS network, the PoDR² starts by slicing the file into multiple fragments, then fault-tolerant Erasure coding is calculated as depicted in the diagram below. The file fragments and Erasure encoded data are then distributed to randomly selected storage miners selected by the CESS network. Metadata of those fragments, such as the segment hash, location of the segment, size, etc., are recorded on the CESS blockchain.

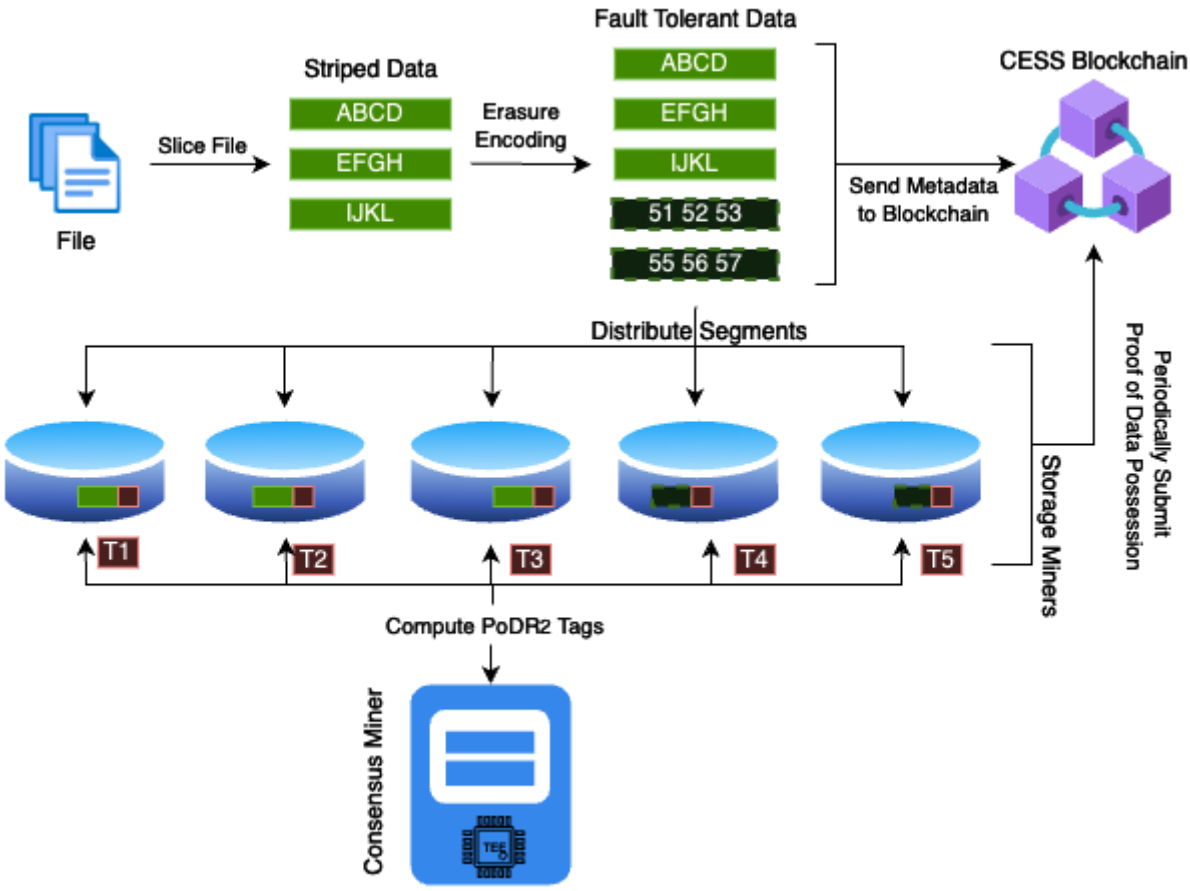

Figure 11. PoDR² Process

When a storage miner receives file fragments, they immediately request the Trusted Execution Environment (TEE) of the consensus miner to compute **PoDR² Tags**. This tag is then stored along with the file segment and used to compute PDP proofs. Once all the storage miners have successfully stored the file fragments, the CESS network periodically challenges the storage miners to compute proofs of randomly selected file fragments and submit them to the blockchain to earn rewards. If the miner loses the file segment or file tags, they will not be able

to compute proofs within the given time frame. This will result in a penalty for not being able to provide the required proof.

In case any file segment is lost either due to a natural hazard or a miner quitting the network, PoDR^2 can detect and recover the missing segment replicating it to a new storage miner, guaranteeing file availability.

5.2.2.1 Handling Large Files

Tagging large files can be a time and resource-consuming task, particularly in a TEE environment. Due to limited computing resources, generating tags for extensive files is not feasible. To address this issue, CESS PoDR^2 divides the data into smaller segments before fragmentation, enabling support for files of any size. This approach simplifies the process of computing Erasure coding, making it easier to handle large files.

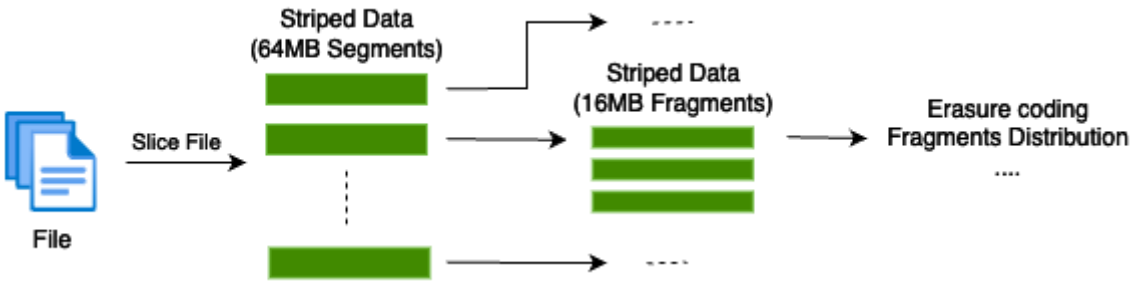

Figure 12. Slicing Files into Fragments

5.2.2.2 Trusted Execution Environment

In the CESS network, consensus miners are configured with a TEE, which is a secure area of a processor that helps code and data loaded inside it to be protected with respect to confidentiality and integrity. The data in the TEE is encrypted and isolated from other programs running on the system, preventing any third party from accessing the data loaded within it. CESS TEE consists of a private key pair that is unknown to external applications. This key is used to compute PoDR^2 Tags for each fragment of the data. The tag contains information like the fragment name and secret information encrypted by the PoDR^2 TEE's private key, which is based on PDP. In addition, computing tags are time and resource-intensive, making storage miners compute their fragment tags in advance. Failing to do so can result in storage miners not being able to produce storage proofs in time and being penalized.

5.3 Proxy Re-encryption

In order to ensure the security of user data, the data stored on CESS are encrypted and distributed to various storage miners. The core goal of CESS is to build a data equity platform around user data, and quickly achieve encrypted data circulation and sharing among different entities. In order to facilitate data sharing among different users on CESS, we will design and implement a decentralized proxy re-encryption mechanism, allowing data owners to exchange

data among data owners without leaking data content. Data uploaded by users to the CESS system will be marked as public and private. If marked as private, each segment of data will be encrypted and then sent to each storage miner for storage. If the data owner authorizes the data to others, the proxy re-encryption mechanism can be used to authorize and encrypt the nodes stored on the storage miner, allowing designated objects to decrypt with their own private keys to achieve secure access to others' data.

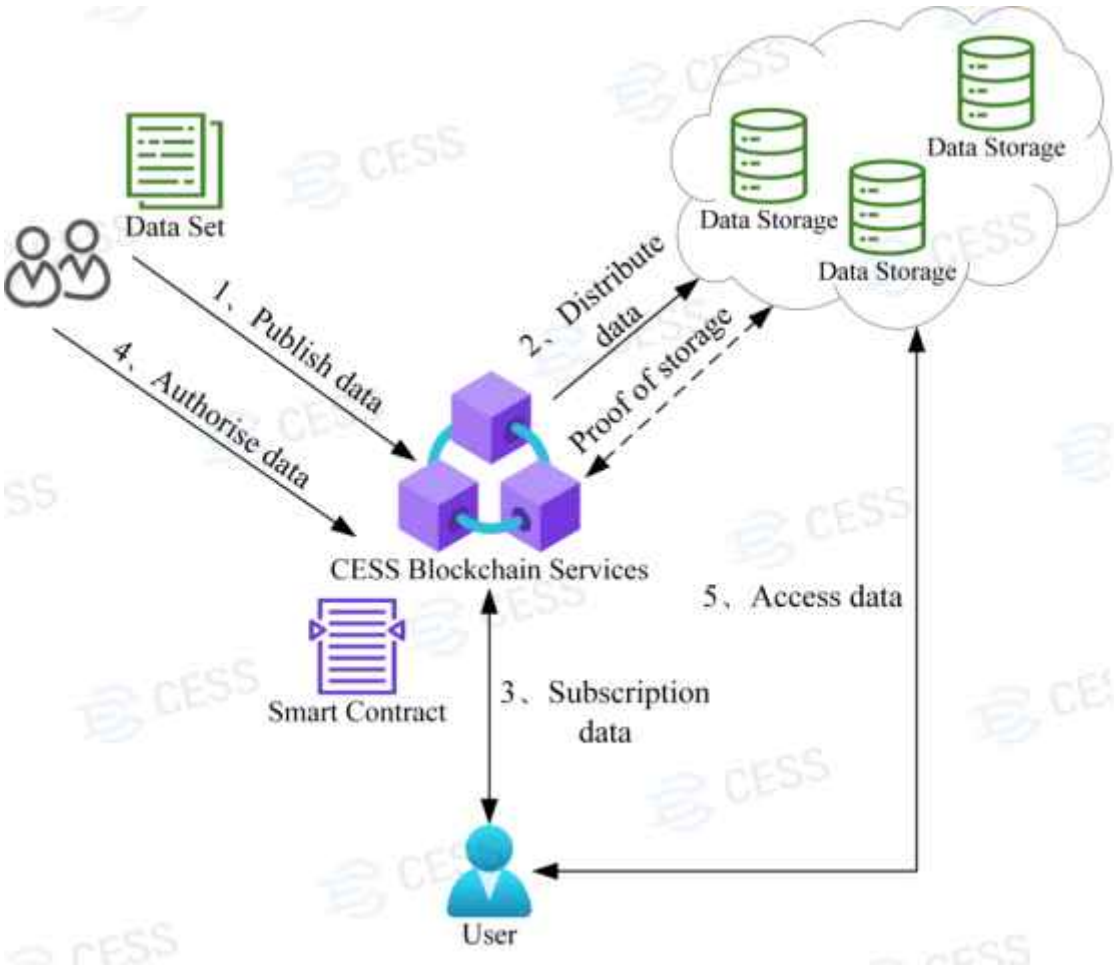

Figure 13. Proxy Re-encryption

5.4 Multiple-Format Data Rights Confirmation Mechanism (MDRC)

CESS is committed to creating a data trade market with copyright protection capabilities for original content creators. In any data circulation system, if valuable data does not receive sufficient copyright protection, it is likely to foster an environment conducive to piracy and plagiarism. Greatly damaging the quality of the market's data content. To address this, CESS has developed a multi-format data rights confirmation mechanism, which ensures that each digital fingerprint serves as copyright identifier for its respective data. By evaluating the resemblance of digital fingerprints, it determines whether the data has a data kinship relationship. When necessary, it can provide data source evidence for digital copyright protection and has certain data copyright protection capabilities. The overall process of MDRC is shown in the following figure:

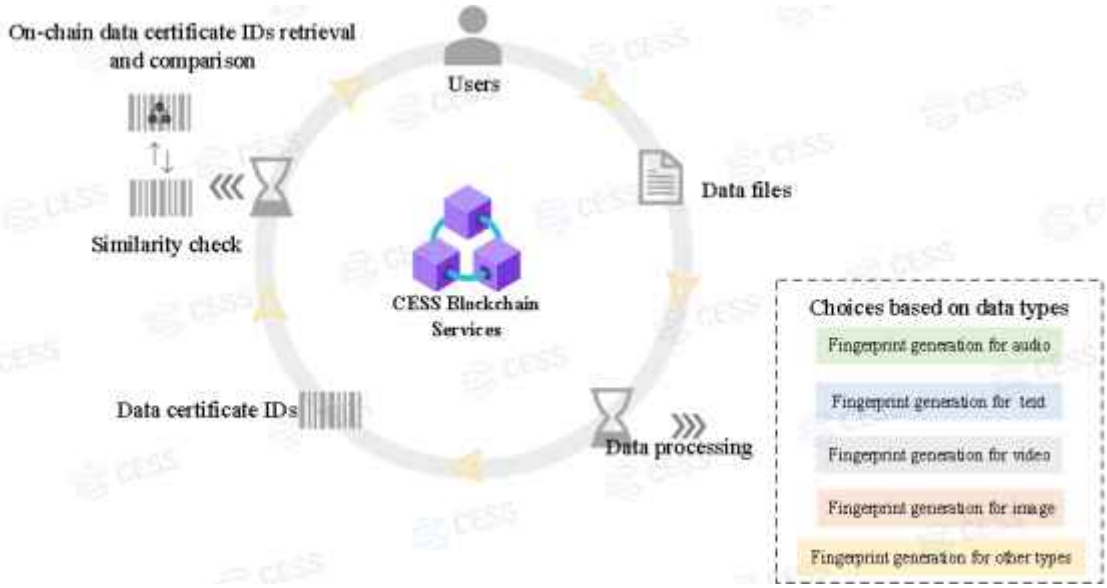

Figure 14. Multiple-Format Data Rights Confirmation Mechanism (MDRC)

The users initially process the data files through the digital fingerprint mechanism to obtain their data fingerprints. The CESS digital fingerprint mechanism operates at the user layer, preprocessing data before uploading it to the storage layer. It involves three main stages: fingerprint extraction, on-chain embedding, and comparison. The digital fingerprint algorithms depend on the data type of the data. For text data, the most advanced segmentation algorithm is selected for different types of natural language, and corresponding fingerprints are generated based on the latent semantic space; for image data, feature extraction such as color features, shape features, texture features, and spatial relationship features is supported, and algorithms are used for feature transformation to improve the accuracy of image digital fingerprints; audio data first samples and quantizes the signal, and then performs fast Fourier transform to achieve feature extraction, supporting energy features, time-domain features, frequency-domain features, music theory features, and perceptual features; feature extraction of video data is mainly achieved by extracting key frames, and then the key Frames are processed by feature extraction from images. In addition to the typical application data content mentioned above, CESS also supports digital fingerprint extraction methods for other data types. Usually, these data can be used as operation credentials for auditing. For example, system operation log data, event behavior trajectory, daily purchase vouchers, etc. For specific Technology Implementation, a simple MD5 value or SHA value will be used as the digital fingerprint of this type of data.

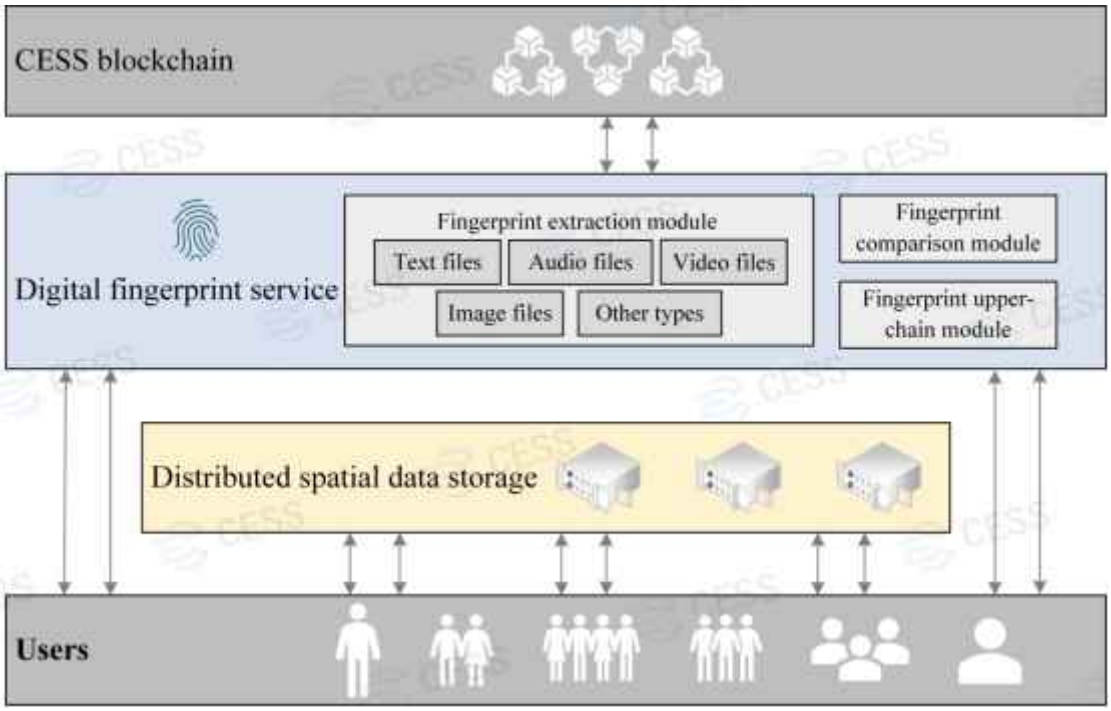

Figure 15.

After obtaining the data fingerprint, the users calculate the data fingerprint through simhash algorithm, perform probabilistic dimensionality reduction on the high-dimensional data and map it into a fingerprint with a small number of fixed digits, and generate a corresponding similarity hash. This similarity hash can be regarded as the copyright mark of the source data. Then, CESS utilizes hamming distance detection technology to compare the copyright identifier with existing identifiers on the blockchain, thereby conducting data lineage and similarity detection to provide users with reference information. Finally, upon uploading the source data to the CESS platform, the copyright identifier is stored on-chain through copyright identification on-chain storage services, facilitated by smart contracts, to provide necessary data support for subsequent data rights confirmation.

5.5 Cumulus Gap

Different organizations each have their own private data. Some(or all) of the private data cannot be shared with the outside world because of sensitivity and some are due to national legal issues. Because of this, valuable data cannot be used for generating AI modules.

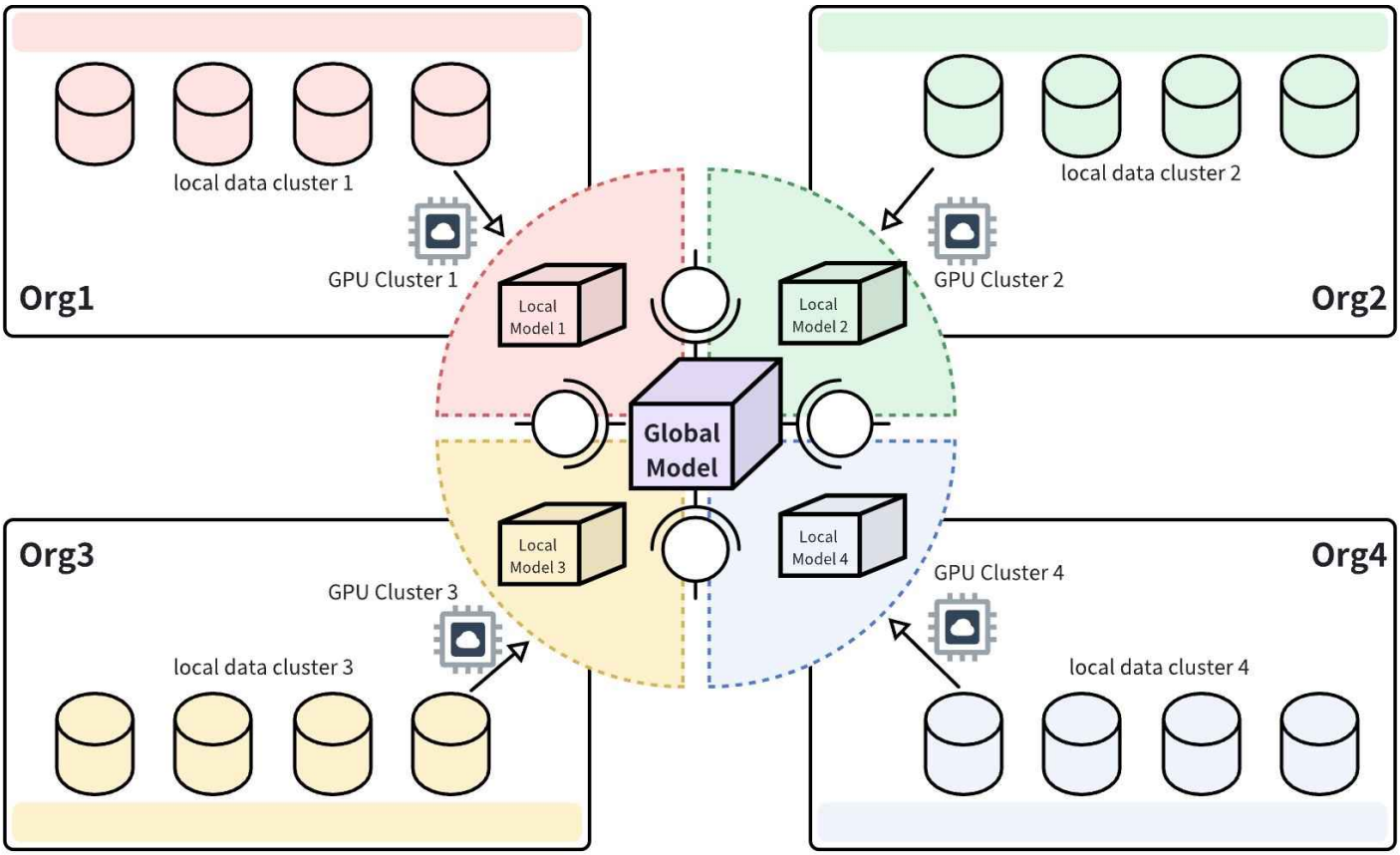

Figure 16. Cumulus Gap

These organizations can use the CESS network to exchange parameters and models, which can be iterated under the encryption mechanism of Cumulus Gap:a byzantine-robust circuit with data privacy. It schedule and coordinate the model training by smart contracts. That is, to establish a shared global model without against data privacy laws. This global model is the optimal model that everyone aggregates data to build. However, when building a global model, the data itself does not need to move outside, nor does it leak privacy or affect data compliance with regulations.

In this way, the generated models can serve not only local training goals, but they can be aggregated into a large model for use around industry-wide，or even around the whole world.

6. Security Mechanisms

CESS takes strict security measures to ensure the integrity and reliability of its stored data, and the security of transactions on blockchain.

6.1 Data Security

CESS guarantees user data security from three aspects: data availability, data integrity and data privacy. As described in previous chapters, all users' data files are stored as encrypted data segments; CESS has implemented a fault tolerant erasure coding method to protect data completeness against node failure and other malicious attacks. The CESS multiple proof schemes ensure the data integrity of each network node.

6.2 Consensus Security

Blockchain, as a de-centralized distributed public database, are maintained jointly by distributed nodes using cryptographic protocols. Byzantine attacks are ones in which an attacker controls a number of authorized nodes in a communication network and arbitrarily interferes with or destroys the network, thus preventing a consensus among the blockchain nodes. It might happen that the data on the CESS blockchain is purged during an attack.

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 a 3f+1 replica node to ensure that the entire network can make correct judgments. This effectively prevents malicious behavior on the chain.

6.3 Transaction Security

To regulate the use of smart contracts and to avoid the abuse of CESS storage and computing resources, CESS has designed a functional module to monitor and audit the blockchain transactions. As shown in the figure below, the CESS transaction audit module is a comprehensive data analysis function based on data on blockchain, private key management and transaction management. It provides a visualization of decentralized transaction execution, and transaction monitoring and auditing functionality.

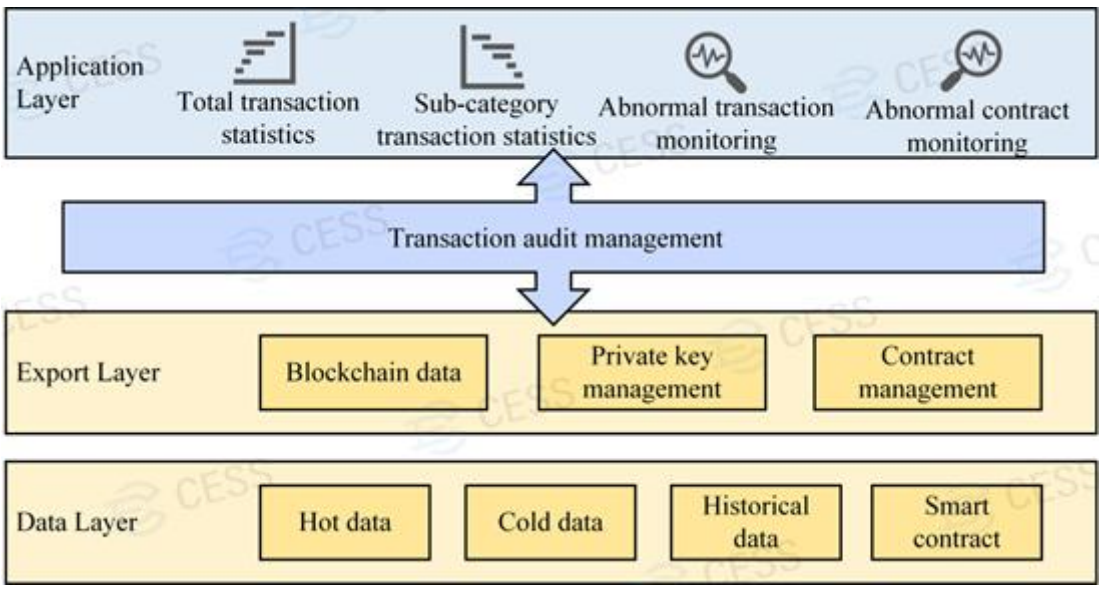

Figure 17. Transaction Security

The transaction audit items include:

1. Total number of user transactions. This is to monitor the daily number of transactions of each external account.
2. Total number of transactions for each transaction type. This is to monitor the daily number of transactions for each transaction type, each external account.

3. Abnormal transaction user accounts. Monitor abnormal user accounts that are not registered on blockchain middleware platform.
4. Abnormal transaction execution. Monitor on-chain transaction execution and transactions not on white-list (not registered on blockchain middleware platform).

7. Economic Model

7.1 Overview

The concept of CESS decentralized cloud storage aims to achieve the goal of allowing ordinary people to meet their data storage needs by renting affordable hard disk space in a peer-to-peer network. The operation of the CESS network requires the participation of multiple roles, such as storage nodes responsible for data storage and distribution and consensus nodes responsible for transaction records. Economic incentives need to solve the modeling of node storage computing power, quantify the network contribution of nodes, and give corresponding incentives according to the different contribution levels of nodes. This is the main link to decentralized storage.

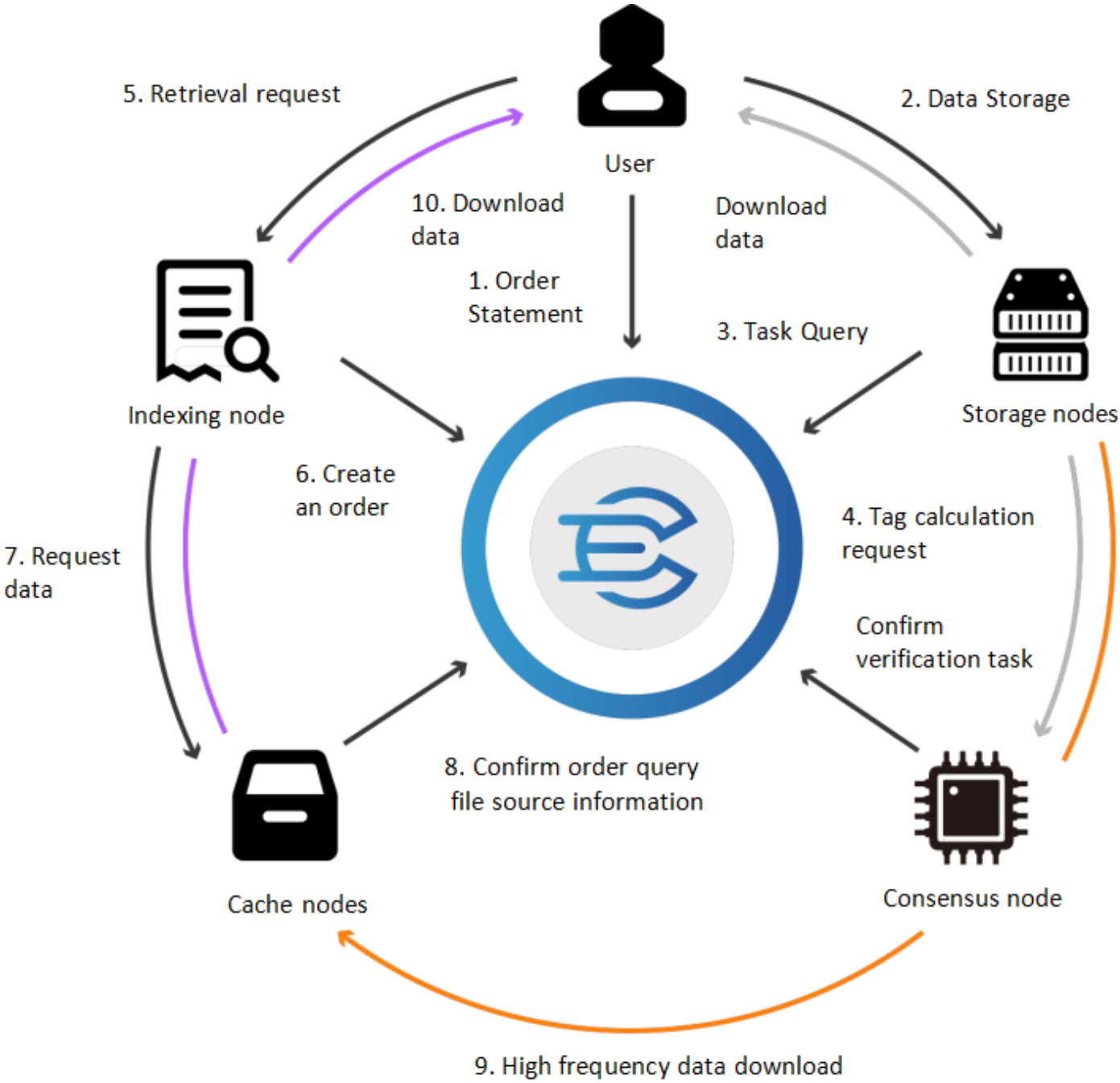

Figure 18. Economic Model

7.2 Roles and Functions

The operation of the CESS network necessitates the participation of various roles to support its upkeep and operation. During the initial stage of network creation, there are four primary roles: Storage-miner, Consensus-miner, Retrieval-miner, and Cache-miner.

7.2.1 Storage miners

During the construction process of the CESS network, content storage miners play a crucial role in providing data storage space for the application layer. The storage performance of these 35miners also directly impacts the overall network performance. To ensure the quality of storage services, a storage-proof mechanism is utilized to verify the storage capabilities of each

miner node. To incentivize storage miners, they are rewarded with mining incentives and a portion of the storage service fees.

7.2.2 Consensus Miners

In the CESS network, consensus miners are responsible for transaction packaging, transaction verification, blockchain data storage across the entire network, and supervising storage miners to verify their integrity. To incentivize consensus miners, they receive mining rewards. To motivate TEE-Workers, they also receive mining rewards and a portion of the storage service fees.

7.2.3 Retrieval Miners

Retrieval miners provide data retrieval services to the network by responding to "Get" requests and retrieving the requested data for users. After receiving a data retrieval request, retrieval miners search for the storage miner with the best overall performance to enhance the efficiency of data retrieval. Unlike storage miners, retrieval miners do not need to provide staking, submit stored data, or provide storage proofs. Unlike consensus miners, retrieval miners are not involved in transaction packaging or transaction verification.

7.2.4 Cache Miners

As the CESS network matures and attracts a large number of storage users, there will be a significant amount of data stored in the network, challenging the performance of data upload, retrieval, and download. To address this, the CESS network incorporates the role of cache miners. Cache miners assist in market transactions by efficiently indexing and distributing data, enabling fast delivery of data to users, consensus miners, and storage miners. Cache miners receive mining rewards as incentives for their role.

7.3 Token Allocation Model

CESS plans to publish a total of 10B tokens, of which 15% will be allocated to initial contributors, 10% to early investors, 10% for community development, incentives, and advertising expenses, 5% for cloud partner business cooperation, and 5% as a foundation reservation for emergency and support for future ecological development. The CESS storage network allocates up to 55% of the tokens as node incentives, of which 30% is allocated to storage nodes, 15% is allocated to consensus nodes, and 10% is used for the construction of the cache layer.

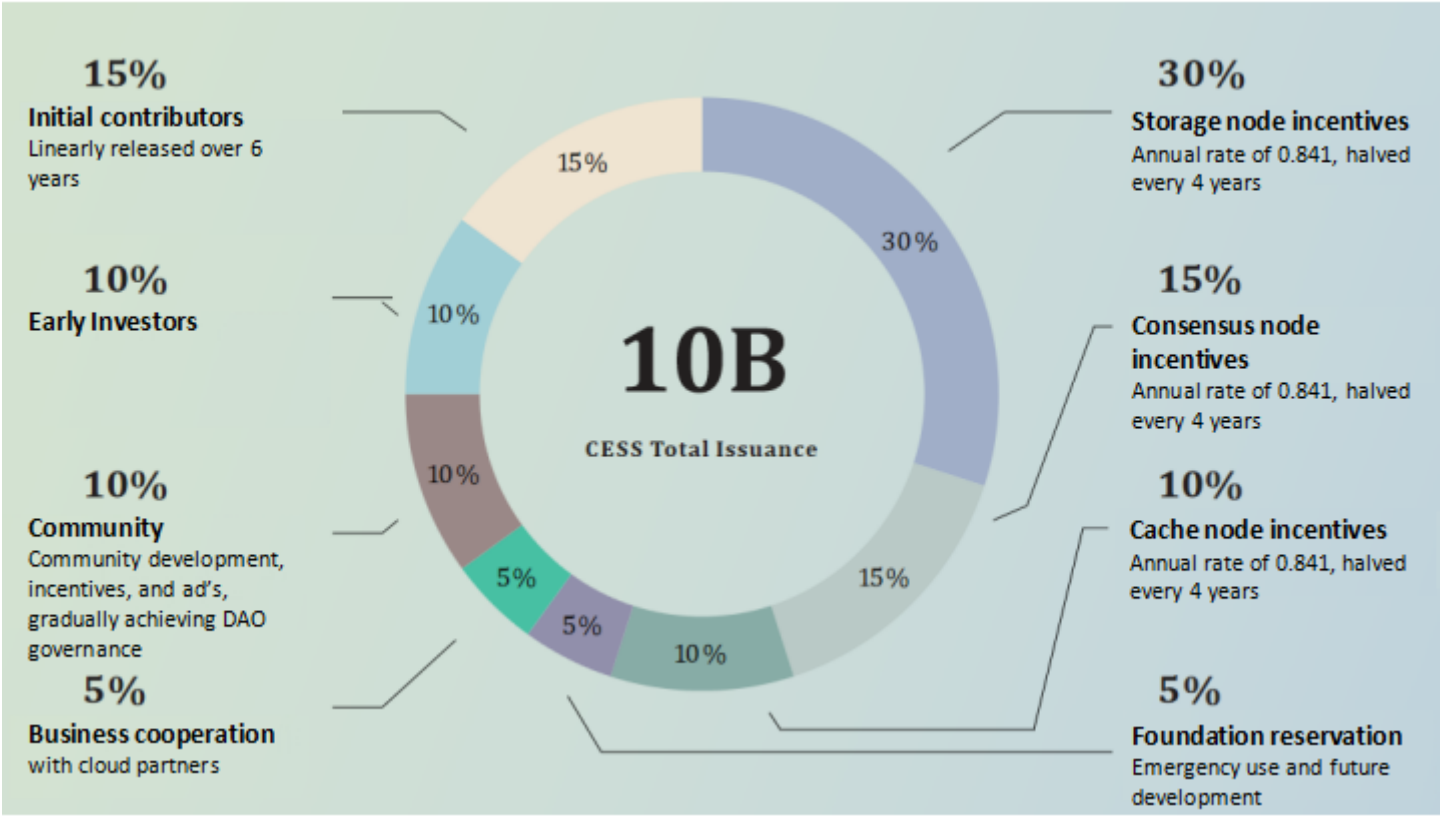

Figure 19. Token Allocation Model

7.4 Consensus Node Incentive

Consensus nodes use a randomly selected rotating consensus node mechanism (R^2S). Within each time period, 11 rotating consensus nodes are randomly selected from a large number of candidate consensus nodes that meet the requirements to provide block production services for the entire blockchain system. Participating consensus nodes will also receive block rewards for that round. Consensus nodes are responsible for packaging, verifying transactions, and storing blockchain data across the entire network. In order to motivate consensus nodes, they will receive mining rewards.

7.5 R^2S consensus algorithm

Each consensus node that joins the CESS network needs to maintain the network state and also undertake the audit work of storing data. In order to motivate consensus nodes to do more of this work, we have designed a reputation model. In the model, each consensus node has a reputation score, which is directly determined by the total workload of TEE Workers bound to the consensus node. Specifically, it includes the following items:

- Total number of bytes to process service files.
- The total number of bytes of authentication/replacement idle space.
- Verify the total number of bytes of service data and idle space in random challenges.

In each round of Era, validators are rotated based on their reputation scores. According to the Random Rotational Selection (R^2S) mechanism, the 11 nodes with the highest scores are

selected as validators for the Era.

The score of R^2S is composed of reputation score and random score, and the calculation process is as follows:

$$R^2S \text{ score} = (\text{Reputation score} * 80\%) + (\text{Random score VRF} * 20\%)$$

7.6 Node Incentive Mechanism

Consensus nodes join the CESS storage network by staking a certain amount of tokens and receive rewards based on the workload of participating in the consensus. For each Era, validators receive rewards in proportion to the Era points they collect. Era points are reward points obtained through payable actions, such as:

- Generate non-uncle block
- Generate a reference to a previously unreferenced uncle block
- Generate a referenced uncle block

(**Note** : Era is a collection of multiple epochs. After one era ends, the reward is settled. An era is about 6 hours.)

Uncle block is a relay chain block that is effective in all aspects, but failed to become the main block. This happens when two or more validators become block producers in the same time slot, and the block produced by one validator reaches the next block producer before the other blocks. We call lagging blocks uncle blocks.

At the end of each Era, rewards will be issued. Regardless of how much the validator has staked, all validators share the block production rewards evenly. However, the rewards for specific validators may vary depending on the Era points, as mentioned above. Although obtaining Era points has a probabilistic component and may be slightly affected by factors such as network connectivity, well-performing validators should usually have a similar sum of Era points on average across a large number of Era points.

In addition, the validator can also receive a "tip" from the transaction sender as an incentive to include the transaction in the block it generates.

When a consensus node exits the network, it needs to go through a cooling-off period. During the cooling-off period, the pledged deposit is frozen and can only be redeemed after it is unfrozen.

7.7 Storage Node Incentive

The main functions of storage nodes include providing storage space, storing data, providing downloads, and calculating data proofs. Storage nodes can control which disk to use and how much space to use for CESS network services. The larger the space provided, the higher the proportion of benefits obtained. Under the same space occupation, the benefits obtained from

storing data are greater than those obtained from providing space. In addition, providing data downloads can not only obtain benefits but also improve their reputation. The higher the reputation, the greater the chance of storing data and obtaining high returns.

Storage nodes prove the validity of their certified idle space and stored service data by completing random challenges. These verified storage computing power will serve as the basis for sharing with the CESS network. After successfully completing the random challenge, storage nodes can share token rewards based on their own storage computing power in the entire network.

The bonus comes from the CESS network. Each Era generates a fixed number of CESS tokens, which will be distributed based on the proportion of the current bonus (total_reward) and the computing power of the storage nodes to the total computing power (total_power) of the current round. The computing power of the storage nodes consists of two parts: idle space (idle_space) and service space (service_space). This design is to incentivize storage nodes to store real data.

Storage node computing power = service data volume(in bytes) * 0.7 + idle space volume(in bytes) * 0.3

This round reward = this round prize pool total_reward * (computing power of storage nodes/total computing power of this round total_power(in bytes))

Storage nodes join the storage network by pledging a certain amount of tokens, and the pledged amount is proportional to the declared storage computing power. After joining the network, storage nodes can apply to exit the network at any time. Before the node goes offline, it needs to assist the CESS network in completing data transfer to ensure the integrity of user data in the storage network.

If a storage node fails to complete random challenges multiple times during the service period (such as shutdown, power outage, network disconnection, killing mining processes, unplugging hard drives, deleting user stored data, etc.), it will be forcibly kicked out of the network, and the pledged funds in the node account will also be deducted.

8. Decentralized Transactions and Storage Mining

8.1 Storage Markets: Verifiable and Trusted Markets

On the commercial storage market, there is an industry chain of "storage suppliers for applications to end users". CESS will improve this industry chain and create an open trading market for clients. 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 CESS storage system to store data.The CESS storage market agreement is based on the following:

- **Placing Orders：** Storage prices are open and transparent. Clients can decide their own order prices based on market situations, and submit orders to the CESS chain. Only blockchain-approved orders can be accepted by the system. Once orders are accepted, they cannot be modified.
- **Storage miners allocating resources：** In order to maintain the stability of the storage market and prevent bad behaviors from storage miners, storage miners must stake a certain number of tokens in proportion to their storage size to system token pool. The transaction fees paid by clients are put in the token pool too. Only after the verification process is completed, the transaction fees will be transferred to the storage miners' accounts.
- **Self-organized processing:** Storage miners must periodically report and prove the integrity of their stored data to verifiers. Verifier nodes must conduct verifications.

8.2 Storage Mining: Commercial Implementation of Decentralized Storage

The implementation of decentralized storage requires miners to store valid data not random data. The storage miners need to become qualified and to stake a collateral in CESS tokens. If miners do not keep their promise and data integrity can't be verified, the system will deduct penalties from their accounts. The transaction fees will be refunded to clients. If data integrity verifications are successful, the miners are rewarded with CESS tokens in their accounts.

8.3 Storage Brokers: Improving Resource Integration in the Economy

Decentralization is not absolutely equivalent to disintermediation, especially in scenarios such as enterprise-level resource allocation, where the matching of individual storage miners and storage demand in the trading market is obviously inefficient and uneconomic. The emergence of storage brokers solves the problem. They can provide and match large scale storage demands with storage resources. The existence of broker service will greatly improve the efficiency of the storage market.

9. Community Governance

The CESS Decentralized Autonomous Organization (DAO) is represented by a set of computer programs with transparency. CESS token holders within the ecosystem can become members of CESS DAO of the highest order authority, independent of any centralized agency. Each member can participate in issuing proposals and voting resulting in the governance of the community. In the CESS ecosystem, the community governance structure is developed with the formation of community consensus to have fair, just, and effective governance, and with important suggestions for ecological development to drive the CESS Decentralized cloud storage and data network ecosystem to their full potential.

The assets management of CESS DAO monitored in the market, is achieved in an open, transparent manner as with community governance. CESS DAO strives to achieve decentralization with the openness of rules, codes, the entire incentive system, and the regulatory mechanism to be publicly available. Everyone in the decentralized community can participate equally with transparency in community governance and operation.

10. Future Outlook

Don Tapscott, one of the world's leading authorities on the impact of technology on business and society, once said, "Technology likely to have the greatest impact on the next few decades has arrived. And it's not social media, it's not big data, it's not robotics, it's not even AI. Imagine a scenario where instead of just the Internet of information, there existed an Internet of value. This would entail an expansive, worldwide, distributed ledger accessible to all, operating across millions of computers. Within this framework, various assets, ranging from currency to creative works like music, arts, and movies, could be securely stored, transferred, traded, and managed without the need for centralized intermediaries. You will be surprised to learn that it's the underlying technology of digital currencies like Bitcoin. It's called Blockchain. Now it's not the most sonorous word in the world, but we believe that this is now the next generation of the Internet, and that it holds vast promise for every business, every society, and for all of you, individually."

Quoting Don Tapscott, CESS, is a firm believer in the digital economy brought forth by Blockchain technology, fully devoted to promoting the interconnection of data in an open, impartial, and secure network environment, and extremely motivated towards the development of Web 3.0.

11. References

1. Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." Decentralized Business Review (2008): 21260.
2. Castro, Miguel, and Barbara Liskov. "Practical byzantine fault tolerance." OSDI. Vol. 99. No. 1999. 1999.
3. Castro, Miguel, and Barbara Liskov. "Practical Byzantine fault tolerance and proactive recovery." ACM Transactions on Computer Systems (TOCS) 20.4 (2002): 398-461.
4. Bach, Leo Maxim, Branko Mihaljevic, and Mario Zagar. "Comparative analysis of blockchain consensus algorithms." 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2018.
5. Gramoli, Vincent. "From blockchain consensus back to Byzantine consensus." Future Generation Computer Systems 107 (2020): 760-769.

6. Milutinovic, Mitar, et al. "Proof of luck: An efficient blockchain consensus protocol." proceedings of the 1st Workshop on System Software for Trusted Execution. 2016.
7. Wood, Gavin. "Ethereum: A secure decentralised generalised transaction ledger." Ethereum project yellow paper 151.2014 (2014): 1-32.
8. Data object store and server for a cloud storage environment, including data deduplication and data management across multiple cloud storage sites, 2012.
9. Lakshman, Avinash, and Prashant Malik. "Cassandra: a decentralized structured storage system." ACM SIGOPS Operating Systems Review 44.2 (2010): 35-40.
10. Lipton, Alexander, and Adrien Treccani. Blockchain and Distributed Ledgers: Mathematics, Technology, and Economics. World Scientific, 2021.
11. Bhutta, Muhammad Nasir Mumtaz, et al. "A Survey on Blockchain Technology: Evolution, Architecture and Security." IEEE Access 9 (2021): 61048-61073.
12. Benet, Juan. "Ipfs-content addressed, versioned, p2p file system." arXiv preprint arXiv:1407.3561 (2014).
13. Karagiannis, Thomas, et al. "Transport layer identification of P2P traffic." Proceedings of the 4th ACM SIGCOMM conference on Internet measurement. 2004.
14. Pouwelse, Johan, et al. "The bittorrent p2p file-sharing system: Measurements and analysis." International workshop on peer-to-peer systems. Springer, Berlin, Heidelberg, 2005.
15. Balakrishnan, Hari, et al. "Looking up data in P2P systems." Communications of the ACM 46.2 (2003): 43-48.
16. Rhea, Sean, et al. "Handling churn in a DHT." Proceedings of the USENIX Annual Technical Conference. Vol. 6. 2004.
17. Rhea, Sean, et al. "OpenDHT: a public DHT service and its uses." Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications. 2005.
18. Micali, Silvio, Michael Rabin, and Salil Vadhan. "Verifiable random functions." 40th annual symposium on foundations of computer science (cat. No. 99CB37039). IEEE, 1999.
19. Dodis, Yevgeniy, and Aleksandr Yampolskiy. "A verifiable random function with short proofs and keys." International Workshop on Public Key Cryptography. Springer, Berlin, Heidelberg, 2005.
20. Bitansky, Nir. "Verifiable random functions from non-interactive witness-indistinguishable proofs." Journal of Cryptology 33.2 (2020): 459-493.

21. David, Bernardo, et al. "Ouroboros praos: An adaptively-secure, semi-synchronous proof-ofstake blockchain." Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2018.
22. Lamport, Leslie. "Paxos made simple." ACM Sigact News 32.4 (2001): 18-25.
23. Chang, Fay, et al. "Bigtable: A distributed storage system for structured data." ACM Transactions on Computer Systems (TOCS) 26.2 (2008): 1-26.
24. Dimakis, Alexandros G., et al. "Network coding for distributed storage systems." IEEE transactions on information theory 56.9 (2010): 4539-4551.
25. Rawat, Ankit Singh, et al. "Locality and availability in distributed storage." IEEE Transactions on Information Theory 62.8 (2016): 4481-4493.
26. Hasan, Ragib, et al. "A survey of peer-to-peer storage techniques for distributed file systems." International Conference on Information Technology: Coding and Computing (ITCC'05)-Volume II. Vol. 2. IEEE, 2005.
27. Herlihy, Maurice. "Atomic cross-chain swaps." Proceedings of the 2018 ACM symposium on principles of distributed computing. 2018.
28. Wood, Gavin. "Polkadot: Vision for a heterogeneous multi-chain framework." White Paper 21 (2016).
29. Amiri, Mohammad Javad, Divyakant Agrawal, and Amr El Abbadi. "Caper: a crossapplication permissioned blockchain." Proceedings of the VLDB Endowment 12.11 (2019): 1385-1398.
30. Garoffolo, Alberto, Dmytro Kaidalov, and Roman Oliynykov. "Zendoo: a zk-SNARK verifiable cross-chain transfer protocol enabling decoupled and decentralized sidechains." 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2020.
31. Green, Matthew, and Giuseppe Ateniese. "Identity-based proxy re-encryption." International Conference on Applied Cryptography and Network Security. Springer, Berlin, Heidelberg, 2007.
32. Ateniese, Giuseppe, et al. "Improved proxy re-encryption schemes with applications to secure distributed storage." ACM Transactions on Information and System Security (TISSEC) 9.1 (2006): 1-30.
33. Ateniese, Giuseppe, Karyn Benson, and Susan Hohenberger. "Key-private proxy reencryption." Cryptographers' Track at the RSA Conference. Springer, Berlin, Heidelberg, 2009.
34. Libert, Benoit, and Damien Vergnaud. "Unidirectional chosen-ciphertext secure proxy reencryption." IEEE Transactions on Information Theory 57.3 (2011): 1786-1802.
35. Rumelhart, David E., et al. "Sequential thought processes in PDP models." Parallel distributed processing: explorations in the microstructures of cognition 2 (1986): 3-57.

36. Curtmola, Reza, et al. "MR-PDP: Multiple-replica provable data possession." 2008 the 28th international conference on distributed computing systems. IEEE, 2008.
37. Shacham, Hovav, and Brent Waters. "Compact proofs of retrievability." International conference on the theory and application of cryptology and information security. Springer, Berlin, Heidelberg, 2008.
38. Juels, Ari, and Burton S. Kaliski Jr. "PORs: Proofs of retrievability for large files." Proceedings of the 14th ACM conference on Computer and communications security. 2007.
39. Barsoum, Ayad F., and M. Anwar Hasan. "Provable multicopy dynamic data possession in cloud computing systems." IEEE Transactions on Information Forensics and Security 10.3 (2014): 485-497.
40. Sadowski, Caitlin, and Greg Levin. "Simhash: Hash-based similarity detection." Technical report, Google (2007).
41. Uddin, Md Sharif, et al. "On the effectiveness of simhash for detecting near-miss clones in large scale software systems." 2011 18th Working Conference on Reverse Engineering. IEEE, 2011.
42. Hovav Shacham and Brent Waters. "Compact proofs of retrievability." International conference on the theory and application of cryptology and information security. Springer, Berlin, Heidelberg, 2008.
43. Qian Wang, Cong Wang, Jin Li, Kui Ren, and Wanjin Lou. "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing", 2009
44. Ren L, Devadas S. Proof of space from stacked expanders[C]//Theory of Cryptography Conference.Springer,Berlin,Heidelberg, 2016: 262-285.
45. Goodrich M T, Tamassia R, Hasic J. An efficient dynamic and distributed RSA accumulator[J]. arXiv preprint arXiv:0905.1307, 2009.
46. Dziembowski S, Faust S, Kolmogorov V, et al. Proofs ofspace[C]//Annual Cryptology Conference.Springer,Berlin,Heidelberg, 2015: 585-605.