Layout guide for *Journal of Physics: Design and Research of Distributed Educational Resource Sharing Platform Based on IPFS*

Qinghe Zhou¹, Ju Lu², Hongjie Yin³ and Dong He⁴

¹,²,³,⁴ School of Software Engineering, Chengdu University of Information Technology, ChengDu, SiChuan, 610225, China

*Corresponding author's e-mail: 3190702027@stu.cuit.edu.cn*

Abstract. In the age of information network, blockchain technology has become a key focus. With the integration of new technologies, blockchain covers many fields such as finance, social welfare, agriculture, smart cities, and education. At the same time, education has undergone fundamental changes along with the continuous advancement of technology, and under the impetus of blockchain and other technologies, "technology better empowers education" is an important goal of education informatization workers. Therefore, this article will explore its application scenarios in the field of "blockchain + education" based on blockchain and IPFS-related technologies and elaborate on how IPFS technology can improve the efficiency of educational resource sharing, optimize the process of sharing and dissemination of educational resources, promote the improvement of the quality of science education and build the application prospects of science education resource sharing.

1. Introduction

In 2020, the sudden outbreak of new coronary pneumonia triggered an "unprecedented, unprecedented" large-scale online teaching practice. "Education is there where the students are", crossing the boundaries of school, family and society, integrating online and offline, making education more possible. In response to the epidemic, schools in 160 countries have been closed, and more than 1.5 billion students have been affected. In the post-epidemic era, China's education has entered a new stage of high-quality development. It is necessary to leverage the advantages of online education and build a learning system for all people with more flexible methods, richer resources, and more convenient learning. The "freshness" is changing to the "new normal". The combination of online teaching and "Internet + Blockchain" technology has become an important development direction of higher education in China and even the world. Therefore, effective sharing of educational resources is extremely critical and important.

P2P network is the basis for the development of blockchain technology, and IPFS[1](InterPlanetary File System) is a distributed, new hypermedia transmission protocol based on content addressing. IPFS is a distributed file system whose purpose is to connect all computer equipment to the same file system, making it a globally unified storage system. This article will be based on the combination of Spring Cloud and IPFS to solve the problem of educational resource sharing.
2. Technical Summary

2.1. IPFS and blockchain technology
Speaking of IPFS, we must mention the blockchain. The blockchain is a block containing specific information (database), but it is grouped in the network in a safe and reliable way (peer-to-peer). In other words, the blockchain is a connected computer, not a central server, which means that the entire network is decentralized and distributed. But this is different from traditional distributed databases. From the purpose of design, the design purpose of blockchain is to solve the problem of trustworthiness of data in a non-trusted environment. A non-trusted environment means that the node responsible for data storage can be free modifying the data and other participating nodes are unable to recognize each other, which creates the problem of mutual mistrust between various nodes. The traditional distributed data storage system is built in the environment of full trust, which is different from blockchain system in data storage and management. It is specifically embodied in:

- The two are different in decentralized topology. Traditional distributed systems do not have the ability to decentralize, while the distributed system built by blockchain adopts a distributed model based on P2P in terms of structure. The joining and exiting of nodes are arbitrary and dynamic.
- Data distribution methods are different. In traditional distributed systems, data is stored on different nodes by horizontal and vertical sharding. In blockchain distributed systems, data storage uses partial replication. Or full copy mode.
- Data consistency protection methods are different. The most important thing in a blockchain system is to maintain data consistency. Usually, efficient algorithms such as Paxos, RAFT[2], and practical Byzantine fault-tolerant PBFT are used to maintain data. Consistency.
- The processing method of data query is different. In the blockchain system, the query of all ledger information is performed locally at the node that participates in the storage of data. Due to its storage structure and lack of index structure, it can only scan all block data sequentially, Currently usually recorded in the Key-Value database.

IPFS integrates the technical advantages of distributed hash tables, BitTorrent, Git, and self-verifying file systems with the blockchain to break through the bottleneck of storage capacity in distributed systems.

2.2. Spring Cloud technology
The application architecture ranges from the first generation of tightly coupled, complex systems, intricate interactions, and the whole body when moving. The complex manufacturing of various middleware is a completely closed monolithic architecture to loosely coupled, usually through ESB for system integration. The SOA architecture is a decoupled, scalable, automatically elastically scalable, and highly available microservice architecture. From the initial single project, it should develop into dozens to hundreds of different projects. For example, the meeting will have the following questions about the dependencies between services, how to unpack the services, internal interface specifications, data transmission, etc., especially service unpacking. In order to ensure that the service granularity after splitting not only conforms to the basic principle of “high cohesion and low coupling”, but also takes into account business development.

Spring Boot[4] is a brand new framework provided by the Spring team, and Spring Cloud is to solve practical application problems on the basis of Spring Boot. It cleverly simplifies the development of distributed system infrastructure such as service discovery and registration, configuration center, message bus, load balancing, circuit breakers, and data monitoring.
3. Design and Implementation of a Distributed Educational Resource Sharing Platform

3.1. Overall architecture design

Through the introduction of the second section, the overall technical architecture of the distributed education system is to solve the problem by IPFS+Spring Cloud, where IPFS is used to solve the problem of educational resource storage, and the various components of Spring Cloud solve the problem between services.

The architecture diagram of the distributed educational resource sharing platform is shown in the figure:

![Architecture Design of Distributed Educational Resource Sharing Platform Based on IPFS](image)

3.1.1. Persistence layer. The sharing of educational resources is nothing more than the shared storage of documents, audios, and videos. Use IPFS to freely create permanent distributed storage and share files on the nodes of the network. More and more students and teachers at school are used as nodes to build a distributed file system. In a peer-to-peer network, after each node successfully uploads a file, it will return a unique CID, and the CID value of the same resource of the content is the same. That is, each CID value corresponds to an educational resource.

3.1.2. Business logic layer. First, the entire educational resource sharing system is divided into different service modules, which are mainly divided into four categories: educational resource provision services, educational resource consumption services, resource search services, and resource backup services. In a P2P[3] network, the resource provider is also the resource consumer, and the unique content identification CID is obtained by uploading the resource, while other nodes obtain the desired resource through the resource search service, that is, the file CID value.

3.1.3. Interface display layer. For the education resource sharing system interface is not complicated, use the elFinder network file manager UI interface to operate the file upload. Realize the display of educational resource service search through ElementUI+vue.js.

3.2. Design and implementation of each service module

For the design and implementation of IPFS's distributed educational resource sharing platform, there are three main problems:

- The IPFS-based distributed educational resource sharing platform needs to meet the closure and security to prevent malicious attacks on its nodes
- At the same time, it is also necessary to solve the temporary unavailability of educational resources in the IPFS network, which needs to be solved by a backup buffer node solution
For the particularity of educational resources, to meet the growing demand for educational resource sharing, however, IPFS only provides a method to accurately find data, that is, each file is uploaded to the IPFS system for storage, and the system returns a unique file identifier CID, this method is not applicable to the requestor for educational resources.

The design and realization of the above three issues have become the key to the educational resource sharing platform.

3.2.1. *IPFS exclusive educational resource network.* IPFS aims to connect global nodes to construct a distributed file system, but in the IPFS network, resources are accessed at any node through the CID value of the resource file. For the IPFS educational resource sharing platform system, it is obviously unreasonable to use such a network. On the one hand, all nodes will be exposed to the global network and provide services that are irrelevant to educational resources. This will inevitably affect the query efficiency within the system. The data of one aspect of the node can be obtained by other nodes that do not belong to the educational resource sharing node, and the security is very low. Therefore, the construction of the educational resource sharing platform must be an IPFS exclusive network.

All IPFS nodes will connect to the IPFS public network by default. In order to build an IPFS proprietary educational resource network, verify the network nodes that join the educational resources. At this time, we need a private key to prevent irrelevant nodes or malicious nodes from joining in the proprietary network.

In order to build a dedicated educational resource sharing network, we will first clear the default connection of IPFS with public network nodes, and then use the official IPFS tool ipfs-swarm-key-gen to create a key that can be used for all nodes in the private network, Nodes that do not use the key cannot join the private network. When an IPFS node requests a connection, the key detection is performed. A node without a key cannot connect to any node in the private network. Finally, the nodes are initialized, and the newly added nodes will be automatically networked and added to the education resource sharing network of IPFS. Each node is decentralized.

| Field Name   | Field description                      |
|--------------|----------------------------------------|
| BootStrap    | Default routing node                   |
| NodeID       | Node ID                                |
| Path         | Storage path                           |
| BlockSize    | storage capacity                       |
| PrivateKey   | IPFS exclusive network private key     |
| IPNS         | Specify the initial configuration of IPNS for IPFS |
| Address      | Specify the IPFS communication address port configuration |
| Swarm        | Specify IPFS node configuration        |

3.2.2. *Highly available for educational resources.* In the exclusive network of educational resource sharing, educational resource providers upload educational resources through IPFS and then go offline. Consumers of educational resources want to obtain the educational resources but cannot obtain them at this time. This is the main part of the educational resource sharing platform. There is a huge problem with the sharing function. In IPFS, data distribution is achieved by exchanging blocks with peers using the BitTorrent heuristic protocol. After reading the IPFS white paper and the BitSwap protocol source code, it is found that the debt \( r \) is defined as:

\[
    r = \frac{\text{byte } \_ \text{send}}{\text{byte } + \text{rev} + 1}
\]

For peer nodes with different debts, the probability of being selected as the node for redundant storage of the current data block:
It can be seen that data caching in IPFS judges whether to cache data on the current node by calculating node debt, that is, a probabilistic caching strategy.

In order to further improve and enhance the overall performance of the entire educational resource sharing platform, cache nodes have become an important solution. After each educational resource providing node uploads resource data, the cache node is called to back up the resource files. Provincial education departments or local operators have deployed IPFS buffer nodes. These buffer nodes form a complete IPFS cache system. Each cache node can back up each other. The more nodes there are, the harder it is to lose data.

3.2.3. Efficient index access to educational resources. The resource requester obtained the corresponding resource file through the original method of IPFS is not suitable for use on the educational resource sharing platform. Since IPFS only supports the acquisition of data resources based on CID, the educational resource requester cannot remember the corresponding resource file that he wants to obtain CID value. When the educational resource requester obtains keywords or other description information, it is difficult to obtain relevant educational resources due to the lack of corresponding search functions, so the application of educational resource sharing is restricted. Here we use Elasticsearch to build a search engine on IPFS. The search server built by Elasticsearch serves all local educational resource servers, thereby achieving high concurrency. To set the ES cluster service as the first IPFS education resource node, you also need to add the same key and clear the IPFS default public network connection method and configure the default bootstrap routing node in the IPFS node configuration file. Research on IPFS data acquisition methods can help users search for data according to their own needs, facilitate the sharing and discovery of IPFS data, and enable IPFS to meet more application scenarios.

4. Summarize
The distributed educational resource sharing platform is different from the previous distributed system. The distributed educational resource sharing platform based on IPFS realizes the integration of IPFS and educational resource storage and constructs an educational resource sharing system in a peer-to-peer network; it combines IPFS and Elasticsearch provides storage and retrieval interfaces for educational resources; it realizes an educational resource sharing with high security, efficient search access and distributed disaster recovery storage.

Acknowledgments
This article is supported by the key R&D project of Sichuan Provincial Department of Science and Technology (No. 21ZDYF0055) and the Soft Science Project of China Meteorological Administration (No. 2021ZZXM17).

References
[1] Kumar Randhir, Tripathi Rakesh, Marchang Ningrinla, Srivastava Gautam, Gadekallu Thippa Reddy, Xiong Neal N.. A secured distributed detection system based on IPFS and blockchain for industrial image and video data security[J]. Journal of Parallel and Distributed Computing, 2021(prepublish).
[2] A Study on PBFT Consensus Process Considering Communication Cost Efficiency of Blockchain Network[J]. The Journal of Korean Institute of Information Technology, 2020, 18(4).
[3] S. Vimal, S. K. Srivatsa. A new cluster P2P file sharing system based on IPFS and blockchain technology[J]. Journal of Ambient Intelligence and Humanized Computing, 2019(prepublish).
[4] Minmin He, Shengli Lv. Research on Civil Aircraft Structural Strength Material Performance Database Based on SpringBoot[A]. Advanced Science and Industry Research Center. Proceedings of 2020 4th International Conference on Electrical, Automation and Mechanical Engineering (EAME2020)[C]. Advanced Science and Industry Research Center: Science and Engineering Research Center, 2020: 6.

[5] Bendechache Malika, Svorobej Sergej, Endo Patricia Takako, Mihai Adrian, Lynn Theo. Simulating and Evaluating a Real-World ElasticSearch System Using the RECAP DES Simulator[J]. Future Internet, 2021, 13(4).