CLC_____ Number___

UDC_____ Available for reference Yes ☐ No ☐


# Undergraduate Thesis

**Thesis Title:** A blockchain - based manufacturing data sharing system

**Student Name:** 潘泰仰

**Student ID:** 11811214

**Department:** 计算机科学与工程系

**Program:** 计算机科学与技术

**Thesis Advisor:** 宋轩

Date: May 25th, 2023

# Letter of Commitment for Integrity

1. I solemnly promise that the paper presented comes from my independent research work under my supervisors supervision. All statistics and images are real and reliable.
2. Except for the annotated reference, the paper contents no other published work or achievement by person or group. All people making important contributions to the study of the paper have been indicated clearly in the paper.
3. I promise that I did not plagiarize other peoples research achievement or forge related data in the process of designing topic and research content.
4. If there is violation of any intellectual property right, I will take legal responsibility myself.

Signature:

Date:

# A blockchain - based manufacturing data sharing system

潘泰仰

计算机科学与工程系　指导教师：宋轩

**[ABSTRACT]:** In this thesis, I proposed a data sharing system based on blockchain for smart manufacturing. The system is based upon a consortium blockchain structure that leverages the benefits of distributed ledger technology to provide a reliable and tamper-proof data sharing mechanism. I also presented a detailed design of the system architecture, identifier, encoding and parsing protocols. The proposed system addresses the challenges of efficient storage and secure and effective sharing of large-scale data in a multi-cloud environment. Future work will include the implementation and testing of the system to evaluate its efficacy and efficiency. In general, the proposed scheme has the potential to improve the transparency, traceability, and efficiency of data sharing in the smart manufacturing industry, leading to increased innovation and economic development.

**[Key words]:** Consortium Blockchain; Smart manufactoring; Data sharing; distributed ledger

[摘要]：在这篇论文中，我提出了一种基于区块链的智能制造数据共享系统。该系统基于联盟区块链结构，利用分布式账本技术的优势，提供了一种可靠且防篡改的数据共享机制。我还详细介绍了系统架构、标识符、编码和解析协议的设计。该系统解决了多云环境中大规模数据的高效存储、安全有效共享的挑战。未来的工作将包括对系统进行实施和测试，以评估其有效性和效率。总体而言，我所提出的方案有潜力提高智能制造行业数据共享的透明度、可追溯性和效率，从而促进创新和经济发展。

[关键词]： 联盟区块链；智能制造；数据共享；分布式账本

# Table of Content

# 1. Introduction

## 1.1 Background and motivation for the research

The manufacturing industry is undergoing a transformation as blockchain technologies have been developed. The blockchain technologies have the potential to improve efficiency, transparency, and traceability in the manufacturing process, leading to better quality products and increased customer satisfaction. However, one challenge the industry is facing is how to efficiently store and securely share large-scale data in a multi-cloud environment.

## 1.2 Research objectives and contributions

To address this challenge, I propose a blockchain-based data sharing system for smart manufacturing. The system is based on a consortial blockchain framework that leverages the benefits of distributed ledger technology to provide a reliable and tamper-proof data sharing mechanism.The proposed system is designed to cater to the specific needs of the manufacturing industry, addressing the challenges of efficient storage and secure and effective sharing of large-scale data in a multi-cloud environment.

## 1.3 Overview of the thesis

In this thesis, the objective of the research is to design a blockchain-based data sharing system to meet specific needs in the smart manufacturing industry. The system aims to address challenges such as efficient storage and secure sharing of large-scale data in a multi-cloud environment. To accomplish this objective, I propose a system design based on a consortium blockchain framework and provide comprehensive descriptions of its architecture, identifiers, encoding, and parsing protocols.

This thesis proposes a blockchain-based data sharing system that provides a reliable and tamper-resistant mechanism for data sharing in the smart manufacturing industry. By designing a consortium blockchain framework specifically tailored to the industry's needs and providing detailed designs of the system architecture, identifiers,

encoding, and parsing protocols, specific technical solutions are offered for implementing the system. This novel approach to data sharing offers economic benefits and competitive advantages to stakeholders in the smart manufacturing industry, fostering innovation and industry growth. The remainder of this thesis is organized as follows: Section 2 provides an overview of blockchain technology and its applications in manufacturing. Section 3 presents my proposed methodology in detail, including the design of the system architecture (Section 3.1) and the identifier design (Section 3.2). The results and discussion are presented in Section 4. Finally, Section 5 concludes with recommendations for later work.

# 2. Literature Review

## 2.1 Overview of blockchain technology and its application

Smart manufacturing is becoming increasingly important in the global manufacturing industry.

Investment in smart manufacturing paves the way for increased efficiency, production, process quality, energy savings, saving on fixed costs in the long term, increased reactivity to changes and increased competitiveness in the global market.

As detailed in the White House Critical and Emerging, the manufacture of Smart is a national focus.

This is an opportunity to create new business value and to dismantle the existing manufacturing value chain through the injection of appropriate technology. The new model of value creation opens up an avenue for myriad innovation opportunities and also encourages organisations to be far more agile.

Blockchain technology has vast opportunities and benefits for improving . Some benefits of using blockchain for data management and sharing in smart manufacturing include improved inventory management, improved data security, improved transparency and traceability, automated vendor payments, and improved customer engagement .

## 2.2 Related works

Related technologies for data sharing fall into two categories. With the former focusing on the security and the latter focusing on frameworks for sharing data. I will introduce them independently in what follows.

### 2.2.1 Security

In[1], this paper proposes an attribute-based dynamic access control scheme . In [2], their proposal involves a solution for lightweight proxy re-encryption, which involves the construction of a pre-encryption algorithm and the development of a certificate-free protocol. This protocol eliminates the need for bilinear pairs and is highly efficient in terms of performance. As an example, Qin's work [3] involves the use of computed tokens that traverse domains in order to minimize computational overheads for users accessing data. Qin also conducted additional research to implement access control mechanisms, as described in [4,5]. This approach can help improve the efficiency and security of data sharing across different domains. Several other studies have also explored the intersection of the internet of things, privacy, and safety. These studies, including [6~9], have investigated various approaches to enhancing the security and privacy of networks.

### 2.2.2 Frameworks

Various frameworks are proposed depending on the scenario. One approach is on the basis a centralized server framework that collects and organizes shared data stored on different cloud platforms into a unified format, and stores it on a centralized server for user access. When a user wants they can share data by retrieving the unified formatted data to be located and obtained the desired information. While this framework is highly effective, it does have a potential issue with a sole point of failure(SPOF). In the event of a centralized server failure, users may experience significant losses in both data and economic value. It is important to consider this potential risk and implement measures to mitigate it.

Another is based on blockchain technology. It stores and manages data using a distributed ledger of blockchain. Unlike centralized server-based frameworks,

blockchain-based frameworks do not require a single central server to store and manage data, but instead manage data through a distributed node network. This decentralized architecture makes data more secure because there is no single point that can be attacked or damaged by a failure. And due to the encryption and verification mechanisms of blockchain technology, data in blockchain-based frameworks is more difficult to tamper with or damage. Therefore, this framework is very suitable for storing sensitive data such as financial and medical data. However, its disadvantages include potential performance issues, high costs, and privacy concerns.

### 2.2.3 Challenges and limitations of existing solutions

Additionally, there is a need to enhance the robustness of these existing methods, as discussed earlier. Secondly, it is worth noting that some existing approaches to data validation rely solely on a single signature check, which may not be sufficient to meet the needs of users in the era of big data[10]. As data volumes continue to grow, it becomes increasingly important to ensure that data is not only authentic but also accurate and reliable.

# 3.Methodology

To overcome the challenges mentioned above, I propose a new system based on the blockchain consortium. It is more efficient than the public chains such as Bitcoin or Ethereum. There is a network of nodes which are responsible for verifying and validating transactions. These nodes are typically owned and operated by different participants in the manufacturing ecosystem, such as manufacturers, suppliers, and distributors. The system is designed with a encoding-parsing algorithm. Also for the performance of storing large data blocks, the system uses cloud storage or distributed databases.

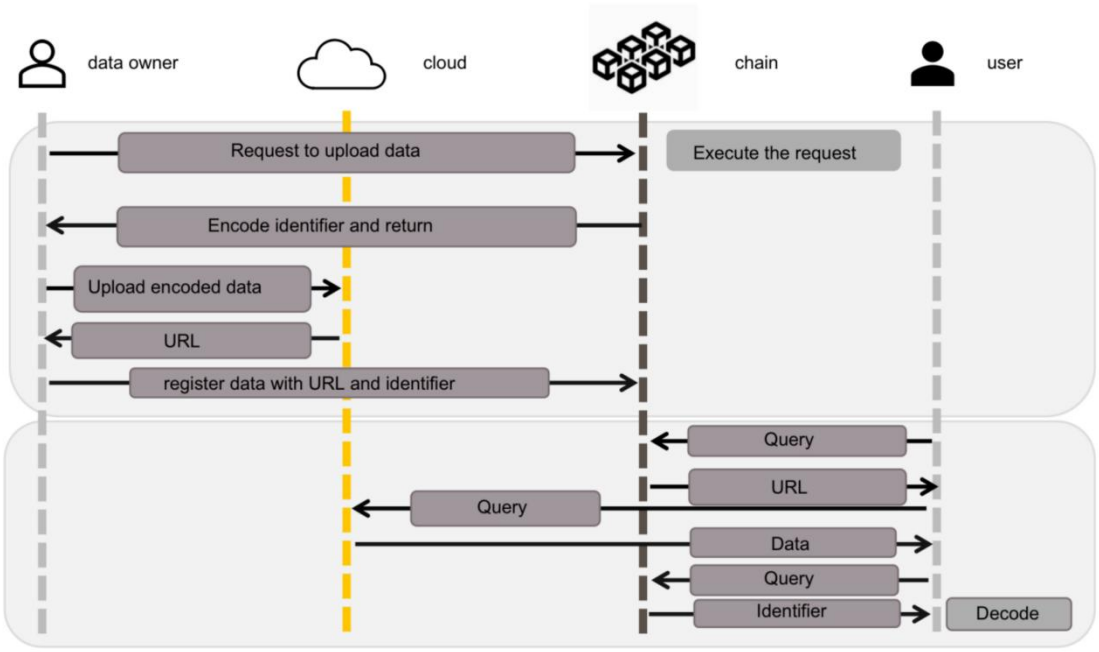


Fig.1

## 3.1 System architecture design

The simplified system architecture depicted in Figure 1 consists of several key components, including data owners, data users, cloud storage, and networks. By leveraging these different elements, the system is able to facilitate the secure and efficient sharing of data across multiple domains. Data owners upload their information by performing the registration process and thus become the managers of their own data in the cloud. Meanwhile, data users can access this information as needed, with appropriate access controls in place to ensure that sensitive data remains protected. The workflow of the system can be described as follows:

a) Data owners are responsible for uploading their data to the cloud and registering it on the blockchain network. This allows them to maintain complete control over their information, including the ability to modify or delete it as needed.

b) Data users are defined as individuals or entities who make use of shared data and have the ability to query and access information uploaded by other users. This allows for a collaborative approach to data sharing, with users able to leverage the insights and knowledge of others in order to achieve their goals.

c) Once data has been recorded on the blockchain network, it can be shared among users and stored in multi-clouds for safekeeping.

d) As a fundamental component of the system, the blockchain is responsible for storing metadata and managing the processes involved in encoding and analyzing data identifiers. This includes recording key information about each piece of data, such as its source, format, and content, as well as ensuring that it is properly encrypted and secured against unauthorized access or manipulation.

The detailed design of the protocols will be discussed in the following sections.

Ethereum can be integrated with any blockchain through smart contracts. However, it is worth noting that Bitcoin may not be the most efficient option for large data-sharing systems due to its inherent limitations. Alternative solutions, such as Permacoin, have been proposed to address these issues[11]. Moreover, selfish mining attacks have been reported to occur in the Bitcoin network[12].

The proposed system differs from existing methods primarily in the way that metadata is stored and managed. Unlike traditional approaches, which often rely on centralized databases or file systems to store metadata, the proposed system leverages the power of blockchain technology to provide a more secure and transparent platform for managing this critical information. In the proposed system, metadata is stored on the blockchain, and the features of blockchains and smart contracts are utilized to implement the systems functions. However, this approach introduces a potential issue: new nodes must obtain a full copy of the metadata before joining the system, which can lead to security and efficiency concerns[13].

To address these challenges, the proposed system could incorporate the following enhancements:

a) Implement a more efficient blockchain network, such as Ethereum or other alternatives, to improve the overall performance of the system.

b) Introduce a mechanism for partial metadata retrieval, allowing new nodes to obtain only the necessary metadata instead of the entire dataset. This can help reduce the burden on new nodes and improve system efficiency.

c) Employ advanced encryption techniques to protect the metadata stored on the blockchain, ensuring the security and privacy of data owners and users.

d) Develop a robust data verification protocol to prevent unauthorized access and

manipulation of data, further enhancing the security of the system.

By incorporating these improvements, the proposed system can effectively address the limitations of existing data-sharing methods while leveraging the advantages of blockchain technology and smart contracts. This will result in a more secure, efficient, and scalable solution for data owners and users alike.

## 3.2 Identifier design

In order to enhance the efficiency of my proposed system, I have designed a unique system of pairs of short and long identifiers. One of the key features of this system is its unique approach to data representation, which involves the use of both short and long identifiers.

By leveraging this approach, users are able to represent data using short identifiers that are easy to manage and share, while also benefiting from more comprehensive descriptions of the data provided by the long identifiers. This allows for a more flexible and efficient approach to data management, with users able to quickly and easily access key information about each piece of data as needed.

When updates are made to the data, the long identifier is modified in order to reflect these changes. However, the short identifier remains constant, providing a stable reference point for users even as the underlying data evolves over time. This helps to ensure that users can easily track changes and updates to the data without having to constantly update their own records or references. Additionally, by separating out metadata from the underlying data itself, this approach helps to improve overall system performance and scalability while reducing complexity and overhead.

Since each piece of data has only one short identifier, it is crucial for it to be unique. In the system, I utilize syndicate chains where each user has a distinct certification. The hash of these certifications, denoted as cert_hash, is used as a prefix for the short identifier. Data owners assign unique names to their data within their User Domain, and these names are appended as a suffix to the cert_hash. This creates a complete short identifier that is unique within the Global Domain. Then the

certification hash is a way to accelarate user interactions and storing public keys.

The long identifier, on the other hand, consolidates all metadata into a single identifier, which also serves as a tool for metadata management. The system is designed to provide users with a seamless and intuitive way to access information by leveraging the power of metadata.

By consolidating all metadata into a single long identifier, the system enables users to easily retrieve key information about each piece of data simply by accessing the long identifier. This system allows for seamless integration and management of metadata within the long identifier.

In summary, the identifier pairs offer several advantages, including uniqueness and human readability. These identifiers cater to the dynamic nature of data, and the short identifier prefix is connected to the owner's identity, thus mitigating security concerns such as counterfeiting[14].

To further elaborate on the benefits and implementation of this system, the following points can be considered:

The dual identifier system allows for efficient data management and retrieval, enabling users to quickly locate and access the information they need.

The unique short identifiers ensure that data remains distinguishable, even when updates or modifications are made, thereby maintaining data integrity within the system.

The long identifiers provide a comprehensive overview of the metadata, making it easier for users to understand the context and relevance of the data they are accessing.

The identifier pairs can be easily integrated with blockchain technology and smart contracts, ensuring secure and transparent data sharing among users.

By linking the short identifier prefix to the owner's identity, the system can effectively prevent unauthorized access or manipulation of data, further enhancing the security and privacy of data owners and users.

Incorporating these features, the proposed system presents a robust and efficient solution for data management and sharing, leveraging the advantages of both short

and long identifiers, as well as the security and transparency offered by blockchain technology and smart contracts.

**Table 1 Identifier pair**

| Field | Description |
|---|---|
| Short Identifier | Short identifier used for data representation and uniqueness |
| Long Identifier | Long identifier consolidating metadata and used for metadata management |
| Cert_hash | Certification hash used as a prefix for the short identifier, representing the user's identity |
| User Domain | User domain, unique names assigned by users within their domain, appended as a suffix to the cert_hash |
| User Information | User information including certificates and public keys, associated with the prefix of the short identifier |
| Metadata | Metadata comprising detailed descriptions and context information about the data, provided by the long identifier |
| Data Integrity | Ensures data integrity through the uniqueness of the short identifier and consolidation provided by the long identifier |
| Data Management | Operations related to data storage, updates, and retrieval |
| Security and Privacy | Security and privacy considerations achieved through the identifier design and integration with blockchain technology |

**Table1 Identifier pair(continued)**

| Field | Description |
|---|---|
| Blockchain Integration | Integration of the identifier system with blockchain technology and smart contracts, enabling secure and transparent data sharing |
| Data Sharing | Secure and transparent data sharing among users facilitated by the short identifier and blockchain technology |

## 3.3 Encoding and parsing protocols design

### 3.3.1 Encoding protocol

The identifier encoding protocol is designed to facilitate data registration and sharing in the system. It provides a standardized process for data owners to register their data in the blockchain system. The protocol consists of four steps:

a) Data owners upload their data to the cloud storage.

b) The cloud storage stores the data and returns the URL of the data to the data owner.

c) Data owners send a registration request to the blockchain network, indicating their intention to share the data with others.

d) The blockchain network generates a short identifier and a long identifier for the data and returns the short identifier to the data owners.

The identifier encoding protocol has several advantages. First, it ensures the uniqueness of identifiers, as the blockchain generates short identifiers using the combination of data name, digital signature, and transaction hash. Second, the protocol allows data owners to manage their data metadata through long identifiers. Third, the protocol is easy to use, as data owners only need to follow the four-step process to register and share their data.

Compared to other systems like Ethereum and Bitcoin, the identifier encoding protocol in the system is designed to overcome their limitations. For example, in Ethereum, users need to repeatedly try to find a unique identifier[15], while in Bitcoin,

users cannot name their own identifiers[16]. the identifier encoding protocol provides a more efficient and user-friendly solution for data registration and sharing.

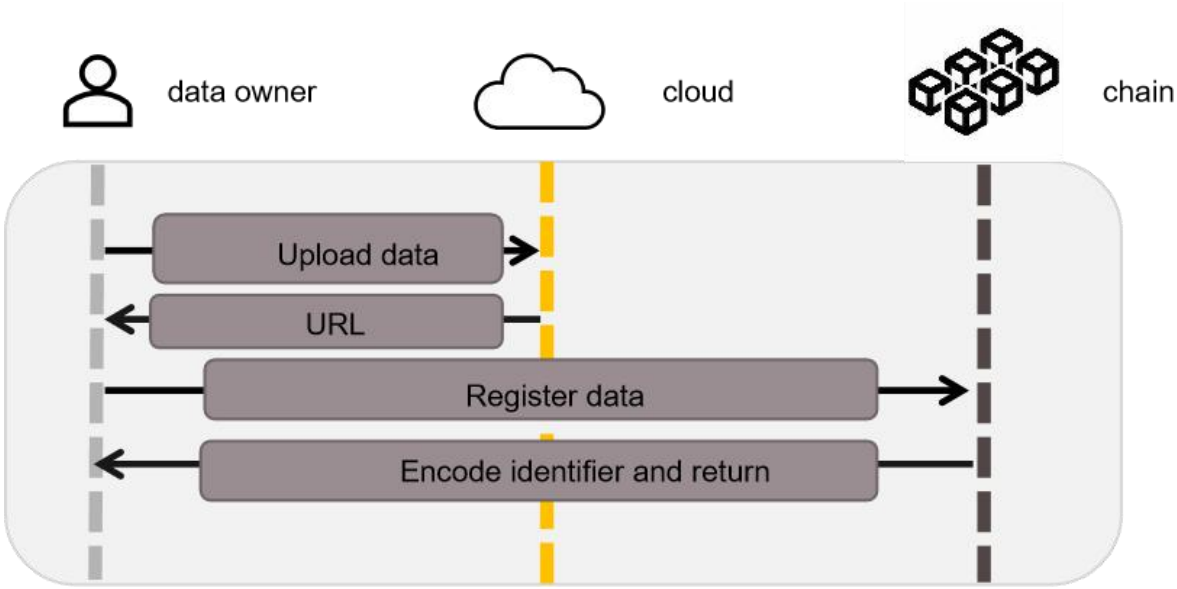


Fig.2 Identifier encoding protocol

3.3.2 Parsing protocol

The identifier parsing protocol is designed to enable data users to query data using short identifiers, while keeping the long identifiers invisible to users. This protocol consists of four steps, where data users first send a query request with a short blockchain identifier; which then executes the parsing process and sends the URL of the metadata back to the users of the data. The data users can then send a query request to the cloud to obtain the data, and the clouds will send the data to the data users.

The parsing process works by retrieving the extract the long identifier from the key pair, use the short identifier as the key, and then retrieve the necessary metadata from the long identifier. The metadata required in this protocol is the URL. This makes it easy for data users to obtain data with only one short identifier, and provides a convenient way to implement extended functions that require certain metadata.

This protocol is important because it ensures that data users can securely and easily access data with short identifiers, while keeping the long identifiers and other sensitive metadata hidden from users. It also provides a secure and efficient way to retrieve metadata, and allows for expanded functions to be easily implemented.

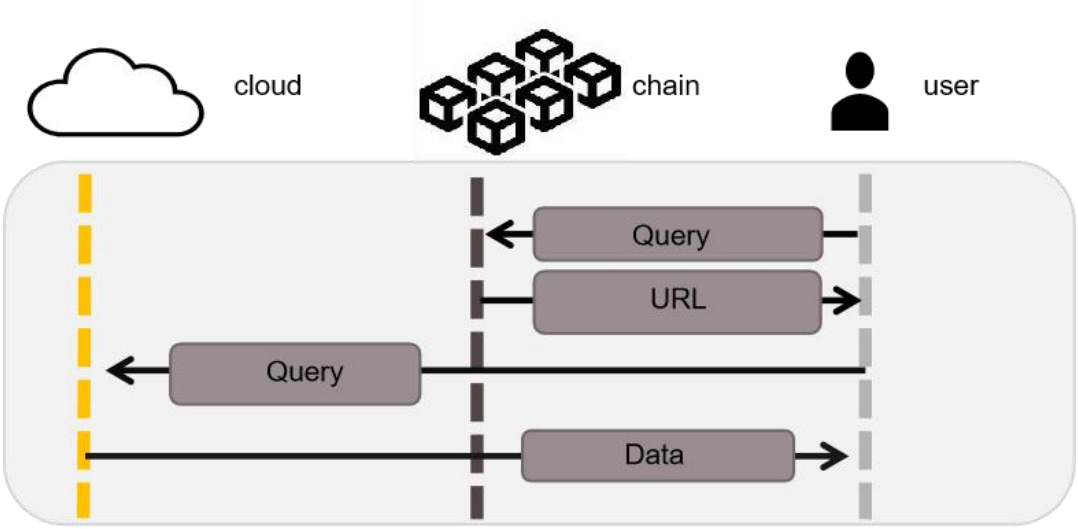


Fig.3 Parsing protocol

3.3.3 Updating protocol

In order to address the dynamic nature of data and the need for frequent updates, the system has been designed with a robust protocol for updating identifiers. This protocol enables data owners to easily update their data both in the cloud and on the blockchain, ensuring that all relevant information is kept up-to-date and accurate. By providing a flexible and scalable approach to data management, this protocol helps to ensure that users can easily adapt to changing circumstances and evolving data requirements. Whether it's a minor update or a major overhaul of the underlying data, this protocol provides a streamlined and efficient way to make changes while minimizing disruption or downtime.

The identifier updating protocol is designed to allow updating the data in the cloud and blockchain. This protocol consists of four steps. In step one, data owners send an updating request to clouds with the new version of data. Then the cloud response with a new URL. And user sand an updating request with the new URL to the blockchain. In the fourth step blockchain reproduces the long identifier and sends reply to the owners of the data.

The system has been designed to provide a robust and secure approach to managing metadata and user information. By leveraging the power of long identifiers, the system enables users to easily store and retrieve key-value pairs along with new values, ensuring that all relevant information is readily accessible and up-to-date. In

addition, the system's use of blockchain technology provides an added layer of security and authentication.

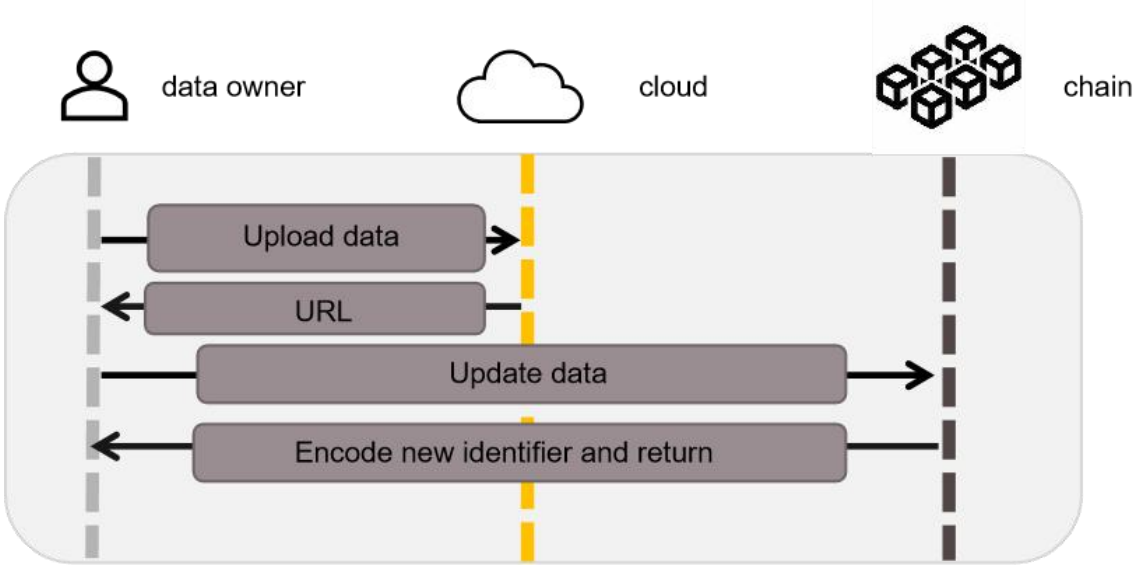


Fig.4 Updating protocol

3.3.4 Data check protocol

After the complete download of data, it is natural to consider data verification. Data users should check the protocol to ensure its validity:

a) whether data has been tampered;

b) Whether the data is indeed provided by the expected data owner.

This verification process is outlined in Figure 4 and consists of three stages:

a) Data users initiate a data verification query to the blockchain.

b) Upon receiving the request, the blockchain retrieves the user_sig and user_pubkey metadata and sends it back to the data users as a response.

c) Successful verification in the third stage confirms that the data corresponds to the short identifier.

This confirmation is based on the fact that the user_sig is generated using the short identifier, providing evidence that the data was provided by the expected data owner. The verification process adds a crucial layer of security and authentication, ensuring that only authorized parties can access or modify the data. The use of user_pubkey in the audit process further supports this confirmation.

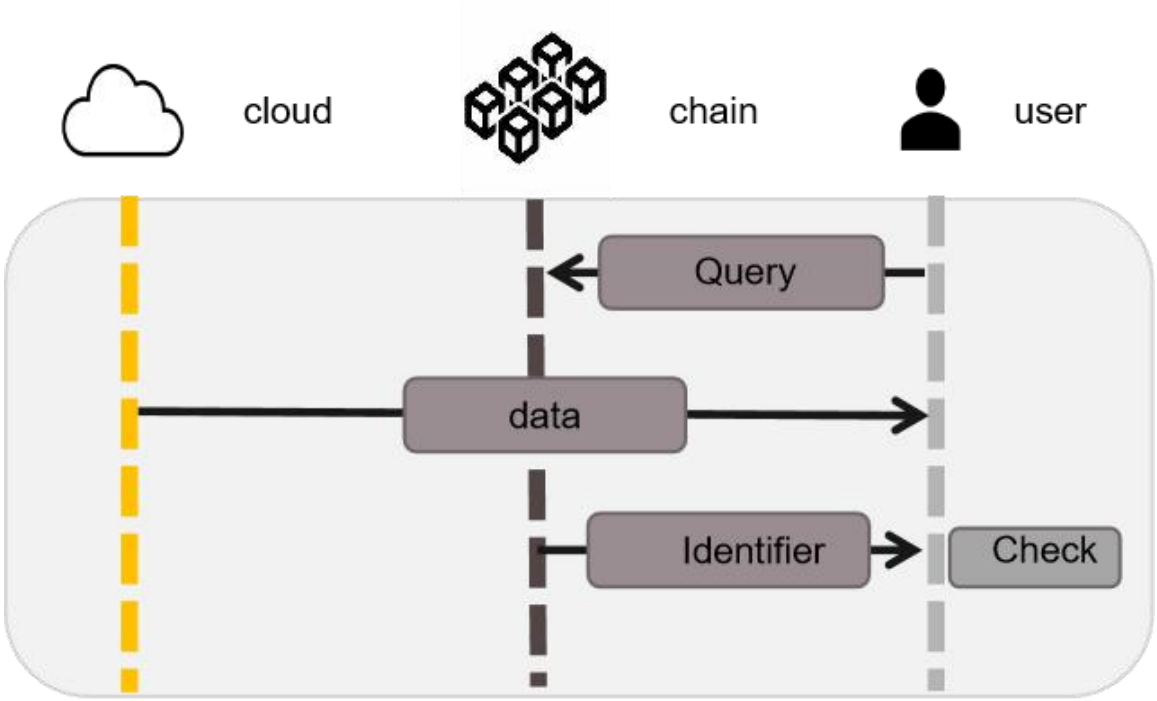


Fig. 5 Data check protocol

# 4. Performance Benchmark and analysis

## 4.1 Performance Benchmark

The proposed system has been benchmarked against two well-established systems, Ethereum[17] and Bitcoin[18], both of which have been successfully implemented and widely accepted by a large user base. Performance and scalability are challenges for the encrypted world, however, I do not have standardized metrics or benchmarks to measure. Data reporting is often inconsistent and incomplete, making accurate comparisons between projects very difficult and often confusing the most important content in practice[19]. In this thesis, I define basic terms, outline the difficult challenges, and provide guidelines and key principles for evaluating blockchain performance to keep in mind.

I propose two metrics: latency, and throughput. Performance indicators may include the number of transactions per second or the median transaction confirmation time. Latency is used to measure the speed of confirming individual transactions, and throughput is used to measure the total rate of transactions over time. In the following tests, I will measure the time from the first broadcast of the user's transaction to the confirmation transaction as an indicator of the delay. Throughput measures the

number of transactions per second. The results of the comparison will be placed in the

**Table 2**

| | My system | Ethereum | Bitcoin |
|---|---|---|---|
| Latency | about 1S (Kafka) | Several mins in PoW and 15s in PoS | about 10mins |
| Throughput | more than 1000 per second | less then 100 per second(Depend on Gas) | less then 10 per second |

In the above comparison, my system is based on the alliance chain of Hyperledger Fabirc, and adopts Kafka as my consensus mechanism. It was proposed and implemented by the Apache Kafka project. The Kafka consensus mechanism is mainly used in distributed systems, aiming to achieve high performance, high scalability and fault tolerance[20]. Kafka consensus mechanism is not based on the traditional consensus algorithm (such as Byzantine fault-tolerant algorithm), but through the partition and copy to achieve the consistency and fault tolerance of data. It is widely used in log transmission, message queue, and event flow processing scenarios in distributed systems, providing reliable data transmission and processing power.

It's worth noting that Bitcoin's transactions are constructed based on its scripting programming language called Script, which is not Turing-complete and based on a stack-like data structure, which will lead to major performance issues with the Bitcoin platform in subsequent tests.

Moreover, my system places a greater emphasis on resource tracing compared to Ethereum and Bitcoin, as this aspect is of significant importance to users. All three systems demonstrate satisfactory performance in query response time.

The following points further elaborate on the advantages of the system compared to Ethereum and Bitcoin:

a) The unique identifier system employed in my solution allows for more efficient data management and retrieval, providing users with faster access to the

information they need.

b) The system's focus on resource tracing ensures that users have better control over their data and can track its usage and modifications more effectively.

c) The integration of blockchain technology and smart contracts in my system enhances its security, transparency, and trustworthiness, setting it apart from Ethereum and Bitcoin.

d) The scalability of my system is superior to that of Bitcoin, as it is designed to accommodate a growing number of users and data without compromising performance or efficiency.

The system's superior theoretical performance in identifier registration and querying ensures that it can handle large volumes of data and user requests more effectively than Ethereum and Bitcoin.

In conclusion, the proposed system demonstrates a more efficient and scalable solution for data management and sharing when benchmarked against existing systems like Ethereum and Bitcoin. By leveraging the advantages of unique identifier pairs, resource tracing, and the integration of blockchain technology and smart contracts, the system offers a robust and secure platform for users to manage and share their data.

## 4.2 Results and analysis

I implemented my system to evaluate its performance and the experimental environment is as follows.

Blockchain: Fabric release 2.4 deployed in docker.

Testing tool: caliper v0.2.0.

Apple M1 Pro chip with 6 Performance core and 2 Efficiency core with 32GB memory.

Golang 1.19.2.

**Table 3. Latency(seconds)**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| My System | 0.785 | 0.932 | 0.652 | 0.412 | 0.895 | 0.743 | 0.587 | 0.918 | 0.976 | 0.536 |
| Ethereum(PoW) | 8.44 | 16.27 | 7.61 | 13.02 | 9.96 | 11.43 | 6.78 | 15.82 | 4.91 | 10.75 |
| Ethereum(PoS) | 15.36 | 14.92 | 15.08 | 14.74 | 15.19 | 15.06 | 15.12 | 14.98 | 15.24 | 15.07 |
| Bitcoin | 764 | 1173 | 423 | 1891 | 988 | 712 | 1557 | 263 | 1412 | 847 |

I conducted tests on the maximum flow rate (number of transactions per second) for each function and then executed these functions with varying flow rates. Table3 displays the response time of functions with varying transaction numbers per second (i.e., throughput).It can be seen that my system delay mainly comes from the delay caused by network fluctuations, concentrated in less than one second. Ethereum based on the PoW consensus mechanism has less stable latency, compared to the PoS consensus mechanism based latency of around 15 seconds, exactly when a block was packaged in the Ethereum 2.0 white paper. The Bitcoin has a less stable delay because it is also based on a consensus mechanism of PoW. In addition, I also performed stress tests, which showed that the system can perform well under constant stress. In contrast, the other two systems experienced network congestion. Based on my past experiences, Ethereum, with its PoW consensus mechanism, even took up to 20 minutes to confirm a block during congestion. The reason behind this is Ethereum's adoption of the Gas mechanism to limit the amount of content that can be included in a block. When dealing with more complex content, a significant amount of Gas is required for packaging, which easily leads to network congestion. On the other hand, the performance issue with Bitcoin has already been mentioned earlier.

It can be seen that the performance advantages of my system mentioned earlier primarily stem from the performance advantages of the consortium blockchain. A consortium blockchain and a public blockchain are two different types of blockchain networks. A consortium blockchain is suitable for scenarios that require establishing trust and sharing data among restricted participants, whereas a public blockchain offers advantages such as decentralization, security, and transparency. However, in

the context of smart manufacturing, participants in the blockchain share a certain level of trust. Smart manufacturing involves multiple participants, including suppliers, manufacturers, and logistics companies, who need to share and exchange data such as supply chain information and production process data. A consortium blockchain can provide a secure and trustworthy platform for participants to share data while ensuring transparency and traceability. Additionally, a consortium blockchain can improve the efficiency and reliability of supply chain management[21]. Participants in the manufacturing supply chain can share data such as orders, delivery status, and inventory information on the consortium blockchain, enabling real-time visualization and collaborative management of the supply chain. This reduces information asymmetry and errors, enhancing overall supply chain efficiency and reliability. Smart manufacturing typically involves a large number of Internet of Things (IoT) devices and machines that require real-time communication and collaboration[22]. A consortium blockchain can provide a distributed device management platform, ensuring device authentication, data transmission security, and trusted interoperability between devices. Furthermore, the adopted system also provides a platform for deploying decentralized applications (chaincode). By combining these advantages and mitigating some of the drawbacks of consortium blockchains, it is more suitable for the smart manufacturing scenario compared to a public blockchain.

There are concerns among some individuals that the use of blockchains may lead to a reduction in efficiency. To address these concerns, an experiment was conducted to evaluate the impact of a blockchain-based system on the efficiency of retrieving shared data stored across four Ali cloud servers. The experiment aimed to determine whether the use of blockchain technology would result in any significant delays or inefficiencies when accessing and sharing data. By measuring key performance indicators such as response time and throughput, it was possible to assess the impact of the blockchain-based system on overall efficiency. With inspirations of how to perform blank control groups [23],I set up a blank control group that downloads data using an SCP command from Linux with URL directly to learn about the system's performance, and the other group obtains the URL of the system before downloading

the data with the SCP command. I repeated this experiment 10 times to reduce errors caused by the network and each file is 1 MB in size. The results are shown in table 3. It can be seen that the difference between the two groups is very small, and this difference decreases as the number of files increases. The results of the experiment indicate that in situations where the system throughput is high, file download latency becomes a critical factor. However, since the data is stored across multiple Ali cloud servers, it can be treated as a multi-cloud environment. This approach does not result in any loss of performance and can actually improve efficiency by distributing the workload across multiple servers.

# 5. Conclusion

In conclusion, the thesis proposes a novel identifier encoding and parsing system, which is engineered to do so address the challenge of data sharing in multi-cloud environments. The system makes use of namespace and identifier pairs to provide readability and ease of use. The corresponding protocols for data registration, updating, and parsing are also proposed. Benefits of the system are analyzed in a comparison with existing systems, such as Ethereum and Bitcoin, and the evaluation results demonstrate that the proposed protocols are efficient and have minimal impact on latency. The system can facilitate data sharing, improve supply chain management efficiency, enable device management, and enhance data security in the context of smart manufacturing. It is suitable for scenarios that require establishing trust and sharing data among restricted participants. However, if the requirement is to build a global open network, achieve higher decentralization that do not require high levels of trust and interoperability, or have higher demands for performance and scalability, then public blockchains like Bitcoin may be more suitable solutions.

# References


[1] X. Qin, Y. Huang, Z. Yang and X. Li, "An access control scheme with fine-grained time constrained attributes based on smart contract and trapdoor."[A] in Proc. of 2019 26th International Conference on Telecommunications (ICT)[C], Hanoi, Vietnam, 2019, 249-253

[2] Qian, Xin, et al., "A No-Pairing Proxy Re-Encryption Scheme for Data Sharing in Untrusted Cloud."[A] in Proc. of International Conference on Artificial Intelligence and Security, Springer[C], Cham, 2019, 85-96.

[3] Qin, Xuanmei, et al., "A Blockchain-based access control scheme with multiple attribute authorities for secure cloud data sharing,"[J] Journal of Systems Architecture, 2020, 101854, 112.

[4] Qin, X., Huang, Y. & Li, X, "An ECC-based access control scheme with lightweight decryption and conditional authentication for data sharing in vehicular networks," Soft Computing, [J]2020, 18881– 18891,24.

[5] Qin, Xuanmei, et al., "LBAC: A Lightweight Blockchain-based Access Control Scheme for the Internet of Things,"[J] Information Sciences, 2021, 554, 222-235.

[6] Le Nguyen, Bao, et al., "Privacy preserving blockchain technique to achieve secure and reliable sharing of IoT data,"[J] CMC-COMPUTERS MATERIALS & CONTINUA, 2020, 65.1, 87-107.

[7] Yang, Zhen, et al., "Protecting personal sensitive data security in the cloud with blockchain,"[J] Advances in Computers, 2021, 120, 195-231.

[8] Yang, Zhen, et al., "Efficient secure data provenance scheme in multimedia outsourcing and sharing,"[J] Computers, Materials & Continua, 2018, 56.1, 1-17.

[9] Li, Jun, et al., "A distributed privacy preservation approach for big data in public health emergencies using smart contract and SGX,"[J] CMC-COMPUTERS MATERIALS &CONTINUA, 2020, 65.1, 723-741.

[10] Li, Z., Barenji, A. V., & Huang, G. Q. "Toward a blockchain cloud manufacturing system as a peer-to-peer distributed network platform."[J], Computers in Industry, 2018, 100, 186-194.

[11]Bordel, Borja, et al., “Trust provision in the internet of things using transversal blockchain networks,”[J] INTELLIGENT AUTOMATION AND SOFT COMPUTING, 2019, 25.1, 155-170.

[12] Lu, Y., & Xu, X. “Blockchain and IoT based food traceability for smart agriculture.”[A] In Proceedings of the 3rd International Conference on Crowd Science and Engineering - ICCSE '18[C], 2018, 1-6.

[13] Tian, F. “An agri-food supply chain traceability system for China based on RFID & blockchain technology”[A]. In 2016 13th International Conference on Service Systems and Service Management(ICSSSM)[C], 2016, 24-26

[14] Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. “Blockchain technology and its relationships to sustainable supply chain management.”[J] International Journal of Production Research, 2019, 57(7), 2117-2135.

[15] Panarello, A., Tapas, N., Merlino, G., Longo, F., & Puliafito, “A. Blockchain and IoT integration: A systematic survey.” [J] Sensors, 2018, 18(8), 2575.

[16] Xu, L. D., Xu, E. L., & Li, L. “Industry 4.0: State of the art and future trends. ”[J] International Journal of Production Research, 2018, 56(8), 2941-2962.

[17]Vitalik Buterin. “Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform.”[OL]. 2014, https://ethereum.org/

[18] Nakamoto, S. “Bitcoin: A Peer-to-Peer Electronic Cash System.”[OL] 2008, https://bitcoin.org/bitcoin.pdf

[19] Rejeb, A., Keogh, J. G., & Treiblmaier, H. “Leveraging the internet of things and blockchain technology in supply chain management.”[J] Future Internet, 2020, 12(7), 118.

[20] Maksimović, M., Vujović, V., Davidović, N., Milošević, V., & Perišić, B.“Blockchain technology, bitcoin, and Ethereum: A brief overview.”[A] In 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)[C],2018.

[21] Abeyratne, S. A., & Monfared, R. P. “Blockchain ready manufacturing supply chain using distributed ledger.”[J] International Journal of Research in Engineering and Technology, 2016, 5(9), 1-10.

[22] Wamba, S. F., Queiroz, M. M., & Machado, M. C. “Addressing the data sharing and privacy paradox in the internet of things: A systematic literature review, synthesis, and future research agenda.”[J] Computers in Industry, 2020, 122, 103306.

[23] Dehao Tao, Zhen Yang, “UEPF：A blockchain based Uniform Encoding and Parsing Framework in multi-cloud environments”[J], KSII Transactions on Internet and Information Systems, 2021, 15(8), 2849-2864

# Acknowledgement

I would like to express my heartfelt gratitude to the following individuals and organizations who have supported and assisted me throughout the completion of this thesis.

First and foremost, I would like to thank my supervisor, Professor Song, for his invaluable guidance and advice throughout the entire research process. His expertise and rigorous approach have had a profound impact on me, helping me improve the quality and depth of my research.

I would also like to thank all the members of the laboratory, for their valuable support and collaboration. Our discussions and exchange of ideas, as well as my collective efforts to overcome research challenges, have played a crucial role in shaping my thesis.

Furthermore, I want to express my gratitude to my family and friends for their unwavering support and encouragement in my academic pursuits. Their presence and understanding have provided me with warmth and emotional support, enabling me to fully dedicate myself to my research work.

Finally, I would like to express my gratitude to all the individuals who have guided and encouraged me throughout my academic journey. Their motivation and support have enabled me to grow and progress continually.

I extend my sincerest thanks to all those who have helped me along the way!