Blockchain solution to enhance supply chain data provenance

S.C.S Sirisooriya
2021

Blockchain solution to enhance supply chain data provenance

A dissertation submitted for the Degree of Master of Computer Science

S.C.S Sirisooriya
University of Colombo School of Computing
2021

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

Student Name: S.C.S Sirisooriya

Registration Number: 2017/MCS/078

Index Number: 17440781

11/30/2021

Signature of the Student & Date

This is to certify that this thesis is based on the work of Ms. S.C.S Sirisooriya under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by,

Supervisor Name: Dr. M.D.J.S Goonetillake

Signature of the Supervisor & Date 30/11/2021

ACKNOWLEDGEMENTS

First, I would like to thank University of Colombo School of Computing for the opportunity given me to study for the Master of Computer Science program.

Foremost I wish to thank Dr. Jeewani Goonetillake, my research project supervisor, for allowing me to expand my knowledge in a cutting-edge technology, and for encouraging me throughout my studies. Her advice was invaluable throughout the research.

In addition, I would like to thank all the other staff members of University of Colombo, especially in the Computing Department, for helping me during my studies.

And I would like to express my heartfelt gratitude to my parents and family for always encouraging me and being there for me.

Finally, I would like to take this opportunity to thank everyone who helped me throughout the course of this MCS project.

ABSTRACT

Consumers are demanding more visibility to the products they purchase. In the past decade, importance of the traceability of products from source to retailer has grown high. Supply chains have become more complicated in structure, difficult in task, and dealing with wide range of stakeholders. Many organizations do not have complete view of the entire supply chain. They only have a system to control their own operations. Due to these limitations supply chain transparency become a complex problem to solve.

Most of the currently available supply chain management systems are built using traditional client server architecture. Information in the traditional systems is controlled by one single authority. Because of that, information can be mutated and there is no assurance of the correctness of available data. Furthermore, traditional systems are not transparent, and consumers don't have the ability to access the information in those systems. This lack of transparency causes numerous issues and difficulties in the supply chain mechanisms in respect to security, traceability, authentication, and verification. As a result, the connection between supply chain entities is lost, making it difficult to trace the product's provenance details.

The main goal of this thesis is to visualize the product's provenance information to the consumers. With the traditional systems visualizing provenance details is hard to achieve due to the limitations mentioned above. In order to overcome these limitations, blockchain based supply chain system has been proposed which developed using smart contracts. The system is built on top of Ethereum Blockchain, a distributed public blockchain. By implementing the system using blockchain, the traceability and the ability to share information among each entity in the supply chain will be made easier and trustworthy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1 Overview

Before reaching the consumer, a product goes through several stages, starting with the sourcing of raw materials through production process connecting complex network of suppliers and distributors. Many suppliers are involved in the process of design, manufacturing, delivery, and sales. But most of the time this journey remains unseen to the consumer. It is not clear to the consumer how value is getting added to a product when move through the supply chain.

There is a growing necessity for more visibility and transparency in the supply chain. Most of the time consumers are unaware of exact product's origin, what happens to it and where it moves over time. Because of that, consumers are eager to know the exact journey of the product they buy. A system is required, which can maintain the product's journey as various parties involved in the production of raw materials, parts, and products, which are then processed and eventually distributed as end products to the consumers.

1.2 Problem Domain

Beside the shops that consumers go to buy products, there is a network of different entities connected with each other, that work together to deliver these products to the consumers. This connected network is called Supply Chain. These connected entities in the supply chain are responsible for extracting raw materials and transform them into finished products, transporting finished products, and delivering them to the consumers.

Supply chain is getting complex day by day. Due to this complexity, supply chain is suffering from three main issues namely, process optimization, data visibility and demand management [31]. Consumers are demanding more visibility in the supply chain, as most of the time consumers are unaware about how products are made, what are the stages product is going through before reaching the consumer. As consumers become increasingly

skeptical about the supply chain transparency, there is a high demand for provenance-based supply chain management systems, which can improve the consumers' provenance knowledge. Supply chain transparency will improve the consumers' provenance knowledge which will then help to increase the consumers' trust by ensuring the origin, authenticity, custody, and integrity of products.

1.3 Problem

Supply chain is a network between different entities who contribute to produce and distribute a specific product or service to the consumers. The entities involved in the supply chain includes producers, vendors, warehouses, transportation companies, distribution centers and retailers.
For example, in the clothing industry, the supply chain network is made by connecting raw material sources from different geographical areas, factories that create final products using these raw materials, distribution centers that deliver these clothes to consumers.

When product manufacturers buy raw materials from primary suppliers, those raw materials can follow multiple paths and destinations before reaching the primary supplier. In the clothing supply chain, factories request fabrics, buttons which are needed to create the final product (garments) from primary suppliers. Fabric suppliers will need to source cotton and dyes from other suppliers to manufacture the fabric. Like so, raw materials transform to different forms and end up as a final product. Since the quality is the main concern when creating the products (i.e., garments in the above case), it is very important to track and view the destination data like **a)** what happens to the raw materials **b)** where it came from **c)** who brought them to the stakeholders in these destinations.

Provenance can be defined as information about the creation, modifications or influences that occurs related to an artifact [32]. Data provenance is the series of changes arise related to any piece of data. Data provenance systems keep track of, from where data originates and moves to, how data has been changed, and who makes changes to the data in the meantime. In supply chain perspective, tracking **a)** what happens to the raw materials **b)** where it came from and **c)** who brought them, throughout the process can be defined as data provenance.

Traceability and transparency are the most essential factors in supply chain management [33]. To increase the visibility in the supply chain, revealing the provenance of a product can be a great choice. But the reliability of information is mostly brought into question as the most commonly used supply chain management systems are managed by the centralized database driven system, which causes following identified problems:

- **Purposely data manipulation (Data tampering).** These can be caused by hackers using attacks or administrators purposefully.
- **Accidently delete or destroy data due to human mistakes, improper data modifications or data center failures (single point of failures due to centralized nature).** Often the whole database or application can be destroyed due to natural disasters, power failures or unexpected hardware failures.
- **Traceability and transparency.** In the traditional database driven systems, traceability and transparency are crucial issues since the users do not have any control over data.

By overcoming above mentioned problems and increasing the traceability and transparency of the supply chain, this research project would address the problem of "**How to manage data provenance with traceability and transparency to represent the product journey in the supply chain process**".

1.4 Exact Computer Science Problem

The most important computer science problem is how to provide a software-based solution to visualize provenance data in the supply chain. Data provenance includes from where data has been originated, how data has been changed and to where it moves with time. Data Provenance enable the visibility of a system while giving the ability to track errors back to the root cause in the data analytics process. In order to visualize the provenance data in supply chain, we need a system where traceability and transparency is possible. With the traditional client server architecture traceability and transparency are hard to achieve. Proper solution is needed to overcome the above problem.

As mentioned in the previous section, there are several problems that provenance system will face if we are to use traditional client server architecture. They are,

- Single point of failure
- Data tampering
- Data lose
- Transparency and traceability

In order to overcome the above-mentioned problems in the traditional supply chain, decentralized solution is needed to eliminate the central authority. This is another computer science problem that will be address within this research project.

1.5 Motivation

Consumers have larger interest to know the provenance information of products they consume [37]. As mentioned in the previous sections, there is no proper way to obtain the provenance information of a product in the traditional supply chain. Motivation for this project is to propose a solution which will be a transparent and traceable decentralized system to visualize the provenance information in the supply chain.

1.6 Objectives

The main objective of this research project is to design and create a transparent and traceable model to enhance data provenance in supply chain using suitable blockchain technology. This is because of:

a) Data stored in the blockchain are immutable and can only be added after the verification using consensus algorithms. Therefore, data cannot be tampered.

b) Data will be encrypted and stored in blocks using smart contracts. Smart contracts are able to grant specific user access levels for the execution.

This model will eliminate the problems in centralized system as follows:

- **Purposely data manipulation (Improper data manipulation)**

This is impossible as persisting data in blockchain is immutable. Therefore data cannot be manipulated upon save.

- **Accidently deleting or destroying data (Single point of failures due to centralized nature)**

Single point of failures is not possible as blockchain technology is decentralized. Therefore, data is persisted in multiple copies by default in multiple nodes.

- **Traceability and transparency**

Blockchain technology will enable the traceability and transparency in the system as provenance information will be tracked in the system.

Through this research following concern to address:

- How far blockchain can be used as a suitable technology to present/reflect data provenance in supply chain.

1.7 Scope

The focus of the project is to design and create a model to track the transformation process of the raw materials during the supply chain life cycle. This will give the ability to trace the path of a product back to its origins. This model will be distributed which gives visibility to the data whoever uses this model in any stage of the supply chain.

The proposed solution contains three main entity types. Those are supply chain users (i.e., suppliers, manufacturers), smart contracts and the blockchain network. In order to increase the supply chain visibility, changes that occur during the product life cycle are tracked and visualized in the proposed system. As traceability and transparency are two critical components in the supply chain, they are combined in the system. Information in the system is permanent, immutable and cannot be changed by anyone in the supply chain network as blockchain technology is used to store the information. Because of that data accuracy is guaranteed. Smart contracts, which are automated programs that runs on top of blockchain, are incorporated into the proposed system that will be executed on each node on the

network to store and retrieve information. As a result, participants of the supply chain network have direct interaction with one another.

2. Literature Review

Nowadays most of the consumers are eager to know the provenance information of products they purchase and consume. Knowing provenance information of a product can reduce the consumers perceive risk. This chapter present important literature on supply chain management, data provenance and related research that use blockchain technology to achieve data provenance in tracking systems.

2.1 Supply Chain Management

Multiple definitions have been proposed in terms of supply chain and supply chain management. According to J. T. Mentzer et. al [1], the supply chain is defined as numerous stakeholders directly involved in the process of moving products, finances, services, and information from a source to a destination. J. T. Mentzer et. al [1] also state that the supply chain management as the key operations that are necessary for managing the process within the supply chain, for the purpose of enhance the entire supply chain and the characteristic of individual components.

As traceability and transparency are the most important factors in supply chain management, during the past years, significant emphasis has been given to those factors. Since the supply chain consumers are willingly following each stage of the production process, traceability become an important factor. Specially, in the food supply chain, traceability is a key factor as knowing the provenance in food supply chain is critical [4].

Traceability is defined as [2] the ability to track the information of products and parts which are in the supply chain, starting with the sourcing of raw materials, throughout the production process, until the product is delivered to the final consumer. Traceability plays an important role in demonstrating the transparency of the supply chain by using verifiable information [3].

L. U. Opara [3] explains the significance of traceability. L. U. Opara [3] has identified six critical elements of traceability, which include:

- Tracking physical location of the product.

- Tracking the manufacturing process of the product.
- Tracking genetic composition of the product.
- Tracking origin and type of food product.
- Tracking origin of the disease from raw product.
- Tracking individual measurement results of the product.

Above mentioned elements can be combined to provide consumers with a comprehensive view of a product's lifecycle.

There is a need for big companies to adapt the traceability and transparency into their supply chains in order to meet consumers demands.

There are three types of traceability systems: centralized, linear, and distributed.

- In the centralized systems, in order to collect data, shared databases are used.
- In the linear systems, information of specific products is being recorded by each partner and transfer from one partner to another.
- In the distributed systems, the entities in the supply chain establish a traceability system to exchange traceability data [5]

C. N. Verdouw et.al [6], proposed a design with distributed architecture, which will be used across the food supply chain in order to track, govern, organize, and optimize processes using the Internet based on virtual objects. W. Wang et.al [7] proposed a solution to monitor the products in wire bond station using RFID. The proposed system has the ability to track and monitor work process delays, qualification lots, semi-product rejection, and wire bond machine maintenance in real time. D. Folinas et.al [8] state that the depend on the ability to trace products, efficiency of a traceability system increase. As stated by the author [8], in order to have an efficient traceable system, constant monitoring is essential from primary production until final disposal by the consumer. C. Shanahan et.al [9] introduce an RFID-based framework in an attempt to increase consumer confidence in beef products. To reduce the risk of dishonesty, the system recognizes all components of beef traceability, including individual cattle recognition and biometric identifiers for confirmation of cattle authenticity from farm to slaughter.

The above-mentioned research work did not specify how they present information to the various parties involved in the chain. Finding the provenance information is not possible using the approaches described above because traditional client-server architecture is used to store the information, resulting in data tampering issues.

According to [35], 75 percent of supply chains in business organizations have reported lack of visibility. Most businesses have less information on the parties involved in the manufacturing process, which contributes to the network's inefficiency. Visibility in the supply chain provides speed, reliability, and flexibility for gaining a competitive advantage through well-controlled and managed supply chain functions [36].

2.2 Data Provenance

The provenance of a data item includes information about the processes and source data items that lead to its creation and current representation.

There are two fundamental perspectives on provenance. The first defines a data item's provenance as the process that led to its creation, while the second focuses on the source data from which the data item is derived [10]. Buneman et al. distinguish between Where Provenance and Why Provenance in [11]. Where Provenance represent the concrete origin of results, whereas Why Provenance represent all source data items that contributed to the creation of a result data item.

Woodruff et al. [12] discussed the provenance for relational databases with respect to visualization. Authors [12] used functions to transform data from one attribute domain to another, then reversed those functions to track the provenance. Groth et al. [13] present another approach to provenance management in which user actions are recorded and visualized as a DAG (directed acyclic graph). Users can return to the previous state of the system by traversing the graph.

To improve privacy, Asghar et al. [14] suggested a cloud-based secure data provenance solution that employs two folder encryption method. To ensure the authenticity and privacy of provenance data, SPROVE [15] employs encryption and digital signatures. SPROVE,

on the other hand, lacks the ability to query provenance data. Progger [16] is a kernel-level logging tool capable of providing evidence of log tampering while violating user privacy. There are also initiatives that use provenance data for cloud environment management, such as discovering usage patterns for cloud resources, popularizing resource reuse, and fault management [17].

2.3 Blockchain

Blockchain was first introduced in 2009 as a distributed peer to peer network underlying Bitcoin transactions, which is a cryptocurrency-based protocol for the trade of digital currency. In 2008, a white paper Bitcoin: A Peer-to-Peer Electronic Cash System, written by Satoshi Nakamoto [18] introduced new way of transferring funds in the form of Bitcoin. He has proposed a specific data structure called blocks which is used to transfer and store data in distributed ledgers. Since bitcoin is the peer-to-peer version of electronic cash, online payments can be made without the involvement of a third party.

Since the financial transactions could be completed without the use of a middleman, the technology underling the Bitcoin network is captivated by the researchers. Following the introduction of Blockchain 1.0 by Satoshi Nakamoto for cryptocurrency exchange, Blockchain 2.0 primarily name as Ethereum was introduced to overcome the challenges that realize in the previous version namely scalability, usability, performance, and limited coding execution ability. Ethereum introduced Smart Contracts which added a computation logic layer to the Blockchain. In Blockchain 3.0, some of the key concepts introduced in the Blockchain 2.0 got matured which enable the development of Ethereum blockchain applications usually referred as decentralized application, that combined front-end and contracts. Even though Ethereum address the usability issues, it didn’t address the scalability and performance issues which led to evolve the Blockchain technology more based on the alternate coins.

The blockchain uses distributed ledger technology to record and store transactions between parties in multiple locations. The distributed ledger technology is a method for recording, sharing, and synchronizing data across a decentralized network of computer systems with varying degrees of ledger control. In the decentralized systems:

- Users have control over data.
- Because the system is decentralized and relies on multiple nodes, there is no single point of failure.
- Malicious attacks are eliminated because the system is not controlled by a single party.

2.3.1 Public Blockchain

In public blockchain, any participant can connect to the network to validate transactions and to access the information on blockchain, as there is no central entity to control the network [19]. One disadvantage of the public blockchain is that it requires more time and energy to synchronize the nodes in the blockchain as the chain grows in size [20]. The network enables unrecognized network users to be inspired and improve the ledger's accuracy [21]. In [22], Linn et al explains the importance of public blockchain for health data. Because anyone can take part in the system and store their health data, public blockchain has been used as an access control manager for health records in their work.

2.3.2 Private Blockchain

Since the network is managed by a single organization, a private blockchain considered as a centralized network. In private blockchain, nodes require permission to access the network and add data to ledgers. Only nodes from the organization that controls the network are permitted to participate in the consensus process. [23]. This capability allows the network to be more easily integrated into regulatory frameworks and institutional arrangements. Private blockchains are also good at dealing with problems such as identity verification and data privacy [21]. However, because permissions are centralized, the decentralized concept is sacrificed [20].

2.4 Data provenance and Blockchain

In the traditional supply chain systems, the risk of data manipulation is significantly high. With the use of blockchain technology in the supply chain, drawbacks in the traditional systems can be overcome as the technology is decentralized and secured. All the involved

parties in the supply chain interact with each other by using public and private keys, which makes the production process paperless. Blockchain technology will revolutionize the supply chain industry as a whole.

By adopting the blockchain technology to the businesses, industries can gain lot of advantages. Those advantages include:

- All the parties can interact with each other real-time which will remove the need for exchange the documents.
- Since the processes are automated, the occurrence of errors are minimized which will increase the accuracy.
- By using the required keys, all the parties can access the information. Since the history of the transactions are available on ledger, the system will be fully transparent.
- As the data in the blockchain are immutable users cannot change the data in the system. This feature prevents the system from tampering. Furthermore, as all the information in the system is encrypted, the system is more secure [24].

There has been some research done to find out how blockchain can be used to leverage the supply chain industry. Below is a detail view of that.

In [25] Malik et al. propose a framework which support provenance in supply chain. The proposed framework [25] includes Food Supply Chain (FSC) entities that collectively governing a permissioned blockchain. Propose framework [25] provides a comprehensive platform for Food Supply Chain entities and administrative bodies to collaborate in order to achieve end to end traceability and monitoring of supply chain events corresponding to product quality. Since the proposed framework makes use of a permissioned blockchain and govern by a consortium of key Food Supply Chain entities, access is limited to authorized parties only.

Kim et al. [26] studied the blockchain to identify how technology can be used to track the provenance of physical goods. On the Ethereum blockchain, a proof-of-concept is built that employs TOVE traceability ontology axioms expressed in smart contracts to carry out source tracking and traceability. However, the study makes no mention of how entities in the supply chain can view provenance data.

Ramachandran et al. [27] proposed a solution for the unavailability of a framework to verify provenance records automatically. Proposed solution by the authors [27] make use of smart contracts and open provenance model (OPM), to capture and verify the provenance through the use of cloud-based verification scripts. Smart Provenance mode used automated verification scripts and automatically discarded invalid changes. The proposed solution focus on verifying the scientific provenance data and hasn't focus on presenting provenance data to the users.

Moeniralam conducted research in [28] to design a Blockchain-based Data Production Traceability System for Sentinel-2 satellite data, in order to track and verify the changes made to the original data set. As stated in the paper, the goal is to design a system that addresses data lineage issues in satellite data by storing the various levels of data, as well as the production environment in which they are processed, and a complete record of all steps taken. Blockchain technology was used to securely store satellite data and to allow the user to reconstruct, analyze, and verify each step. The research hasn't focus on presenting provenance data to the users and the model can be used only to trace the provenance of satellite data.

Rodrigo et al. [29] propose a graph database architecture for data provenance in bioinformatics workflows. In [29], Rodrigo et al present a ProvBio, an architecture that can automatically perform the data provenance of scientific experiments in bioinformatics using the PROV-DM provenance data model and a graph database. The architecture is capable of performing automatic provenance types prospectively, retrospectively, and with user-defined data. As a result, the architecture stores and captures information obtained during the execution of the data generation processes, such as features and versions of the programs used, with user-defined data information. For storing data provenance, a graph model based on the PROV-DM model was proposed. The PROV-DM can be represented as a graph, allowing for more natural modeling, as well as more natural query expression and the implementation of efficient algorithms to perform specific operations. However, because the proposed solution is not decentralized, it must rely on centralized authority to ensure its security. Another disadvantage is that the proposed solution [29] can only be used to store Bioinformatics data.

Marten et al. proposed a generic blockchain-based data provenance framework for the IoT in [30], which can be applied to a wide range of use cases. The benefits of a generic framework include the ease with which new use cases can adopt provenance concepts and the interoperability of applications that use the framework. Create and deploy an IoT data provenance framework comprised of smart contracts based on a generic data model to provide provenance functionality for a wide range of IoT use cases. However, the research doesn't focus on presenting the provenance data to the users.

Tian [38] presents a blockchain based architecture for the agri-food supply chain. The high-level conceptual design is based on public blockchain and IoT sensor data. The authors extend their work [39] to address scalability by leveraging distributed databases such as BigChain. However, their approach uses blockchain as a blackbox and thus they do not elaborate on implementation details or consider challenges associated with incorporating blockchain in the supply chain context such as the type of blockchain used, availability and traceability.

A blockchain based wine traceability system is proposed in [40]. In order to verify the origin and purchase history of wine bottles, reliable information is collected using blockchain. Multichain, an open platform for implementing private blockchain solutions, is used to implement the proposed framework. Their design proposes using specific supply chain entities as miners, such as wine producers and bulk distributors, who would be responsible for verifying blocks. Information such as origin, production, and purchase history is available for the other entities in the supply chain, once verified blocks are added to the ledger and make them publicly available. Their approach, however, does not address scalability. Furthermore, it is unclear how wine ingredients can be traced apart from the provenance of each individual bottle.

To prevent counterfeiting, K. Toyoda et al. [41] propose storing ownership information about manufactured goods on the blockchain as the product move through the distribution chain. As the products leave the manufacturer, RFID tags are attached to them, and the information in the RFID tag is updated when the product is traded among different entities, storing ownership details on each transit. When a product arrives at a retailer, a consumer may reject it if the product history lacks ownership information of the seller from whom

the consumer is purchasing. While it is an effective method for tracing products from manufacturer to retailer, it does not allow for the acquisition of information prior to manufacturing, such as the originator of raw materials.

2.5 Summary

Most of the research that mentioned above used blockchain to store the provenance data, because it's a decentralized network and data stored in the blockchain can't be changed. Some of the identified research use graph databases to store provenance data. When it comes to querying provenance data graph databases will be more convenient than blockchain. However, data stored in the blockchain is more secure.

In the above mentioned research following research gaps exist.

1. Ability for supply chain entities to read and view provenance information. (Forward and backward tracking)
2. Ability to track the entire supply chain information using a single system.
3. Ability to access the system without the permission from third party organization.
4. Using traditional client server architecture which will lead to data tampering, single point of failure and traceability and transparency issues.
5. Proposed models can be only used in one industry (e.g., in food supply chain). There is no general solution which can use in every industry.

	Ability for supply chain entities to read and view provenance information	**Ability to track the entire supply chain information using a single system**	**Accessibility is limited to the authorized parties**	**Using traditional client server architecture**	**Proposed models can be only used in one industry**
Malik et al [25]	The information can only be read and viewed by those who have the access rights.	This is possible with the proposed system.	The proposed framework makes use of permissioned blockchain, which limits accessibility.	As the blockchain technology used in the proposed system drawbacks in client server architecture is eliminated.	The proposed framework is a generalized framework.
Kim et al. [26]	This is not emphasized in the study.	This is not emphasized in the study.	The proposed design makes use of a public blockchain. As a result, anyone can access the data.	As the blockchain technology used in the proposed system drawbacks in client server architecture is eliminated.	This is not emphasized in the study.
Ramachandran et al. [27]	This is not emphasized in the study.	This is not highlighted in the study because the study focuses on tracking the provenance of scientific data.	The study has limited the accessibility for provenance data.	As the blockchain technology used in the proposed system drawbacks in client server architecture is eliminated.	The proposed solution is only applicable for tracking the provenance of scientific data.
Moeniralam [28]	This is not emphasized in the study.	Because the study focuses on satellite data provenance,	The study has limited the accessibility for	As the blockchain technology used in the proposed	The proposed solution is only applicable

		this is not emphasized.	provenance data.	system drawbacks in client server architecture is eliminated.	for tracking the provenance of satellite data.
Rodrigo et al. [29]	This is not emphasized in the study.	The study has focus on provenance in bioinformatics data	This is not emphasized in the study.	The proposed solution uses graph databases. So, the data tampering issues are possible. Relies on central authority.	The proposed solution is only applicable for tracking the provenance of bioinformatic data
Marten et al. [30]	This is not emphasized in the study.	This is possible with the proposed system.	This is not emphasized in the study.	The study uses decentralized solution that eliminates the drawbacks in traditional client server architecture.	The proposed framework is a generalized framework.
Tian [38]	This is not emphasized in the study.	This is not emphasized in the study.	This is not emphasized in the study.	This is not emphasized in the study.	This is not emphasized in the study.
K. Biswas et al. [40]	This is partially possible with the proposed framework.	The study doesn't focus on tracing the entire supply chain.	The proposed framework uses private blockchain. Hence the access is limited to authorized parties.	As the blockchain technology used in the proposed system drawbacks in client server architecture is eliminated.	The proposed solution can only be used in wine supply chain.
K. Toyoda et al. [41]	This is possible with the proposed	The study doesn't focus on tracing the	The proposed design makes use	As the blockchain technology used in the	The proposed framework is a

		entire supply chain.	of a public blockchain. As a result, anyone can access the data.	proposed system drawbacks in client server architecture is eliminated.	generalized framework.
S. Abeyratne et al. [42]	This is not emphasized in the study.	This is hypotheticall y emphasized in the study.	This is not emphasized in the study.	The study emphasized of using a decentralize d architecture.	This is not emphasized in the study.
G. Baralla et al. [43]	This is not emphasized in the study.	This is not emphasized in the study.	The proposed design makes use of a public blockchain. As a result, anyone can access the data.	The study emphasized of using a decentralize d architecture.	This is not emphasized in the study.

Table 2.1: Summary of the literature

3. Methodology

3.1 Problem Analysis

Before consumer can consume, a product is traveling different stages in supply chain.

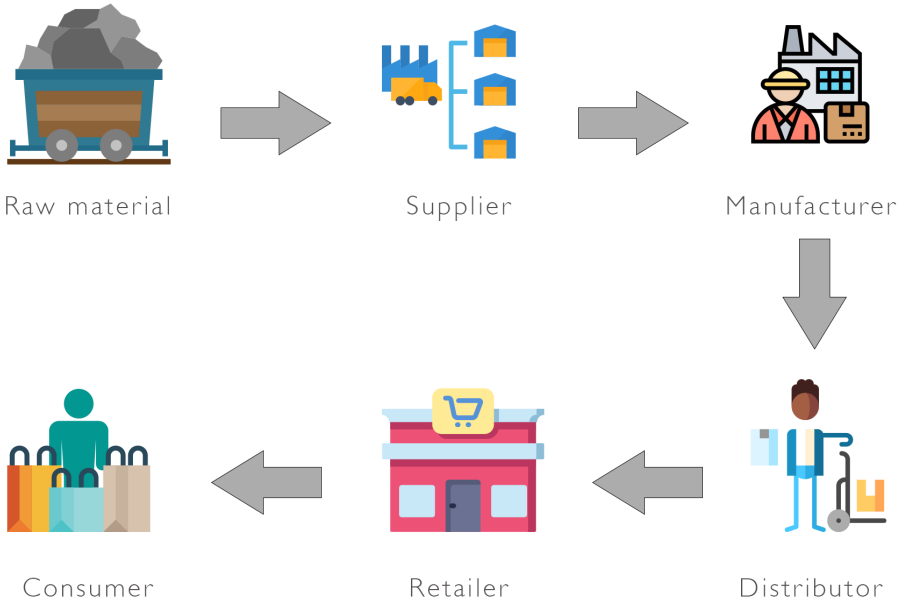


Figure 3.1:The process of supply chain

Figure 3.1 shows the life cycle of a product as it moves through different stages in the supply chain. Primary producers derived the raw materials and sold to wholesale suppliers. The manufacturers obtain the primary ingredients from the suppliers and process them to manufacture the product. Products are then passed on to the distributor and shipped to the retailer where they are eventually purchased by the consumers.

Since products are shipped all over the world which involves physical movements of the products through supply chain, in order to protect the health and safety of the consumer ensuring product quality and safety is essential.

It is often unclear to the consumers how value is added to the product along the supply chain. Most of the time consumers are relying on the country-of-origin labels, brands,

certifications when purchasing a product. However, due to increase number of product recalls in recent time which involves sourcing and supplying of products, consumers are questioning the reliability of these matrices. For an example, River Island, famous clothing brand has recalled number of garments over fears they have been made using dangerous levels of harmful chemicals [34]. Due to these incidents, consumers are taking unprecedent interest in the way product is produced, processed and handle, and want to know the origin of the product or whether the product is handled appropriately.

Currently, supply chains are becoming complex in structure, difficult in terms of task, and diverse in terms of stakeholders. Many organizations do not have the complete view of the entire supply chain. Due to this low transparency many problems and difficulties arise in the supply chain mechanism in terms of security, traceability, authentication, and verification.

Unfortunately, finding provenance information of a product is not straightforward due to the usage of different repositories in each stage of supply chain and complexity of data aggregation. Moreover, a product moves through a complex supply chain involving many entities with distinct operational practices and procedures.

The current traceability solutions suffer from issues such as scattering of information across multiple places and susceptibility in recording erroneous data and thus are often unable to produce reliable information of a product from its raw material extraction and production phase to the final consumer.

The process of collecting data from different repositories to build a product story and ensuring the integrity of data are key challenges, especially at the manufacturer stage where various raw materials are collected to form one product. There is a possibility that data inaccuracies may arise either because of errors or malicious intent. Although there are traceability systems based on centralized repositories, since the organizations are working independently from each other, the process of tracing provenance information from these independent sources are tiresome and filled with errors.

New supply chain traceability systems are needed which provide a detailed product story, manage the data in a decentralized manner and make sure the confidentiality of the data. Data provenance systems are the way to achieve traceability in supply chain. Data provenance systems tracks the origin and evolution of data such as data creation, data modifications, who initiated them, and when and how those modifications occur. It is necessary to store the information in a tamper-proof and replicable way, in order to trust the information provided by the provenance system.

3.2 Proposing model/design

The main objective of this research project is to give opportunity to visualize provenance information of a product to the consumer which is registered in a supply chain. In order to achieve this, main steps of the supply chain process have been identified and implemented using blockchain technology to record the provenance information and then reveal the provenance of a product to the consumer.

In this research project, blockchain technology has been chosen over traditional client-server architecture in order to develop the provenance system. Main reason to use blockchain technology is the issues that exist in traditional systems. Those are as follows.

- Central point of failure as system is relying on one single node.
- Purposely data manipulation by hackers or administrators.
- Performance is limited as central server should handle all incoming requests.
- Traceability and transparency issues since the users do not have any control over data.

Supply chain has a complicated structure. In order to simplify the domain, simple clothing manufacturing supply chain has been chosen.

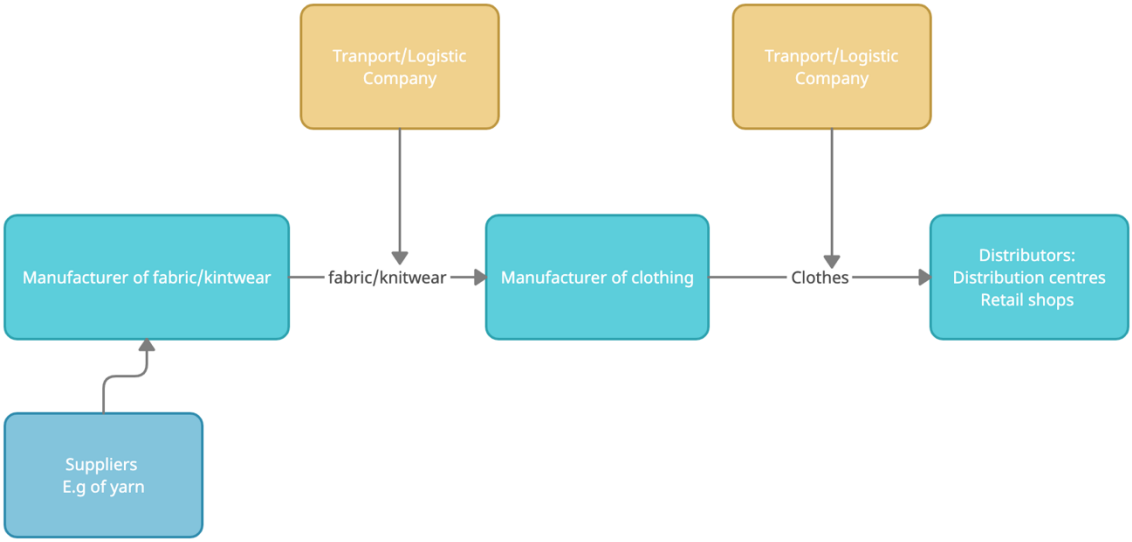


Figure 3.2:Simple Clothing Supply Chain

As shown in Figure 3.2, three major users of the system have been identified: Supplier, Manufacturer, and Distributor. Before using the system, each of these users must register on the same blockchain. Users will receive an address after registering on the blockchain, which will be used to uniquely identify users in the system.

Suppliers are responsible for producing primary ingredients such as yarn, cotton, wool, and silk for the manufacturing process. The following attributes will be captured in the supplier stage to maintain provenance details.

- Supplier details (company/individual name, contact details)
- Batch number
- Raw material type (Cotton, Wool, Yarn, Silk, Leather)
- Quantity
- Description
- Source location
- Date

Once the supplier enters the above details to the system, a hash value is generated to uniquely identify the entered record. Furthermore, the supplier states the ownership of the raw materials to himself.

Moving up the supply chain, manufacturers purchase ingredients from the primary suppliers to manufacture products. Primary suppliers will transfer the ownership of the primary ingredients to manufacturers. Following attributes must be captured at the manufacturer stage in order to maintain provenance details.

- Manufacture details (company/individual name, contact details)
- Serial number
- Batch number
- Quantity
- Description
- Manufacturer location
- Product type (Fabric, Garment)
- Date

With enough raw materials, manufacturer can eventually begin the production process. Once the production is done, manufacturer register the product in the system and set the ownership of the product to himself. Hash value is generated to uniquely identify the product in the system.

After the production of the garments is done, distributors can buy garments from the manufacturer. Ownership of the product is transfer to the distributor by manufacturer. Users will be able to see the provenance details at each stage of the production process.

The blockchain stores every transaction that involved in the production process, so the distributor or any other party can track everything that a specific item has gone through. Using the events generated by every transaction we can achieve this. Figure 3.3 shows the sequence diagram for the above mentioned scenario.

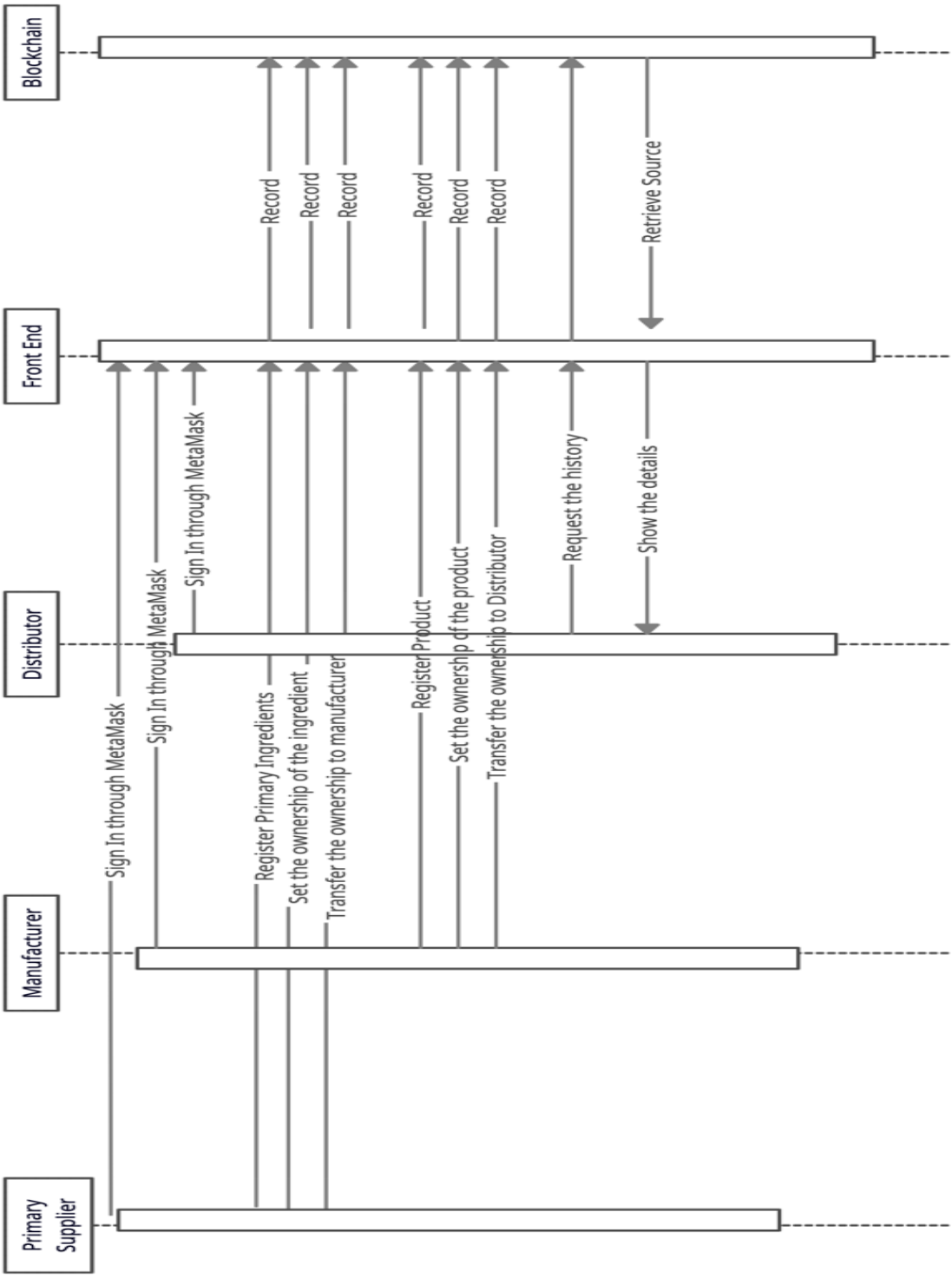


Figure 3.3:Sequence Diagram

In order to provide secure decentralized provenance system, blockchain based system is introduced. Main reason to use the blockchain technology is to take the advantages of its features like transparency, traceability, and immutability. The provenance system architecture is based on Ethereum blockchain in order to overcome the importance of an administration body to control the system. Even though there is no administration body, since the data manipulation is not possible with the blockchain technology, consumers of the system can trust the information in the system.

For the development purpose Ethereum test network has been used. The test network, same as the main network can have one or more nodes that are connected directly to each other. Each node is a machine, running an Ethereum client, which has the complete copy of the blockchain. It is same as a database which store detail of every transaction that has ever occur.

Figure 3.4 shows the proposed system architecture.

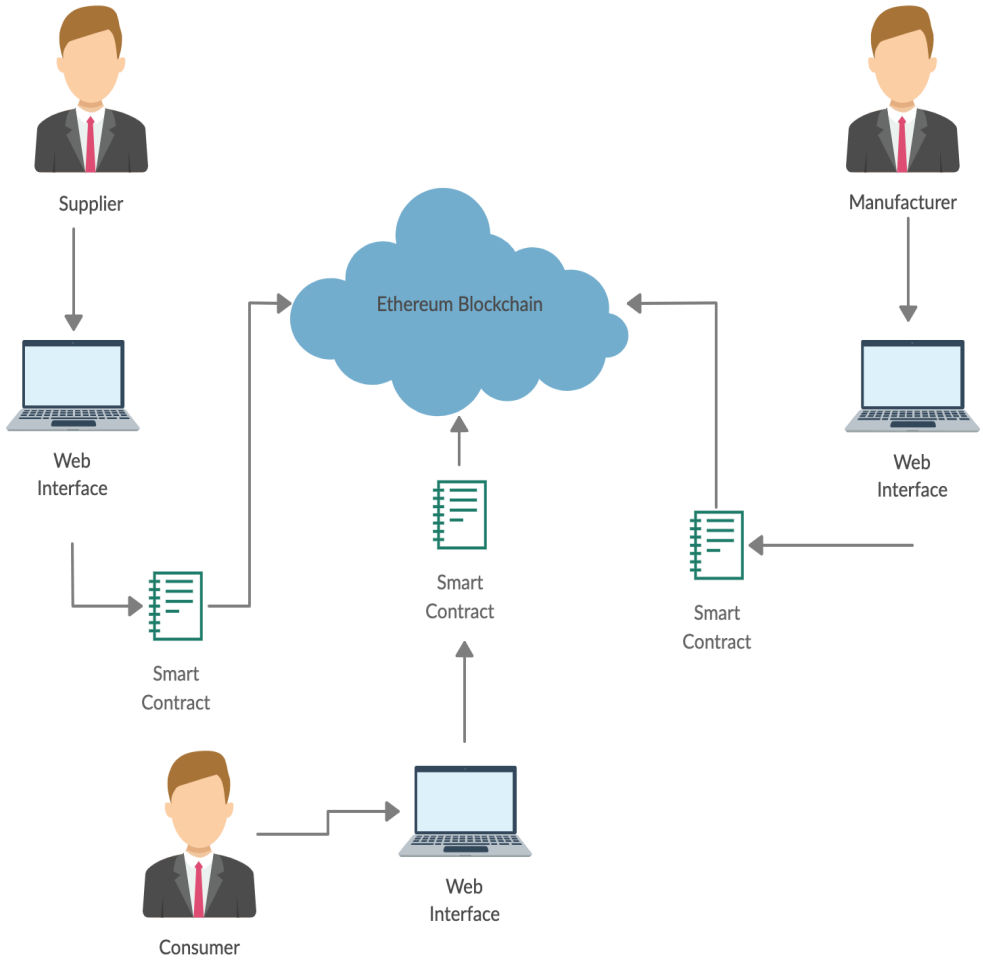


Figure 3.4:System Architecture

3.2.1 System Components

As shown in Figure 3.4, main components of the system architecture described below.

(1) Blockchain

Blockchain is an important technology which provide data transparency and integrity with its prominent features which include decentralization, immutability, and auditability. It is a distributed ledger technology which has the ability to build powerful supply chain traceability system by tracking information that is tamper proof, secure and highly available. Blockchain used as the main component in the system in order to record the provenance details of the products permanently and immutably. With the use of blockchain technology, the issues that exist with the traditional systems such as purposely data

manipulation can be eliminated. Since the blockchain systems are decentralized by nature, all involving entities can have one to one communication. Apart from that, since the records in the blockchain are immutable no one can alter or delete transaction data which helps to prevent data manipulation.

The primary purpose of the solution is to reveal the information of the products which travel through each stage of supply chain. In order to achieve this, keeping the data without transforming is essential. Once the data is added to the system nobody should be able to manipulate them. There are several blockchain platforms available, such as Hyperledger and Ethereum. Hyperledger is a general purpose permissioned blockchain-based distributed ledger that improve many aspects of performance and reliability.

Ethereum is the most widely used public blockchain platform, which provide decentralized virtual machines (VMs) to run smart contracts. Solidity is the built-in general-purpose scripting language for Ethereum. Ethereum can be used by anyone to build a blockchain-based decentralized system at a low cost and on a per-contract and per-byte basis. On the other hand, since the Ethereum is a public blockchain, following benefits can be obtained.

- Network security is high as large number of nodes helps to maintain the immutability.
- Unlike permissioned blockchain, there are no restrictions on who can participate on the network.
- Public blockchain is decentralized, while private blockchain is more centralized as it controls by one or more entities.

In addition, all the information about accounts or transactions that have ever occurred on the Ethereum network can be viewed in Etherscan.io using the account address. So, the Ethereum platform selected over Hyperledger.

(2) Smart Contract

Smart Contracts are written in the Solidity scripting language, which is where the system's entire logic is defined. Two smart contracts have been used to handle the system's logic.

- **ProductManagement Smart Contract**

  Following are the responsibilities of the smart contract.

  - Registering primary ingredients to the blockchain.
  - Registering products to the blockchain.
  - Retrieving provenance information of the primary ingredients.
  - Retrieving provenance information of the products.
  - Generate hash to uniquely identify the product or primary ingredients.

- **ChangeOwnership Smart Contract**

  Following are the responsibilities of the smart contract.

  - Adding the ownership of the product and primary ingredients to the owner.
  - Changing the ownership of the items when moving through the supply chain.

(3) Front End

The front end of the system was developed using JavaScript. JavaScript will help to show the content of the interface to the user faster. It is a framework that widely used in front end frameworks. This is because of it is simple and easy to learn, fast and scalable compared to other libraries.

By using the system users will be able to do the following.

- **Registering raw materials:** Allow the users to register raw materials by giving batch number, raw material type, quantity, description, source location
- **Registering products:** Allow the users to register product by giving serial number, batch number, quantity, description, manufacture location
- **Change the Ownership:** Setting and transferring the ownership to the supply chain entities.
- **View History:** Allow consumers of the system to view the history of product.

3.3 Implementation

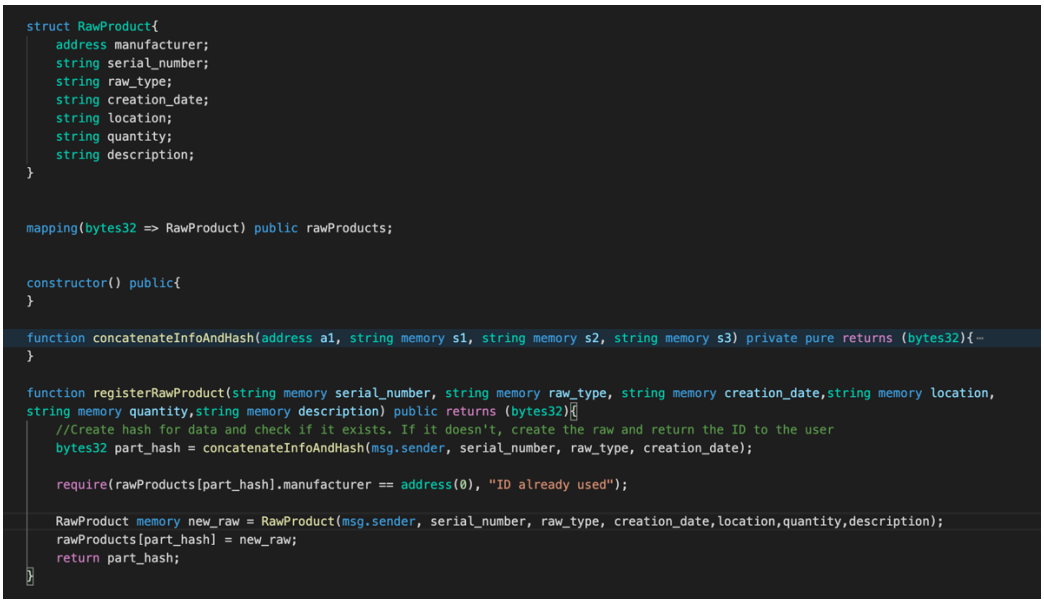
This section describes the implementation details of the system.

3.3.1 Contract Development

Two smart contracts, written using Solidity have been used to handle the system's logic, as described in the design section. Logic of each contract shown below.

(1) ProductManagement Smart Contract

- Registering primary ingredients



```
struct RawProduct{
    address manufacturer;
    string serial_number;
    string raw_type;
    string creation_date;
    string location;
    string quantity;
    string description;
}


mapping(bytes32 => RawProduct) public rawProducts;


constructor() public{
}

function concatenateInfoAndHash(address a1, string memory s1, string memory s2, string memory s3) private pure returns (bytes32){…
}

function registerRawProduct(string memory serial_number, string memory raw_type, string memory creation_date,string memory location,
string memory quantity,string memory description) public returns (bytes32){
    //Create hash for data and check if it exists. If it doesn't, create the raw and return the ID to the user
    bytes32 part_hash = concatenateInfoAndHash(msg.sender, serial_number, raw_type, creation_date);

    require(rawProducts[part_hash].manufacturer == address(0), "ID already used");

    RawProduct memory new_raw = RawProduct(msg.sender, serial_number, raw_type, creation_date,location,quantity,description);
    rawProducts[part_hash] = new_raw;
    return part_hash;
}
```

Figure 3.5:Registering primary ingredients

Mapping in solidity act likes a hash table in any other language to store the data in the form of key-value pair.

In order to uniquely identify a record in the mapping, unique key will be generated. Using a helper function the owner id, raw material type, batch number and supplier date will be concatenate to a bytes variable and calculate a hash. To calculate the hash *keccak256()* in built function will be used which computes the keccak-256 of the bytes input and return bytes32 variable. This hash is the key used when registering and querying the data.

```
function concatenateInfoAndHash(address a1, string memory s1, string memory s2, string memory s3) private pure returns (bytes32){
    //First, get all values as bytes
    bytes20 b_a1 = bytes20(a1);
    bytes memory b_s1 = bytes(s1);
    bytes memory b_s2 = bytes(s2);
    bytes memory b_s3 = bytes(s3);

    //Then calculate and reserve a space for the full string
    string memory s_full = new string(b_a1.length + b_s1.length + b_s2.length + b_s3.length);
    bytes memory b_full = bytes(s_full);
    uint j = 0;
    uint i;
    for(i = 0; i < b_a1.length; i++){
        b_full[j++] = b_a1[i];
    }
    for(i = 0; i < b_s1.length; i++){
        b_full[j++] = b_s1[i];
    }
    for(i = 0; i < b_s2.length; i++){
        b_full[j++] = b_s2[i];
    }
    for(i = 0; i < b_s3.length; i++){
        b_full[j++] = b_s3[i];
    }

    //Hash the result and return
    return keccak256(b_full);
}
```

Figure 3.6:Generate Hash

- Registering products

```
struct Product{
    address manufacturer;
    string serial_number;
    string location;
    string product_type;
    string creation_date;
    string batch;
    string quantity;
    string description;
    bytes32[4] parts;
}

mapping(bytes32 => Product) public products;

function registerProduct(string memory serial_number, string memory manufacturer_location,string memory product_type, string memory creation_date,
    string memory batch, string memory quantity, string memory description, bytes32[4] memory part_array) public returns (bytes32){
    //Check if all the parts exist, hash values and add to product mapping.
    uint i;
    for(i = 0;i < part_array.length; i++){
        require(parts[part_array[i]].manufacturer != address(0), "Inexistent part used on product");
    }

    //Create hash for data and check if exists. If it doesn't, create the part and return the ID to the user
    bytes32 product_hash = concatenateInfoAndHash(msg.sender, serial_number, product_type, creation_date);

    require(products[product_hash].manufacturer == address(0), "Product ID already used");

    Product memory new_product = Product(msg.sender, serial_number, manufacturer_location, product_type, creation_date, batch, quantity, description, part_array);
    products[product_hash] = new_product;
    return product_hash;
}
```

Figure 3.7:Registering Products

(2) ChangeOwnership Smart Contract

When setting the ownership, smart contract will verify that the item exists, check if it is still unregistered and belongs to the user requesting ownership. An event will be emitted when one user transfers the ownership to another user successfully. Events can be logged

with transactions, which will be used to track an item on the chain and visualize the details to the user.

```
enum OperationType {PART, PRODUCT}
mapping(bytes32 => address) public currentPartOwner;
mapping(bytes32 => address) public currentProductOwner;

event TransferPartOwnership(bytes32 indexed p, address indexed account);
event TransferProductOwnership(bytes32 indexed p, address indexed account);
ProductManagement private pm;

constructor(address prod_contract_addr) public {…
}

function addOwnership(uint op_type, bytes32 p_hash) public returns (bool) {
    if(op_type == uint(OperationType.PART)){
        address manufacturer;
        (manufacturer, , , ) = pm.parts(p_hash);
        require(currentPartOwner[p_hash] == address(0), "Part was already registered");
        require(manufacturer == msg.sender, "Part was not made by requester");
        currentPartOwner[p_hash] = msg.sender;
        emit TransferPartOwnership(p_hash, msg.sender);
    } else if (op_type == uint(OperationType.PRODUCT)){
        address manufacturer;
        (manufacturer, , , ) = pm.products(p_hash);
        require(currentProductOwner[p_hash] == address(0), "Product was already registered");
        require(manufacturer == msg.sender, "Product was not made by requester");
        currentProductOwner[p_hash] = msg.sender;
        emit TransferProductOwnership(p_hash, msg.sender);
    }
}

function changeOwnership(uint op_type, bytes32 p_hash, address to) public returns (bool) {
  //Check if the element exists and belongs to the user requesting ownership change
    if(op_type == uint(OperationType.PART)){
        require(currentPartOwner[p_hash] == msg.sender, "Part is not owned by requester");
        currentPartOwner[p_hash] = to;
        emit TransferPartOwnership(p_hash, to);
    } else if (op_type == uint(OperationType.PRODUCT)){
        require(currentProductOwner[p_hash] == msg.sender, "Product is not owned by requester");
        currentProductOwner[p_hash] = to;
        emit TransferProductOwnership(p_hash, to);
        //Change part ownership too
        bytes32[4] memory part_list = pm.getParts(p_hash);
        for(uint i = 0; i < part_list.length; i++){
            currentPartOwner[part_list[i]] = to;
            emit TransferPartOwnership(part_list[i], to);
        }

    }
}
```

Figure 3.8:Change Ownership

3.3.2 Compiling the contract

In order to compile the smart contract, Truffle framework has been used.

After compiling smart contract using truffle framework, new file is generated which contains the JSON format of the smart contract. This file is the ABI (Abstract Binary Interface) file of the smart contract. This file can be used for following.

- ABI file consist of the compiled bytecode version of the Solidity smart contract code which can be run on Ethereum Virtual Machine (EVM), i.e., an Ethereum Node
- It contains a JSON representation of the smart contract functions and arguments.

The contract is then deployed to the Ganache personal blockchain which will mimic the behavior of the Ethereum public blockchain.

Figure 9 shows the smart contract compile and deployment workflow in more detail view.

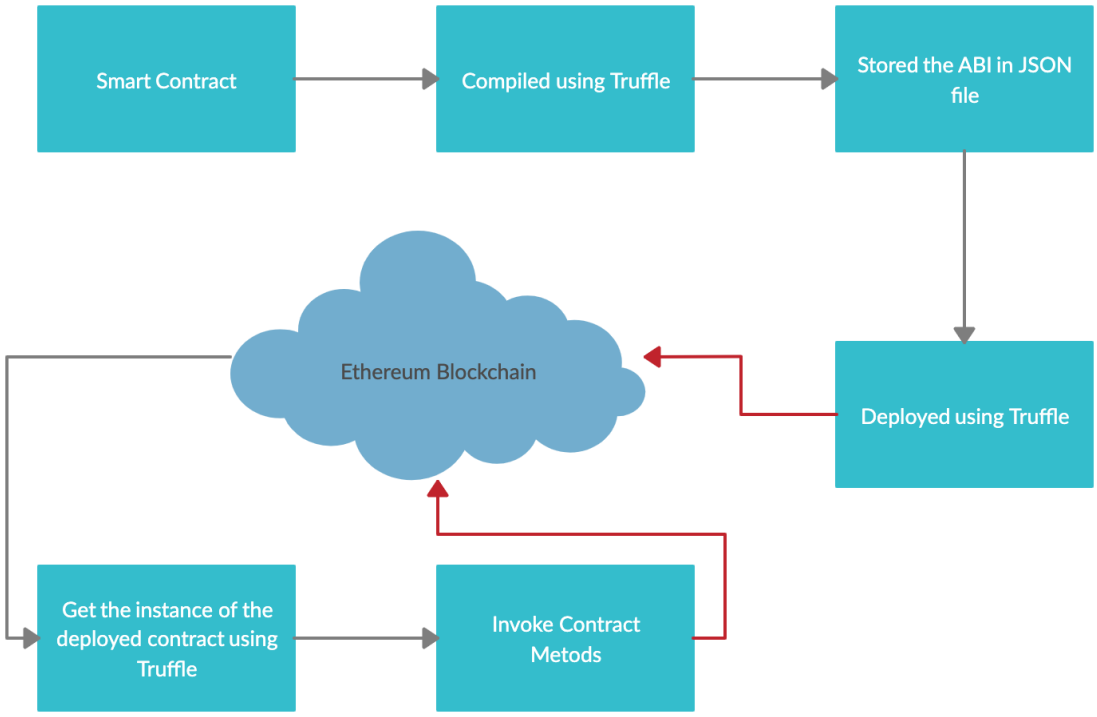


Figure 3.9:Deployment workflow of Smart Contracts

3.3.3 Deploying the contract

After compiling the smart contract, we can deploy the smart contract to the personal blockchain using truffle framework.

We can create a migration script file for this. Script file will look like below

```
var Migrations = artifacts.require("./Migrations.sol");
var ProductManagement = artifacts.require("./ProductManagement.sol");
var ChangeOwnership = artifacts.require("./ChangeOwnership.sol");

module.exports = function(deployer) {
  deployer.deploy(Migrations);
  deployer.deploy(ProductManagement)
  .then(function(){
    return deployer.deploy(ChangeOwnership, ProductManagement.address);
  })
};
```

Figure 3.10:Contract Deployment

3.3.4 Web3

Web 3 is a JavaScript library that allows client-side application to communicate with Blockchain network in order to store and retrieve information. Web3 uses JSON RPC (Remote Procedure Call) to communicate with the network. Web3 send the request to the network node using JSON RPC, when it is necessary to store or retrieve information from blockchain. Figure 11 shows the script implement to create a web3 instance.

```
async function init_web3() {
    //Web3 init
    if (typeof web3 != 'undefined') {
        web3 = new Web3(web3.currentProvider) // what Metamask injected
    } else {
        // Instantiate and set Ganache as your provider
        web3 = new Web3(new Web3.providers.HttpProvider("http://localhost:8545"));
    }
    //Load accounts
    window.accounts = await web3.eth.getAccounts()
    console.log("Loaded accounts")

}
```

Figure 3.11:Web3

3.3.5 User Authentication

In order to handle the user authentication and user registration MetaMask is used in the system. MetaMask is a cryptocurrency wallet and an open-source web browser extension that stores users' public and private keys. It is used to connect to the Ethereum blockchain in order to access the decentralized applications. Parties in the network can exchange ethers using public and private keys. Anyone in the network can see the public key, while the private key should be kept secret because that is the only way for users to prove the ownership of the cryptocurrencies they have in their accounts.

Because of the hierarchical deterministic nature of this online wallet, MetaMask was used in this system. Users can use a 12-word phrase or account mnemonic to protect their accounts. Since the users' public and private keys are stored in the browser using MetaMask, need for the remote server is eliminated. It will improve the privacy as well. Another reason to use MetaMask rather than traditional login methods is its ability to interact with the web page by injecting web3 into the browser. To sign data, some web3 functions require the use of private keys. In that case, MetaMask will request user approval before signing data with their private keys.

4. Evaluation Plan

The primary goal of this research project is to manage provenance data throughout the product journey in the supply chain process and present it to supply chain consumers. With the existing systems following gaps have been identified.

1. Ability for supply chain entities to read and view provenance information. (Forward and backward tracking)
2. Ability to track the entire supply chain information using a single system.
3. Ability to access the system without the permission from third party organization.
4. Using traditional client server architecture which will lead to data tampering, single point of failure and traceability and transparency issues.
5. Proposed models can be only used in one industry (e.g., in food supply chain). There is no general solution which can use in every industry.

In order to overcome the above motioned problems, an architecture have been proposed and implemented. In this chapter main goal is to evaluate the proposed architecture to determine whether the identified gaps have been addressed.

In order to simplify the supply chain process, three main entity types have been identified. These entity types include supplier, manufacture, and distributor.

- **Ability for supply chain entities to read and view provenance information. (Forward and backward tracking)**

Every user in the system should be able to view the provenance information.

As the manufacturer or distributor, user should be able to view the provenance information of the products and primary ingredients that they own.

As the supplier, user should be able to view the provenance information of the primary ingredients that they own.

As the customer, user should be to search for the provenance data of a product.

Below figures shows how the proposed system has achieved this.

Distributor view

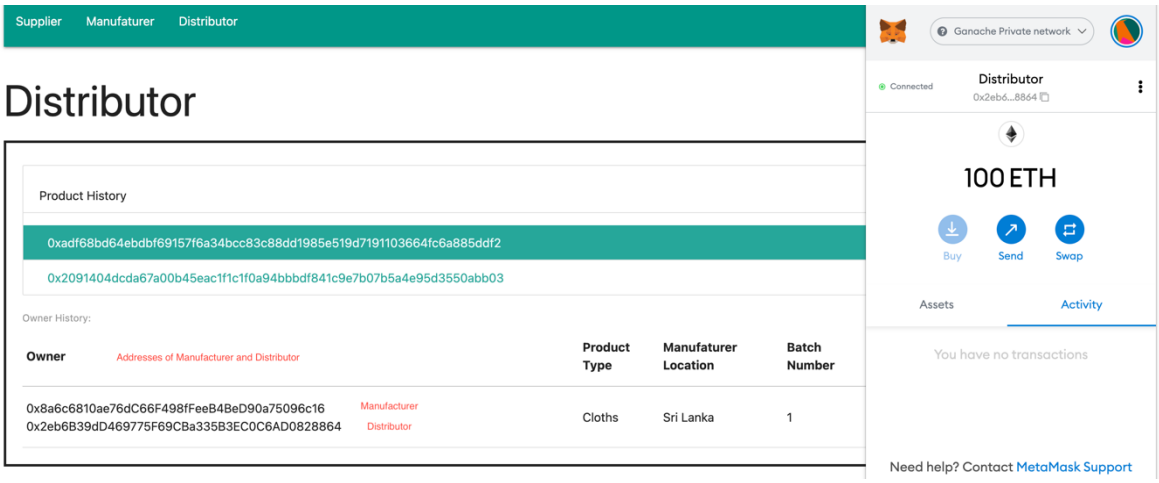


Figure 4.1:Distributer View 1

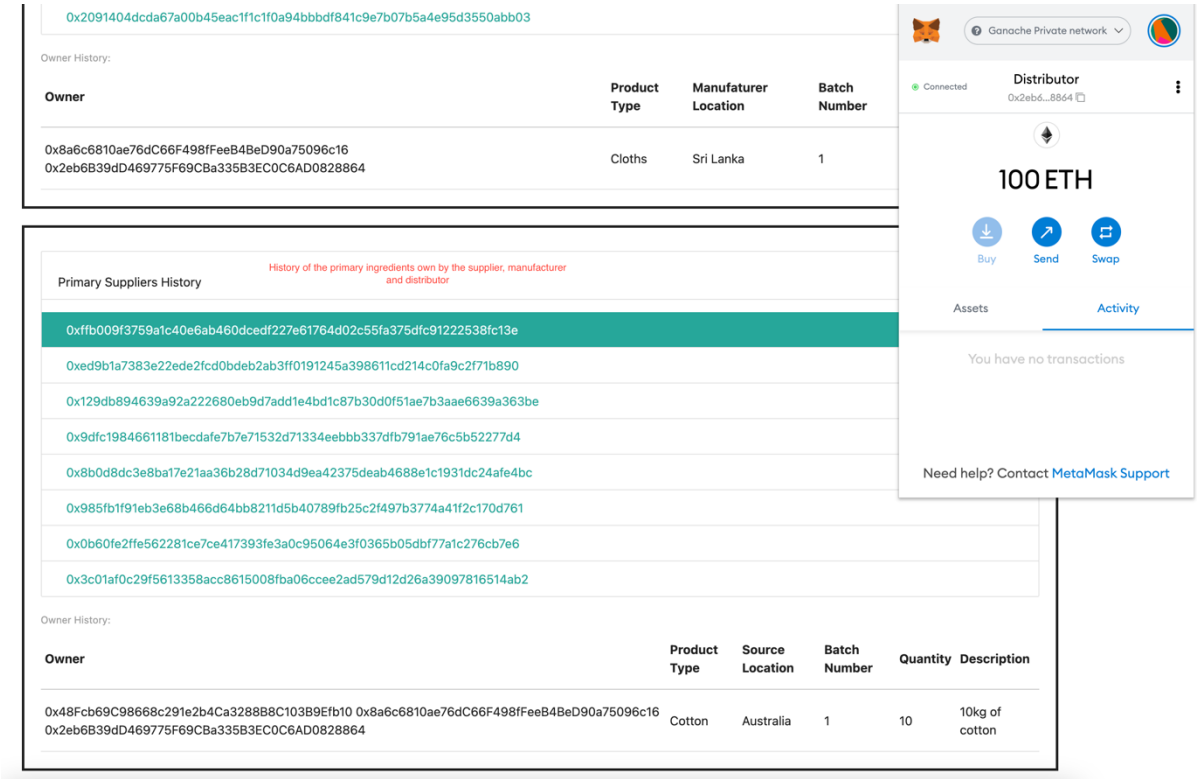


Figure 4.2:Distributor View 2

Manufacture view

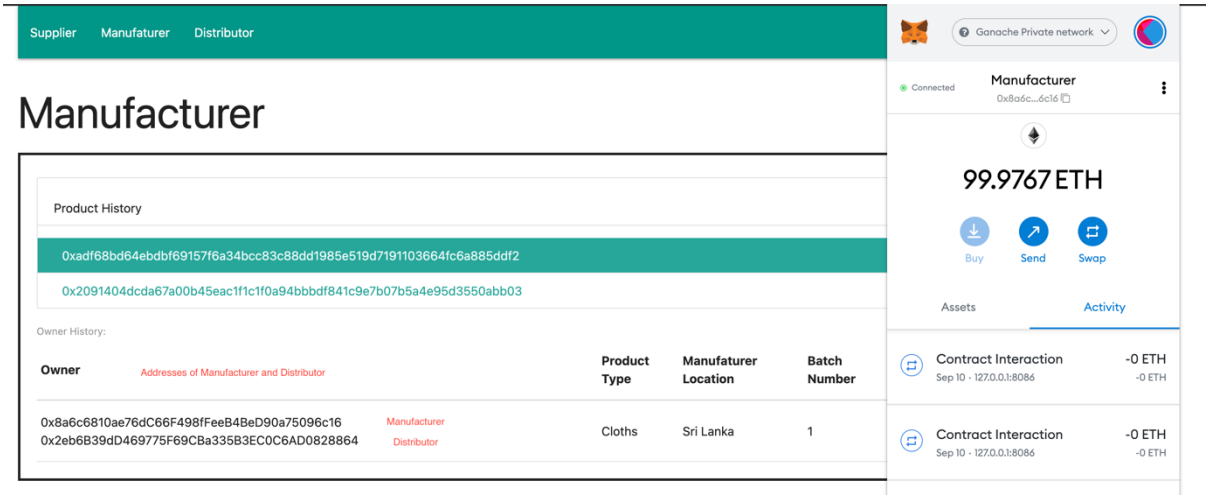


Figure 4.3:Manufacturer View 1

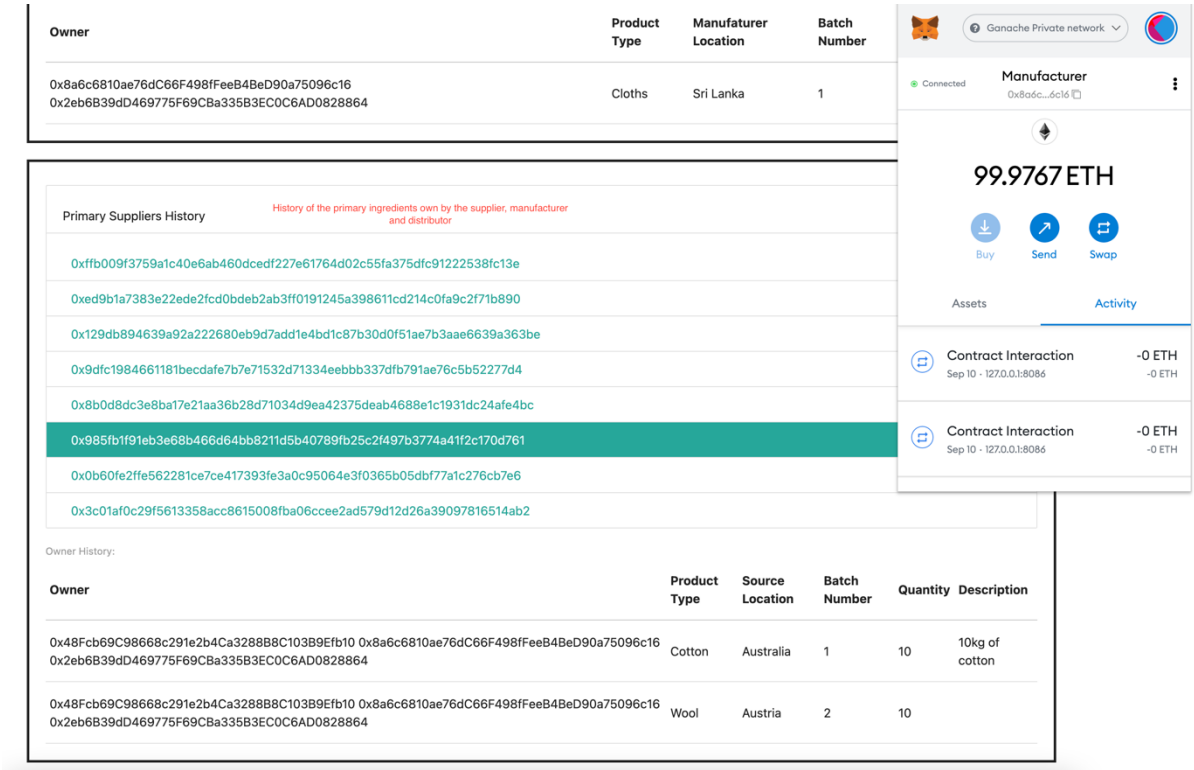


Figure 4.4:Manufacturer View 2

Supplier view

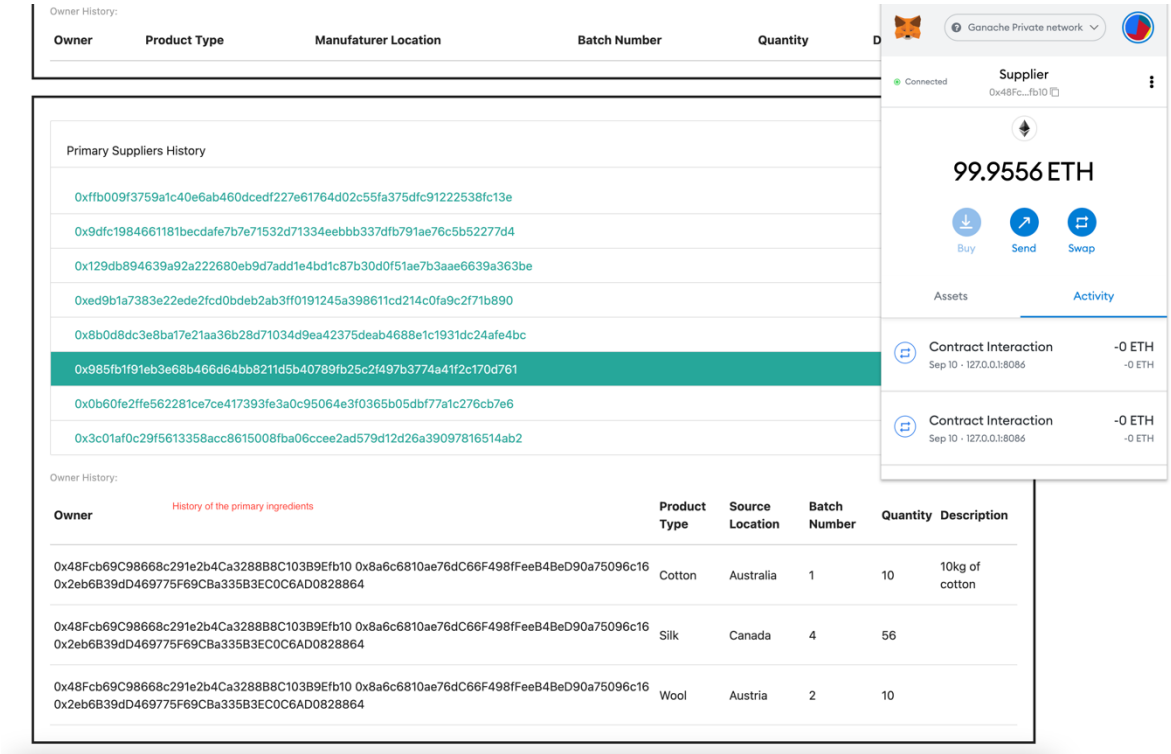


Figure 4.5:Supplier View

Customer View

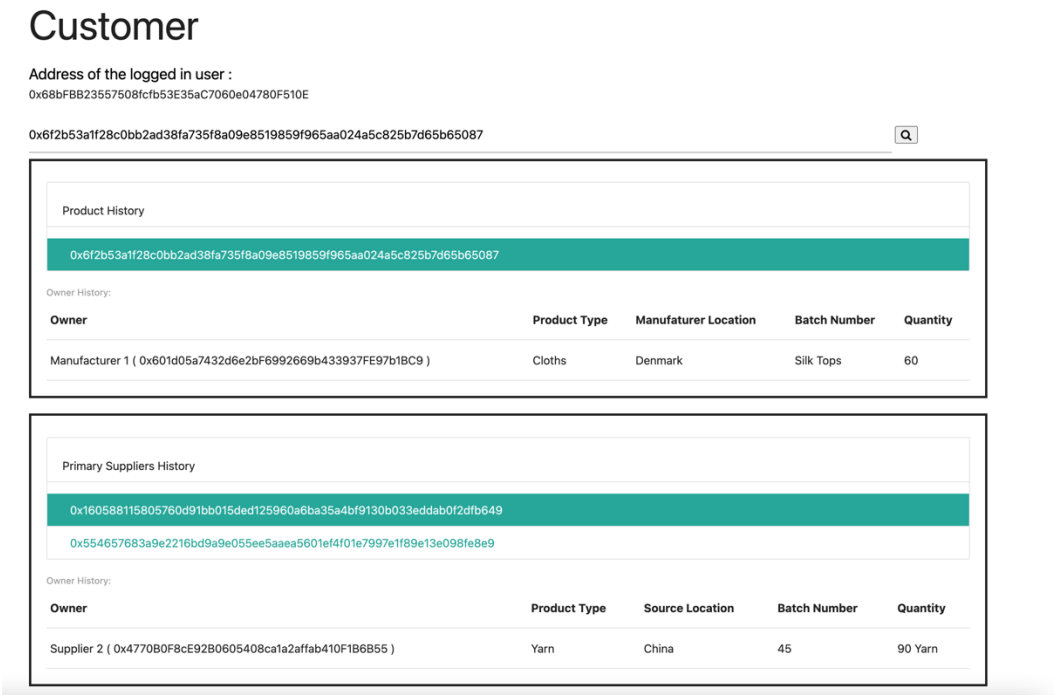


Figure 4.6:Customer View

- **Ability to track the entire supply chain information using a single system.**

As mentioned above only three main entity types used in order to simplify the supply chain process. Based on the above figures, the entire flow has been tracked and visualized, beginning with the raw material supplier, and continuing through the manufacturer until the product arrives at the distributor.

- **Accessibility is limited to the authorized parties.**

The Ethereum platform was used to construct the proposed system. Ethereum is a public blockchain network. Anyone can join the network and participate within the blockchain. Like in the permissioned blockchain, accessibility is not limited to authorized parties.

- **Using traditional client server architecture which will lead to data tampering, single point of failure and traceability and transparency issues.**

The proposed system achieved decentralized architecture using blockchain technology. So, the drawbacks in the client server architecture have been eliminated.

- **Proposed models can be only used in one industry**

The proposed model is a generalized solution for managing the provenance of any supply chain in any industry. Adaptation for Chocolate supply chain is shown below.

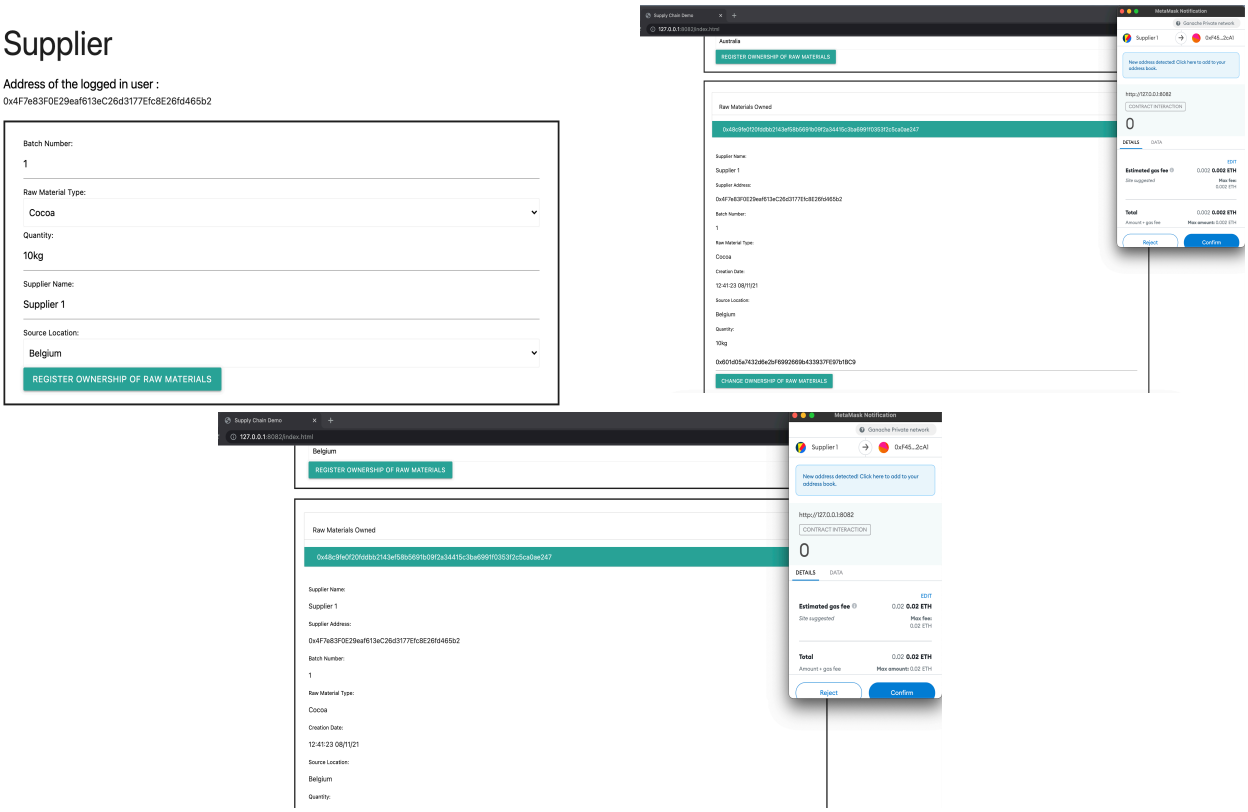


Figure 4.7:Adaptation for Chocolate Manufacturing Process

5. Conclusion and Future Work

5.1 Conclusion

In this work main objective was to increase transparency of supply chain by providing product prevenance. Transparency was the main concern here, as it provides numerous benefits to all the parties in the supply chain. When transparency of supply chain increases, it will automatically help to improve the consumers trust on products.

In order to visualize provenance of product blockchain based solution has been proposed and developed. Blockchain technology choose over other technologies due to following reasons.

- Decentralized - Since the system is decentralized having only one authority to control data can be eliminated.

- Immutable - Once data is entered to the system it cannot be changed by any of the users in the system which make data permanently available.

- Transparent – As data in the system are always available, visibility of the system increases, which makes the system more reliable and trustworthy.

By using the proposed system, users can view provenance information of a product which will improve the transparency of the supply chain.

5.2 Future Work

As mentioned in the previous section, main objective of the project is to increase transparency by visualizing provenance of products in a supply chain. Blockchain based solution has been proposed and developed using Ethereum Platform. Following are the few aspects that this project can be improve in the future.

- Proposed decentralized solution developed using Ethereum blockchain. There are several other blockchain platforms that are available like Hyperledger and Multichain. By using these platforms, reimplementing the proposed solution is possible with few changes. By reimplementing the same application using other platforms, efficiency of these platforms can be compared with each other.

- In the proposed solution, users enter the information of production process manually. Entering information can be automated using barcodes and RFIDs which will eventually improve the accuracy of data.
- Proposed solution can be improved to give the organizations ability to examine the data in the production process. This will help the organizations when making decisions in future.

6. References

[1] J. T. Mentzer, W. DeWitt, J. S. Keebler, S. Min, N. W. Nix, C. D. Smith, and Z. G. Zacharia, “Defining supply chain management,” Journal of Business logistics, vol. 22, no. 2, pp. 1–25, 2001.

[2] Shields, T. N. J. B. L. (2014). 'A Guide to trAceAbility' A Practical Approach to Advance Sustainability in Global Supply Chains. Retrieved from New York, NY https://bit.ly/1tQEnn0

[3] L. U.Opara, “Traceability in agriculture and food supply chain: a review of basic concepts, technological implications, and future prospects,” Journal of Food Agriculture and Environ- ment, vol. 1, pp. 101–106, 2003.

[4] D. Galvin, “Ibm and walmart: Blockchain for food safety,” PowerPoint presentation, 2017.

[5] A. Kassahun, R. Hartog, T. Sadowski, H. Scholten, T. Bartram, S. Wolfert, and A. Beulens, “Enabling chain-wide transparency in meat supply chains based on the epcis global standard and cloud-based services,” Computers and electronics in agriculture, vol. 109, pp. 179–190, 2014.

[6] C. N. Verdouw, J. Wolfert, A. Beulens, and A. Rialland, “Virtualization of food supply chains with the internet of things,” Journal of Food Engineering, vol. 176, pp. 128–136, 2016.

[7] W. Wang, Y. Liu, and S. Wang, “A rfid-enabled tracking system in wire bond station of an ic packaging assemble line,” in RFID-Technology and Applications (RFID-TA), 2010 IEEE International Conference on, pp. 170–175, IEEE, 2010.

[8] D. Folinas, I. Manikas, and B. Manos, “Traceability data management for food chains,” British Food Journal, vol. 108, no. 8, pp. 622–633, 2006.

[9] C. Shanahan, B. Kernan, G. Ayalew, K. McDonnell, F. Butler, and S. Ward, “A framework for beef traceability from farm to slaughter using global standards: an irish perspective,” Computers and electronics in agriculture, vol. 66, no. 1, pp. 62–69, 2009.

[10] Wang-Chiew Tan. Research Problems in Data Provenance. IEEE Data Engineering Bulletin, 27(4):42–52, 2004.

[11] Peter Buneman, Sanjeev Khanna, and Wang Chiew Tan. Why and Where: A Characterization of Data Provenance. In ICDT ’01: Proceedings of the 8th International Conference on Database Theory, pages 316–330, London, UK, 2001. Springer-Verlag.

[12] Allison Woodruff and Michael Stonebraker. Supporting Fine-grained Data Lineage in a Database Visualization Environment. In ICDE ’97: Proceedings of the Thir- teenth International Conference on Data Engineering, pages 91–102, Washington, DC, USA, 1997. IEEE Computer Society.

[13] Dennis P. Groth. Information Provenance and the Knowledge Rediscovery Problem. In IV, pages 345–351. IEEE Computer Society, 2004.

[14] M. R. Asghar, M. Ion, G. Russello, and B. Crispo, “Securing data provenance in the cloud,” in Open Problems in Network Security. Springer, 2012, pp. 145–160.

[15] R. Hasan, R. Sion, and M. Winslett, “Sprov 2.0: A highly configurable platform independent library for secure provenance,” ACM, CCS, Chicago, IL, USA, 2009.

[16] R. K. Ko and M. A. Will, “Progger: An efficient, tamper- evident kernel-space logger for cloud data provenance track- ing,” in 2014 IEEE 7th International Conference on Cloud Computing. IEEE, 2014, pp. 881–889.

[17] M. Imran and H. Hlavacs, “Applications of provenance data for cloud infrastructure,” in Semantics, Knowledge and Grids (SKG), 2012 Eighth International Conference on. IEEE, 2012, pp. 16–23

[18] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008.

[19] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, “Medrec: Using blockchain for medical data access and permission management,” in Open and Big Data (OBD), International Conference on, pp. 25–30, IEEE, 2016.

[20] S. Rouhani and R. Deters, “Performance analysis of ethereum transactions in private blockchain,” in 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 70–74, IEEE, 2017.

[21] H. Natarajan, S. K. Krause, and H. L. Gradstein, “Distributed ledger technology (dlt) and blockchain,” FinTech note, 2017.

[22] L. A. Linn and M. B. Koo, “Blockchain for health data and its potential use in health it and health care related research,” in ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST, 2016.

[23] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, “An overview of blockchain technology: Architecture, consensus, and future trends,” in Big Data (BigData Congress), 2017 IEEE International Congress on, pp. 557–564, IEEE, 2017.

[24] OpenSea.pro, “How can the shipping industry take advantage of the blockchain technology?.” https://opensea.pro/blog/blockchain-for-shipping-industry

[25] Sidra Malik, Salil S. Kanhere, Raja Jurdak. ProductChain: Scalable Blockchain Framework to Support Provenance in Supply Chains

[26] Henry M. Kim, Marek Laskowski. Toward an ontology-driven blockchain design for supply-chain provenance

[27] Ramachandran, A. and Kantarcioglu, M. SmartProvenance: A Distributed, Blockchain Based DataProvenance System. In CODASPY ’18: Eighth ACM Conference on Data and Application Security and Privacy, March 19–21, 2018.

[28] Sandino Moeniralam, A Blockchain based Data Production Traceability System.

[29] Rodrigo Almeida, Waldeyr da Silva, Klayton Castro, Maria Em´ılia Walter, Aleteia Araujo, Maristela Holanda and Sergio Lifschitz. AProvBio: An Architecture for Data Provenance in Bioinformatics Workflows using Graph Database. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).

[30] Marten Sigwart, Michael Borkowski, Marco Peise, Stefan Schulte, and Stefan Tai. 2019. Blockchain-based Data Provenance for the Internet of Things. In Proceedings of 9th International Conference on the Internet of Things (IOT'19).

[31] Patrick, K. Buzzword of the Year: Blockchain. https://www.supplychaindive.com/news/buzzword-of-the-year-blockchain/511022/

[32] Cheney, J., Chong, S., Foster, N., Seltzer, M., & Vansummeren, S. (2009). Provenance: A future history. Proceedings of the 24th ACM SIGPLAN Conference Companion on Object Oriented Programming Systems Languages and Applications (pp. 957—964).

[33] Bart Slob 2008. Global supply chains: the importance of traceability and transparency

[34] River Island recalls clothes over fears they contain harmful levels of dangerous chemicals. Available at https://www.independent.co.uk/life-style/fashion/river-island-product-recall-lead-cadmium-clothes-harmful-a8992946.html

[35] Sourcemap, "End-to-end supply chain visualization white paper," tech. rep., [Accessed May 08, 2019].

[36] G. Bhosle, P. Kumar, B. Griffin-Cryan, R. van Doesburg, M. Sparks, and A. Paton, "Global supply chain control towers: Achieving end-to-end supply chain visibility," Capgemini Con- sulting White Paper, 2011.

[37] F. Dabbene, P. Gay, and C. Tortia, “Traceability issues in food supply chain management: A review,” Biosystems Engineering, vol. 120, pp. 65 – 80, 2014.

[38] F. Tian, “An agri-food supply chain traceability system for china based on rfid & blockchain technology,” in Service Systems and Service Management (ICSSSM), 2016 13th International Conference on. IEEE, 2016, pp. 1–6.

[39] “A supply chain traceability system for food safety based on haccp, blockchain internet of things,” in 2017 International Conference on Service Systems and Service Management, June 2017, pp. 1–6.

[40] K. Biswas, V. Muthukkumarasamy, and W. L. Tan, “Blockchain based wine supply chain traceability system,” in Future Technologies Conference, 2017.

[41] K. Toyoda, P. T. Mathiopoulos, I. Sasase, and T. Ohtsuki, “A novel blockchain-based product ownership management system (poms) for anti-counterfeits in the post supply chain,” IEEE Access, vol. 5, pp. 17 465–17 477, 2017.

[42] S. Abeyratne, R. Monfared, “Blockchain Ready Manufacturing Supply Chain Using Distributed Ledger”.

[43] G. Baralla, S. Ibba, M. Marchesi, R. Tonelli, and S. Missineo, “A Blockchain Based System to Ensure Transparency and Reliability in Food Supply Chain”