Gao Z, Xu L, Chen L *et al.* CoC: A unified distributed ledger based supply chain management system. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 33(2): 237–248 Mar. 2018. DOI 10.1007/s11390-018-1816-5

CoC: A Unified Distributed Ledger Based Supply Chain Management System

Zhimin Gao[1], *Member*, *IEEE*, Lei Xu[1], Lin Chen[1], Xi Zhao[2], *Member*, *IEEE*, Yang Lu[1]
and Weidong Shi[1], *Member*, *IEEE*

[1] *Department of Computer Science, University of Houston, Houston, Texas 77204-3010, U.S.A.*
[2] *School of Management, Xi'an Jiaotong University, Xi'an 710049, China*

E-mail: mtion@msn.com; {xuleimath, chenlin198662, zhaoxi1}@gmail.com; {ylu17, wshi3}@central.uh.edu

Received July 9, 2017; revised January 18, 2018.

Abstract Modern supply chain is a complex system and plays an important role for different sectors under the globalization economic integration background. Supply chain management system is proposed to handle the increasing complexity and improve the efficiency of flows of goods. It is also useful to prevent potential frauds and guarantee trade compliance. Currently, most companies maintain their own IT systems for supply chain management. However, it is hard for these isolated systems to work together and provide a global view of the status of the highly distributed supply chain system. Using emerging decentralized ledger/blockchain technology, which is a special type of distributed system in essence, to build supply chain management system is a promising direction to go. Decentralized ledger usually suffers from low performance and lack of capability to protect information stored on the ledger. To overcome these challenges, we propose CoC (supply chain on blockchain), a novel supply chain management system based on a hybrid decentralized ledger with a novel two-step block construction mechanism. We also design an efficient storage scheme and information protection method that satisfy requirements of supply chain management. These techniques can also be applied to other decentralized ledger based applications with requirements similar to supply chain management.

Keywords blockchain, distributed system, supply chain management, security

1 Introduction

Modern economy heavily depends on global collaboration. According to a World Trade Organization (WTO) report, the international trade volume keeps increasing at a high rate in the past decades and merchandise exports from WTO members achieved US$ 18.0 trillion in 2014①. Behind this explosive growth, supply chain plays a critical role. Besides classical functions such as making movements of goods smoother and reducing the cost of international transportation, modern supply chain system is becoming the center of various business activities such as planning/forecasting, procurement, customer services, and performance measurement. It becomes a challenge to manage modern supply chain efficiently due to its large scale and complex functionalities. In response to such demands, the concept of supply chain management system was introduced by Oliver in 1982[1], and the market of supply chain management software outpaced most software

Regular Paper

Special Section on Computer Networks and Distributed Computing

A preliminary version of the paper was published in the Proceedings of ICCCN 2017.

This material is based upon the work supported by the U.S. Department of Homeland Security (DHS) under Grant Award No. 2015-ST-061-BSH001 and the National Nature Science Foundation of China under Grant No. 91746111. The grant from DHS is awarded to the Borders, Trade, and Immigration (BTI) Institute: A DHS Center of Excellence led by the University of Houston, and includes support for the project "Secure and Transparent Cargo Supply Chain: Enabling Chain-of-custody with Economical and Privacy Respecting Biometrics, and Blockchain Technology" awarded to University of Houston. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

① World Trade Organization. International trade statistics 2015. https://www.wto.org/english/res_e/statis_e/its2015_e/its2015_e.pdf, Jan. 2018.

markets to total US$ 9.9 billion in 2014[2]. A lot of work has been done to improve the efficiency of supply chain management system and add more features. For example, researchers proposed to integrate sensors (e.g., GPS receiver[2] and radio-frequency identification/RFID[3]) into the supply chain to provide more information to the end user, and bind the cyber world and the physical world more tightly[4]. As cloud computing technology emerges, cloud-based supply chain management system is also developed to improve the reliability and reduce the cost[5].

However, existing supply chain management systems suffer from some limitations that prevent users from achieving most out of the value of supply chain information. The two major issues are as follows. 1) Supply chain in nature involves multiple parties and is a distributed system. However, most companies and stakeholders nowadays use their own supply chain management systems, which are difficult to be integrated together to provide a unified platform. Therefore, it is not convenient to offer end-to-end tracking and share information to enable new functionalities and services. Furthermore, supply chain information is sensitive and the companies may not be willing to disclose and share with others. 2) As an IT system, supply chain management system faces all types of cyber threats, which may lead to breach of the integrity of supply chain information and cause fraud, losses of goods, and incompliance in trading. The recent rising of ransomware attack also poses a significant risk to supply chain management system as losing access to historical data can cause financial damages[6][3].

Decentralized ledger technology (DLT) provides a way to organize records in a distributed manner through consensus mechanism. It has been used in Bitcoin and other similar cryptocurrency systems for recording and sharing transaction history[4][5] and is constructed by a group of users together, and each of them maintains a local copy of the ledger. A group of records are embedded into a block and blocks are linked through hash values. A consensus mechanism helps these users achieve agreement when a new block is added to the system. If there is more than one branch on the chain, usually the "longest-chain" principle is used, i.e., users will follow the branch with more blocks and add new blocks on this branch. In order to alter an existing block, an adversary has to compete with all honest users to construct a longer branch[7-8]. Therefore, DLT provides a collaboration mechanism that can protect historical data.

These features make decentralized ledger a promising technology for global supply chain management system, which is in essence a distributed environment. Both technology startups and transnational corporations start to experiment supply chain management systems based on distributed ledger[9-10][6]. However, most of the existing efforts on creating supply chain management system with DLT are straightforward, i.e., they just use DLT as a decentralized storage system to store supply chain related information in blocks to replace traditional file system, but ignore downsides of the technology listed as follows. 1) Decentralized ledger usually has performance issues such as limited throughput/long latency for adding new blocks and inefficient storage, which may not be sufficient to support application scenarios with requirements to store high volume of supply chain operation records and support high transaction throughput. 2) Information stored in decentralized ledger is distributed to and maintained by different nodes. There is a lack of mechanism to protect supply chain related information stored in the distributed ledger from unauthorized access.

To address these shortcomings, we propose CoC (supply chain on blockchain), a novel supply chain management system which leverages the decentralized ledger technology. CoC uses a hybrid model and two-step block construction method for the underlying distributed ledger, which achieves a good balance between security and performance. In addition, CoC introduces a new storage scheme that reduces data redundancy without affecting distributed ledger related operations. Because supply chain management system plays a central role in business operations that involve sensitive information, a protection mechanism is built on top of the hybrid model and the storage scheme to guaran-

[2] Gartner says worldwide supply chain management and procurement software market grew 10.8 percent in 2014, May 2015. http://www.gartner.com/newsroom/id/3050617, Jan. 2018.

[3] Ransomware: A growing menace. https://www.symantec.com/connect/blogs/ransomware-growing-menace, Jan. 2018.

[4] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system, 2008. https://bitcoin.org/bitcoin.pdf, Jan. 2018.

[5] King S. Primecoin: Cryptocurrency with prime number proof-of-work, 2013. http://primecoin.io/bin/primecoin-paper.pdf, Jan. 2018.

[6] Parker L. Blockchain tech companies focus on the $40 trillion supply chain market, 2016. https://bravenewcoin.com/news/blockchain-tech-companies-focus-on-the-40-trillion-supply-chain-market/, Jan. 2018.

tee that only authorized users can access corresponding data on the ledger.

Our contributions in this work are summarized as follows.

• We propose a novel design of supply chain management system based on public ledger that serves as a unified platform for different parties and stakeholders involved in the supply chain ecosystem to conduct transactions and share information.

• We develop a two-step block generation method for the system which has low latency, and an efficient storage scheme that alleviates the concern of storage overhead of decentralized ledger technology.

• We also provide the design of identity management and data protection scheme that addresses security issues for decentralized ledger based supply chain management system.

The remainder of this paper is organized as follows. Section 2 briefly describes the supply chain management system and decentralized ledger technology. In Section 3, we provide an overview of the proposed CoC system and the hybrid model for ledger construction. Detailed design of critical components of CoC is given in Section 4, and we analyze the security/performance of CoC in Section 5. Section 6 reviews related work and the work is concluded in Section 7.

2 Background

In this section, we briefly review the supply chain management system and DLT.

2.1 Supply Chain Management System

Supply chain management is not a single extension of logistics management, but an integration of business processes from end users through original suppliers that provides products, services, and information that add value for customers[11]. Typical supply chain management functions include ordering/receipt of raw materials/products, supporting customer services, and performance measurements. The coordination of multiple functions across the enterprise is required to provide rapid and quality response to supply chain events[12]. Fig.1 depicts the functions of supply chain management and its position in business operations. Besides handling physical cargos, supply chain system is now also used for data transfer (e.g., Fedex is helping Amazon customers to move a giant amount of data).

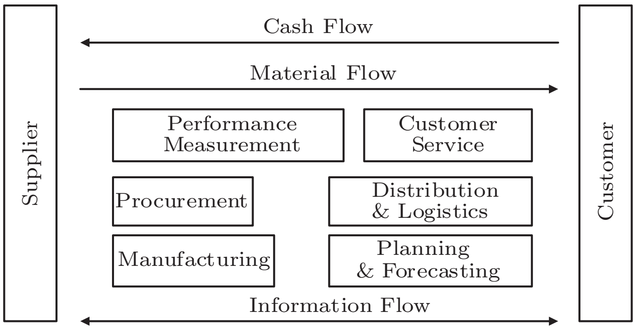


Fig.1. Role of supply chain in business operations. It manages the information flow and provides the foundation for various functions[12].

2.2 Decentralized Ledger Technology

Decentralized ledger or blockchain technology was first introduced by Bitcoin as a distributed bookkeeping system[7]. As each user keeps a local copy of the ledger, he/she has access to all historical transaction information and detects double-spending without relying on a trusted third party.

Bitcoin uses proof-of-work to control the construction of blocks, which is depicted in Fig.2. Information is embedded into a block, which also contains a hash value from the previous block and a magic number. The magic number is found out through a brute-force searching process, i.e., one searches all possible values of magic number to make sure the hash value of the triple (previous hash value, embedded information, and magic number) satisfies pre-defined condition (e.g., the hash value has a certain number of leading zeros). Specifically, in order to create a new block, one has to find a magic number to make sure the hash value of the block satisfies the pre-defined condition (e.g., smaller than a constant value). When more than one valid block is added to the ledger that causes branches, users will follow the "longest-chain" that contains more blocks. If an attacker wants to replace or remove an existing block in the ledger, he/she has to compete with all honest participants of the system to generate more valid blocks to make sure his/her branch is longer.

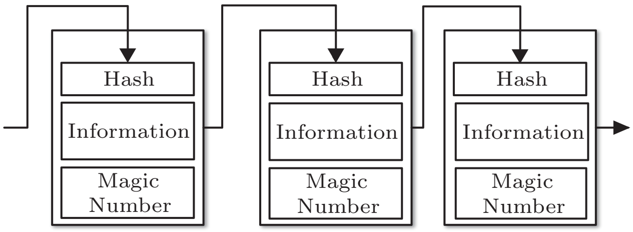


Fig.2. Basic working principles of decentralized ledger with proof-of-work.

[7] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system, 2008. https://bitcoin.org/bitcoin.pdf, Jan. 2018.

In summary, distributed ledger has three key features.

1) *Public Accessibility*. All information stored with blockchain is publicly accessible to everyone.

2) *Immutability*. It is impossible to modify, alter, or remove information that has been added to the blockchain.

3) *Resilience*. Each participant of the system keeps a whole copy of the blockchain and no single point of failure can affect the availability of the stored information.

One major issue of proof-of-work based distributed ledger construction is the high latency of block generation, which is caused by the expensive mining process, e.g., brute-force searching for pre-image of a hash function. Different approaches have been proposed to improve the performance, and proof-of-stake and permissioned distributed ledger are two major schemes. We summarize these three methods as follows.

- *Proof-of-Work*[7]. In order to construct a new block and add it to the distributed ledger, a participant has to solve a computation intensive problem and attach the result to the new block as proof of his/her work.

Pros. The mechanism is simple and fair.

Cons. It wastes a lot of computation resources and has relatively high latency.

- *Proof-of-Stake*[13]. Participants accumulate stake according to the pre-defined accumulation scheme, and a certain amount of stake has to be used to create a new block. Therefore, any participant who has enough stake can generate a new block instantly.

Pros. This approach can generate blocks with very low latency when the system has enough stake available.

Cons. It is a challenge to design a stable stake accumulation scheme, and the system may go to two extreme statuses: no one has enough stake to generate a block, or everyone has enough stake to generate a block.

- *Permissioned*[14]. A set of trusted parties is responsible for block generation. One party that belongs to the set can attach a signature to the block and the block is recognized as a valid one.

Pros. The mechanism is simple and new blocks can be generated very fast.

Cons. This strategy requires a different security model (e.g., some nodes are trusted and it is not public) and only fits certain scenarios like transactions between financial institutes.

Both proof-of-work and proof-of-stake based distributed ledger use the longest-chain strategy to resolve disagreements, and permissioned distributed ledger can leverage consensus protocols like Byzantine fault tolerant protocol to avoid disagreements[15-16].

3 Overview of CoC

In this section, we give an overview of CoC and describe the hybrid model that CoC leverages for block construction.

3.1 Participants in CoC

As a unified supply chain management platform, CoC needs to support different types of participants including factories, supply chain operators, financial institutes, insurance companies, and customs. According to their roles in the supply chain management, we divide all participants into three groups.

- *Ordinary Users*. An ordinary user can use CoC for different supply chain related operations, e.g., submitting new request for raw material, tracking transportation information, processing bill of lading, and analyzing historical data related to the user. Supply chain is a complex system and CoC supports multiple ordinary users to collaborate with each other. Ordinary users are the major information contributors to CoC.
- *Third Party Users*. Besides ordinary users, there is another group of users, third party users, who mainly monitoring supply chain information with CoC. Typical third party users include government entities such as customs and insurance companies who need to monitor the status of the goods.
- *Supporting Entities*. CoC also includes some supporting entities for supply chain operations. Two of the main supporting entities are identity management component and financial institutions. Here identity management can be part of CoC, while financial institutions have their own IT system and only interact with CoC to provide required services such as payment processing.

In the following of the paper, if not explicitly stated, the term "user" stands for ordinary user, third party user, or both of them.

3.2 Hybrid Model of CoC

Existing models of decentralized ledger do not fit the requirements of supply chain management very well.

- Proof-of-work involves heavy computation and is usually slow, which may not be able to satisfy the demands of supply chain management.

- Proof-of-stake is not stable for supply chain management system as it is hard to predict the demands of blocks.
- CoC aims at providing a unified supply chain management platform that can serve multiple entities that do not need to fully trust each other, and it is hard to achieve an agreement on the nodes that compromise the permissioned network for block construction.

Considering all the limitations of existing models and the special requirements of supply chain management, CoC separates the right to submit records and the right to build blocks by using a hybrid model to organize the underlying distributed ledger. Specifically, CoC allows only users, third party users, and supporting entities to submit supply chain related records to the system, but the block construction is open to the public and based on proof-of-work. Those who contribute their computation resources to help to build and maintain the distributed ledger are called "helpers". The number of helpers is relatively large and driven by the demands. CoC does not put much restriction on helpers. Anyone with reasonable computation resources can join the system to contribute to block construction, as described in Subsection 4.1. They can also leave the system freely. Fig.3 illustrates the system and different types of entities involved. Users (e.g., factories, transportation companies) use the system for supply chain information management. Third party users include insurance companies and government regulators. In most cases, they just monitor information stored in CoC and do not add new information. Support entities include financial institutions for payment service and identity management component. The system also involves a large number of helpers, who facilitate the generation of blocks that are used to hold supply chain information. Helpers play the role similar to miners of cryptocurrency systems like Bitcoin[8] and Ethereum[9].

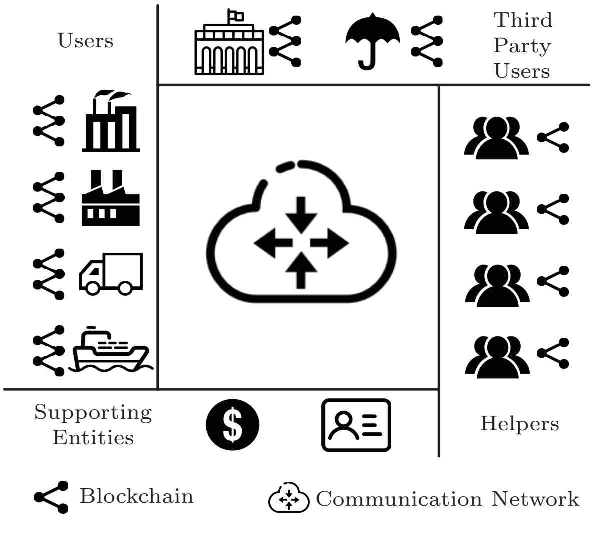


Fig.3. Overview of CoC.

Security Model. We assume supporting entities are fully trusted, e.g., they will follow pre-defined protocols to collaborate with other parties and will not try to inject faked information into the system. Third party users are usually large companies and government agencies, and also trusted, and they will follow the policies to perform their tasks (e.g., generating certificate of compliance or insurance). Any individual helper is not trusted, and he/she may try to compromise the system using different ways. However, the number of helpers is usually large, and the majority of them are honest and will follow pre-defined protocols. The users are not fully trusted. Although they have the incentive to keep accurate information to support their business activities, it is hard to guarantee that all of the users have adequate cyber protection and they may be compromised (e.g., loss of private key, infected by Trojan or viruses). A compromised user may try to generate invalid supply chain information and/or try to modify historical data. We also assume communications between different parties are secure, i.e., an attacker cannot tamper or eavesdrop the exchanged messages between any two parties, which can be achieved by using SSL (Security Socket Layer).

4 Detailed Design of Key Components of CoC

In this section, we describe the design of key components of CoC, including ledger construction, storage scheme, identity management, and information protection.

4.1 Block Construction in CoC

As discussed earlier, one of the main challenges of using decentralized ledger for supply chain management system is to support a large number of operations in a short time. According to the overview of CoC given in Section 3, users are not fully trusted and permissioned

[8] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system, 2008. https://bitcoin.org/bitcoin.pdf, Jan. 2018.

[9] Wood G. Ethereum: A secure decentralised generalised transaction ledger. http://www.cryptopapers.net/papers/ethereum-yellowpaper.pdf, Jan. 2018.

blockchain system cannot be used to reach low latency block construction. The proof-of-stake strategy does not work well either for supply chain management because the amount of transactions is not fixed and it is very likely that the stake system goes to two extreme cases (i.e., no one has enough stake or everyone has enough stake to create a valid block).

Two-Step Block Construction. To overcome the performance obstacles of DLT while taking supply chain management characters into consideration, we propose a novel two-step approach for block construction for CoC. The basic idea is to allow users to reserve blocks for near future usage based on their prediction, and then the users can use reserved blocks immediately when they are needed (as depicted in Fig.4).

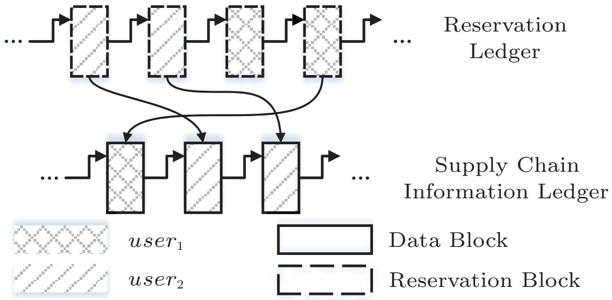


Fig.4. Two-step block generation. Before the user can put a supply chain record into the chain, it has to make a reservation in another chain. The reservation is confirmed by proof-of-work, i.e., someone has to complete a computation intensive task for a reservation. As showed in the figure, $user_1$ and $user_2$ reserve two blocks for their supply chain information in the reservation ledger respectively. $user_1$ uses one of his/her reservations and $user_2$ uses both. If $user_2$ wants to put more information to the supply chain information chain, he/she has to make extra reservations.

Specifically, the two-step block construction mechanism works as follows.

- *Step* 1: *Generation of Reservation Blocks.* When a user submits his/her reservation request to the system, the request is distributed to all helpers through gossip protocol[17]. Helpers who receive the request try to create a block through mining. Fig.5(a) depicts an example structure of the reservation block. For each block included in the reservation ledger, it contains the information of the user who wants to reserve the block, the fee the user wants to pay for the block, the identity of the helper who creates it, and other essential information. Note that all helpers have to reach a consensus on the reservation chain. Specifically, everyone checks whether a block is on the current longest-chain to determine whether to accept it or not. For a block b just added, helpers wait for a certain number of new blocks to be added after b. Satoshi proved that if six blocks are added after b, it is very likely that b is on the longest chain[10].

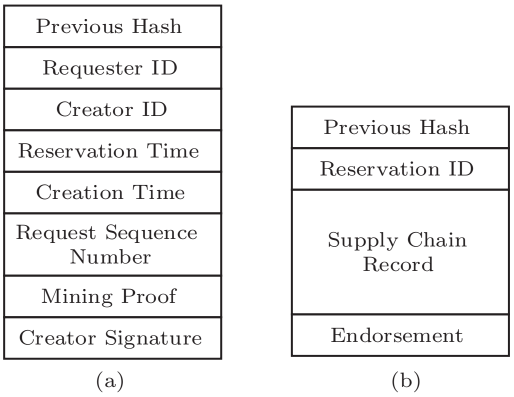


Fig.5. Two block structures for reservation and supply chain data. For data block, the field "supply chain record" is used to hold various kinds of information from order, payment, to bill of lading. (a) Reservation block. (b) Data block.

- *Step* 2: *Generation of Data Blocks.* When a user has one supply chain record that needs to be put into the distributed ledger holding real data, he/she first checks the reservation ledger to see whether he/she has available reservations for block generation. If he/she has an available reservation, a data block is constructed for supply chain record and the proof of reservation is included in the block. Fig.5(b) shows the structure of a data block. Putting this block in the data ledger does not require proof-of-work. When other peers receive the new block, they first check its validity: whether the block is properly constructed and whether attached reservation information is valid. If the new block passes all the checks, it is accepted and added to the ledger. The system also needs to achieve a consensus on all accepted blocks and different consensus protocols such as Paxos[18] can be used for this purpose. Note that if the record embedded in the new block involves multiple users, all of them need to sign the record to prevent faking information.

The two-step block construction method does not reduce the overall work load or latency compared with proof-of-work based approach. In fact, the work load and latency for the first step are very similar to those of classical proof-of-work based blockchain construction. But it provides a mechanism to shift the latency: as long as a user has enough reservation, the latency of adding a new supply chain record can be very low.

[10] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system, 2008. https://bitcoin.org/bitcoin.pdf, Jan. 2018.

For the reservation step, the latency is determined by both the demands (the number of reservation requests) and the supply (the number of reservations that can be generated in a given period). This is a typical supply-demand equilibrium problem. From supply perspective, reservation blocks are generated through mining, and by leveraging throughput scalable proof-of-work protocol[11][19], the supply increases when more helpers join the system. Because users pay for reservation, the market mechanism can automatically adjust the supply and demand of reservations.

The two-step block construction method can also be applied to other distributed ledger applications where the requirements are similar to those of the supply chain management system.

4.2 Storage Design of CoC

Supply chain management is in essence the management of corresponding information. Therefore, it is critical to have an efficient way to organize the information that is flexible enough to support various operations.

According to the design of CoC, it needs to maintain two decentralized ledgers: the reservation ledger (RL) and the data ledger (DL). Although decentralized ledger technology brings many useful features, it is not easy to manage them efficiently. The simplest approach to maintaining the two ledgers is to let everyone in CoC keep full copies of both of them. However, this is a waste of storage resources as different players in the system need different information. We design a more efficient storage scheme for CoC to manage the two ledgers, which allows different players to store ledgers in different ways.

Reservation Ledger Storage. The construction of reservation ledger involves users and helpers, where users submit reservation requests and helpers conduct mining to build blocks.

For helpers, they play a similar role as miners in Bitcoin system. But unlike Bitcoin and other cryptocurrency systems, blocks stored in RL are independent, i.e., when a new block is created, helpers do not need to check previous blocks to verify its validity. Therefore, a helper can keep headers of blocks instead of the whole blocks to reduce the storage cost, as depicted in Fig.6. With block headers, a helper can still determine which branch to follow by using the longest-chain principle, and check whether a given block is valid or not. But when a helper broadcasts a new block to be added to RL, he/she still needs to provide the complete block and thus other helpers can verify its validity.

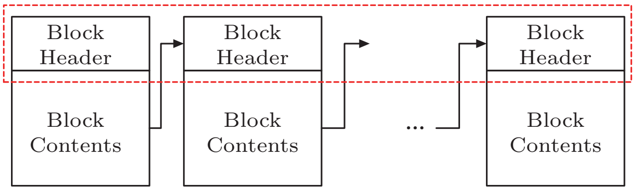


Fig.6. Helpers can store headers of blocks in the dotted box to reduce storage cost.

For users, they need to access RL for two purposes: 1) obtaining blocks containing their own reservation information to create new blocks in DL; 2) verifying whether blocks submitted to DL have a valid reservation. Therefore, a user can keep blocks that contain his/her own reservation and ignore other blocks on DL. To determine whether to accept a new block on DL, he/she can query helpers to check corresponding reservation block.

Data Ledger Storage. DL is used to store real supply chain information, and its construction relies on RL. As supporting entities and helpers do not need track supply chain information, they do not store blocks on DL. Third party users usually need to monitor supply chain information of different ordinary users, and thus they keep a full copy of DL and can serve as full nodes like in the Bitcoin system. For ordinary users, they only care about supply chain information related to them and keep these blocks contain such information. In addition to these blocks, they also store all headers of DL to facilitate adding new blocks to DL. Supporting entities can choose to store blocks related to them and headers of DL to verify the validity of other blocks.

Table 1 summarizes storage strategies for different types of participants.

Table 1. CoC Storage Strategies for Different Parties

Role	RL	DL
Helper	Headers of RL	NA
Ordinary user	His/her own reservation blocks	Related blocks and headers of DL*
Third party user	NA	Complete ledger
Supporting entity	NA	Related blocks and headers of DL*

Note: ∗: if ordinary users and supporting entities want to further reduce the storage cost, they can choose to trust third party users and discard all local storage related to DL.

[11] Luu L, Narayanan V, Baweja K, Zheng C D, Gilbert S, Saxena P. SCP: A computationally-scalable byzantine consensus protocol for blockchains. https://www.weusecoins.com/assets/pdf/library/SCP%20-%20%20A%20Computationally-Scalable%20Byzantine.pdf, Jan. 2018.

4.3 Identity Management of CoC

Decentralized ledgers used in Bitcoin and other fully open systems do not have a centralized identity management component, and each participant can generate his/her own credential, e.g., public/private key pair. However, the supply chain management scenario is not a complete open environment, and the participants are not equal and play different roles (as depicted in Fig.3). Therefore, it is necessary to have an identity management mechanism, and CoC uses "supporting entity" for this purpose.

"Helpers" are the largest group in CoC, and this group is usually quite dynamic and expensive to manage in a centralized way. Furthermore, helpers only contribute their computation resources to maintain CoC and there is no need to authenticate their identities. Therefore, CoC does not need to manage helpers, and they can generate their own public/private key pairs without notifying others. Their identities are used to receive rewards from users.

"Users" generate supply chain information and thus it is necessary to bind information with its creator. CoC uses a centralized identity management component (as part of supporting entities) to generate public/private key pairs for users and they use the keys to generate digital signatures for the information they submit to CoC to guarantee the authenticity/integrity. There are some on-going studies on building PKI with decentralized ledger[20-21], which can be used to replace a centralized identity management system in the future. For finical institutions that work as supporting entities, they maintain their own identity management system as they usually have their own standards and compliance requirements.

A centralized identity management does not mean that it has to be operated by a single entity. Multiple identity management systems can be integrated as long as they can collaborate with others. Besides using public/private key pair to identify a user, CoC also supports using biometrics for identity management.

4.4 Information Protection of CoC

When multiple companies are using CoC, their supply chain management related records are mixed and stored on the same distributed ledger. However, they do not want to disclose information to unrelated parties. To address this problem, encryption is used to protect supply chain management records on the ledger.

- *Record Encryption.* The creator of a record selects a random AES key *dek* to encrypt the record. It is the creator's responsibility to select adequate attributes of the record to encrypt and keep other parts in plain-text.
- *Authorizing Access.* The creator also creates a list of users/supporting entities, e.g., involved companies, government agencies, and financial institutions. By working together with the identity management component, the creator further encrypts *dek* with public keys of users/supporting entities in the list. Ciphertexts of *dek* can be stored together with the encrypted record on the distributed ledger as an evidence that the creator has allowed these access.

With this design, helpers and unrelated users/supporting entities are not able to learn useful information by observing the distributed ledger because they do not have access to the key *dek*. This approach is independent of the underlying decentralized ledger and can support flexible record level access control. If a group of records are shared with the same set of users/supporting entities, the creator can also use the same *dek* to avoid multiple time key distribution. Other encryption techniques that are used for secure data distribution can also be used, e.g., attribute encryption and proxy re-encryption[22-23].

5 Evaluation of CoC

In this section, we analyze the security of CoC, i.e., whether an attacker can alter historical data or insert fake data to the ledgers used by CoC.

5.1 Security Analysis

Security of RL. Since RL is built with proof-of-work, an attacker cannot alter historical data unless he/she controls more computation power than all honest helpers[12][24]. An attacker cannot insert reservations to RL without authorization from a user either because digital signature is used to issue a reservation request.

Security of DL. According to the design of CoC, only authorized users are allowed to add new blocks to RL, and it is more like permissioned ledger[25]. If a malicious user wants to alter an existing block, he/she needs to compete with all honest users for reservations. As reservation requests are not free, this is equivalent to the case that the malicious user pays more than all other users together.

Note that unlike cryptocurrency systems, it is easier for CoC to prevent invalid blocks because cryptocur-

[12] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system, 2008. https://bitcoin.org/bitcoin.pdf, Jan. 2018.

rency system handles purely information in cyber space but CoC has connections with physical world. For example, if user A wants to add a block indicating that he/she has transferred a container to user B, the block must be signed by both of them and they do not need to scan previous blocks to check whether this activity is valid.

5.2 Performance Analysis

Latency. Fig.7 demonstrates the relationship between latencies of making reservation and generating a data block. As long as $t_2 \leqslant t_3$, the latency of adding a new data block $(t_4 - t_3)$ is independent of the latency of making a reservation $(t_2 - t_1)$.

The major factors that affect the value of $(t_4 - t_3)$ are as follows.

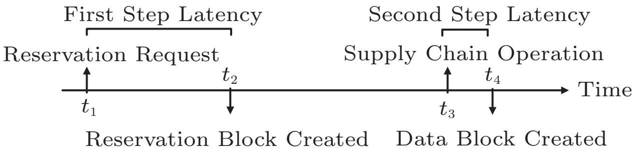


Fig.7. Latency of two-step block generation.

- *Latency to Verify a Block.* When receiving a data block, the user needs to verify whether it is valid or not. The verification is further divided into two parts: verifying the block and verifying the reservation. The first operation only involves the verification of digital signatures, and is not a problem for modern computers. To verify the reservation, the user needs to query helpers who maintain headers of the reservation ledger, which is also very cheap.

- *Latency to Achieve Consensus.* Because the second step uses classical consensus protocol such as BFT protocol, the latency to achieve consensus is much lower than that to use use proof-of-work and longest-chain[24].

Throughput. The throughput of CoC is determined by the minimal throughput of the reservation ledger and the data ledger. Because the data ledger uses BFT protocol, it can achieve very high throughput[26–27]. For the reservation ledger generated by proof-of-work, there are many techniques available to improve its throughput such as using larger block size to hold more reservation requests[28] and divide-and-conquer strategy to make it scalable[13].

Note that the two-step block construction allows the reservation ledger to focus on throughput improving without considering latency too much. This is much easier than improving both of them.

5.3 Experimental Results

As we discussed in above subsections, the performance of CoC is determined by the second step of block construction. Therefore, we focus on the performance of this step. We implement key components of CoC using code base of Hyperledger Fabric[29], and utilize practical BFT for the second step of block construction[30], where users (e.g., factories and transportation companies) submit their records to the system.

Because helpers scatter around the globe, it is better to conduct the experiments using machines in different physical locations. Therefore, we use Amazon cloud as the testbed and its machines are from multiple data centers. Specifically, we use EC2 t2.micro instances running Ubuntu 14.04 and each instance has one CPU core and 1 GB memory. All instances are evenly distributed in four data centers located in California, Virginia, Ohio, and London respectively. Fig.8 shows the latency of the second step of block construction with different numbers of users in different locations, and Fig.9 shows the throughput. The latency of the second step is roughly linear to the number of helpers in the system. When there are 100 helpers in the system, we achieve a latency about 16 seconds, which is much better than purely proof-of-work based system like Bitcoin. It is unsurprisingly that the throughput is in inverse proportion to the latency. When we have 100

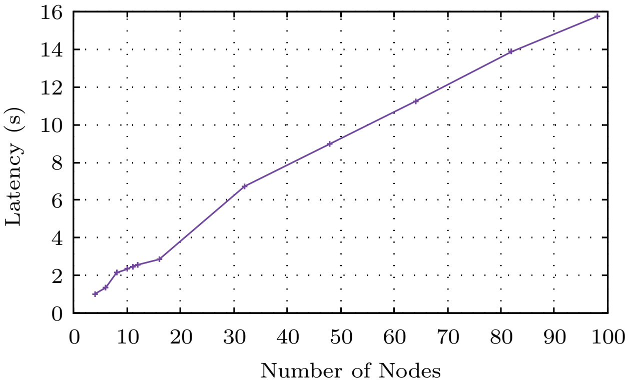


Fig.8. Latency of the second step of block construction with different numbers of helpers. These helpers reside in different Amazon data centers.

[13] Luu L, Narayanan V, Baweja K, Zheng C D, Gilbert S, Saxena P. SCP: A computationally-scalable byzantine consensus protocol for blockchains. https://www.weusecoins.com/assets/pdf/library/SCP%20-%20%20A%20Computationally-Scalable%20Byzantine.pdf, Jan. 2018.

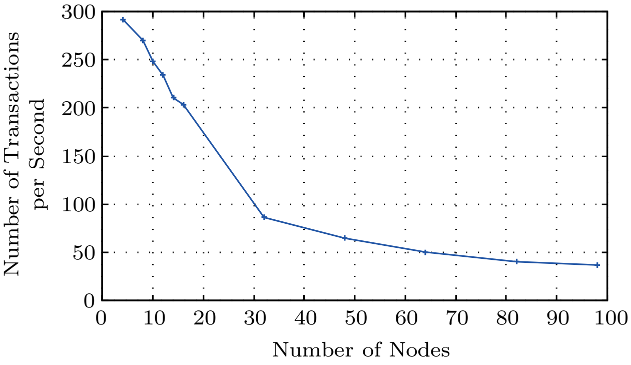


Fig.9. Throughput of the second step of block construction with different numbers of helpers. These helpers reside in different Amazon data centers.

helpers, the system can process about 40 transactions in one second. Note that a transaction can be a block with multiple records, thereby if we put 10 records in a single block, the system can process 400 records in one second. Using a larger block size can further improve the throughput.

6 Related Work

In this section, we briefly review related work.

Using DLT for Supply Chain Management. Korpela *et al.* noticed that blockchain technology offers a public model to connect different stakeholders and provided a set of factors that affect the adoption of such system[31]. Tian[32] proposed a design of agri-food supply chain that combines RFID and decentralized ledger. This work mentioned some performance limitations of blockchain but did not give any solution[32]. IBM also introduced its blockchain-based supply chain management system and blockchain-based bill of lading system on top of the Hyperledger project, which is a purely permissioned decentralized ledger platform[29]. There are other studies along this direction[9-10,33][14]. However, most of these studies ignore the limitations of distributed ledger and just use it as a storage mechanism to replace existing file system/database.

DLT Performance. Distributed ledger technology finds various applications in different sectors, and many efforts have been spent on improving its performance. One direction is to replace proof-of-work/longest-chain with Byzantine fault tolerant protocols[24], which works well in a closed environment but not suitable for supply chain management. Trusted computing technology is also used for distributed ledger construction[34-35]. This approach achieves high throughput and low latency at the same time but requires special hardware that supports trusted computing.

7 Conclusions

Supply chain management plays an important role in the modern economy, especially when business entities are more dependent on each other. CoC leverages the emerging distributed ledger technology to build a unified supply chain management system, and uses a series of novel techniques to overcome the limitations of distributed ledger, including the two-step block construction method under hybrid model, efficient ledger storage, and information protection. Besides the basic cargo tracing capability, CoC can support various supply chain management tasks such as bill of lading, international trade compliance, and customs clearance. We also analyzed the security and performance of CoC to show that it satisfies the major requirements of supply chain management. For the next step, we plan to keep improving the prototype and evaluate its effectiveness in production environment.

References

[1] Laseter T, Oliver K. When will supply chain management grow up? *Strategy + Business*, 2003. https://www.strategy-business.com/article/03304, Jan. 2018.

[2] Dai J, Ding Z M, Xu J J. Context-based moving object trajectory uncertainty reduction and ranking in road network. *Journal of Computer Science and Technology*, 2016, 31(1): 167-184.

[3] Liu H L, Chen Q, Li Z H. Optimization techniques for RFID complex event processing. *Journal of Computer Science and Technology*, 2009, 24(4): 723-733.

[4] He W, Tan E L, Lee E W, Li T Y. A solution for integrated track and trace in supply chain based on RFID & GPS. In *Proc. IEEE Conf. Emerging Technologies & Factory Automation*, September 2009.

[5] Lindner M, Marquez F G, Chapman C, Clayman S, Henriksson D, Elmroth E. The cloud supply chain: A framework for information, monitoring, accounting and billing. In *Proc. the 2nd Int. ICST Conf. Cloud Computing*, October 2010.

[6] Gazet A. Comparative analysis of various ransomware virii. *Journal in Computer Virology*, 2010, 6(1): 77-90.

[7] Garay J, Kiayias A, Leonardos N. The Bitcoin backbone protocol: Analysis and applications. In *Proc. the 34th Annual Int. Conf. the Theory and Applications of Cryptographic Techniques*, April 2015, pp.281-310.

[8] Lemieux V L. Trusting records: Is blockchain technology the answer? *Records Management Journal*, 2016, 26(2): 110-139.

[14] Parker L. Blockchain tech companies focus on the $40 trillion supply chain market, 2016. https://bravenewcoin.com/news/blockchain-tech-companies-focus-on-the-40-trillion-supply-chain-market/, Jan. 2018.

[9] Morabito V. Blockchain practices. In *Business Innovation Through Blockchain: The B^3 Perspective*, Morabito V (ed.), Springer, 2017, pp.145-166.

[10] Lehmacher W. Global dynamics and key trends. In *The Global Supply Chain: How Technology and Circular Thinking Transform Our Future*, Lehmacher W (ed.), Springer, 2017, pp.67-112.

[11] Cooper M C, Lambert D M, Pagh J D. Supply chain management: More than a new name for logistics. *The International Journal of Logistics Management*, 1997, 8(1): 1-14.

[12] Fox M S, Chionglo J F, Barbuceanu M. The integrated supply chain management system. Technical Report, Department of Industrial Engineering, University of Toronto, 1993.

[13] Buterin V. What proof of stake is and why it matters. *Bitcoin Magazine*, 2013. https://bitcoinmagazine.com/articles/what-proof-of-stake-is-and-why-it-matters-1377531463, Jan. 2018.

[14] Xu X W, Pautasso C, Zhu L M, Gramoli V, Ponomarev A, Tran A B, Chen S P. The blockchain as a software connector. In *Proc the 13th Working IEEE/IFIP Conf. Software Architecture*, April 2016, pp.182-191.

[15] Castro M, Liskov B. Practical Byzantine fault tolerance and proactive recovery. *ACM Trans. Computer Systems*, 2002, 20(4): 398-461.

[16] Lamport L, Shostak R, Pease M. The Byzantine Generals Problem. *ACM Trans. Programming Languages and Systems*, 1982, 4(3): 382-401.

[17] Kermarrec A M, van Steen M. Gossiping in distributed systems. *ACM SIGOPS Operating Systems Review*, 2007, 41(5): 2-7.

[18] Lamport L. The part-time parliament. *ACM Trans. Computer Systems*, 1998, 16(2): 133-169.

[19] Eyal I, Gencer A E, Sirer E G, van Renesse R. Bitcoin-NG: A scalable blockchain protocol. In *Proc. the 13th USENIX Conf. Networked Systems Design and Implementation*, March 2016, pp.45-59.

[20] Lewison K, Corella F. Backing rich credentials with a blockchain PKI. Technical Report, Pomcor, 2016. https://pomcor.com/techreports/BlockchainPKI.pdf, Jan. 2018.

[21] Al-Bassam M. SCPKI: A smart contract-based PKI and identity system. In *Proc. the ACM Workshop on Blockchain Cryptocurrencies and Contracts*, April 2017, pp.35-40.

[22] Xu L, Wu X X, Zhang X W. CL-PRE: A certificateless proxy re-encryption scheme for secure data sharing with public cloud. In *Proc. the 7th ACM Symp. Information Computer and Communications Security*, May 2012, pp.87-88.

[23] Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute-based encryption. In *Proc. IEEE Symp. Security and Privacy*, May 2007, pp.321-334.

[24] Vukolić M. The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication. In *Proc. the Int. Workshop on Open Problems in Network Security*, October 2015, pp.112-125.

[25] Vukolić M. Rethinking permissioned blockchains. In *Proc. the ACM Workshop on Blockchain Cryptocurrencies and Contracts*, April 2017, pp.3-7.

[26] Guerraoui R, Knežević N, Quéma V, Vukolić M. The next 700 BFT protocols. In *Proc. the 5th European Conf. Computer Systems*, April 2010, pp.363-376.

[27] Kotla R, Dahlin M. High throughput Byzantine fault tolerance. In *Proc. Int. Conf. Dependable Systems and Networks*, July 2004, pp.575-584.

[28] Croman K, Decker C, Eyal I, Gencer A E, Juels A, Kosba A, Miller A, Saxena P, Shi E, Sirer E G, Song D, Wattenhofer R. On scaling decentralized blockchains. In *Proc. Int. Conf. Financial Cryptography and Data Security*, February 2016, pp.106-125.

[29] Cachin C. Architecture of the hyperledger blockchain fabric. In *Proc. the Workshop on Distributed Cryptocurrencies and Consensus Ledgers*, July 2016.

[30] Wood T, Singh R, Venkataramani A, Shenoy P, Cecchet E. ZZ and the art of practical BFT execution. In *Proc. the 6th Conf. Computer Systems*, April 2011, pp.123-138.

[31] Korpela K, Hallikas J, Dahlberg T. Digital supply chain transformation toward blockchain integration. In *Proc. the 50th Hawaii Int. Conf. System Sciences*, Jan. 2017, pp.4182-4191.

[32] Tian F. An agri-food supply chain traceability system for China based on RFID & blockchain technology. In *Proc. the 13th Int. Conf. Service Systems and Service Management*, June 2016, pp.1-6.

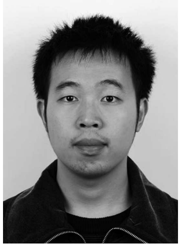
[33] Abeyratne S A, Monfared R P. Blockchain ready manufacturing supply chain using distributed ledger. *International Journal of Research in Engineering and Technology*, 2016, 5(9): 1-10.

[34] Milutinovic M, He W, Wu H, Kanwal M. Proof of luck: An efficient blockchain consensus protocol. In *Proc. the 1st Workshop on System Software for Trusted Execution*, Dec. 2016, Article No. 2.

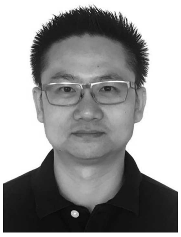
[35] Intel. Blockchain and its emerging role in healthcare and health-related research. Technical Report 4150-45-P, 2016. https://s3.amazonaws.com/public-inspection.federalregister.gov/2016-16133.pdf, Jan. 2018.

Zhimin Gao received his B.S. degree in software engineering from South China Agricultural University, Guangzhou, in 2009, and his Ph.D. degree in computer science from University of Houston, Houston, in 2017. He is currently working as a post-doctoral fellow at University of Houston, Houston. His research interests include blockchain, high-performance computing and cloud computing.

Lei Xu received his B.S. degree in applied mathematics from Hebei University, Baoding, in 2004, and his Ph.D. degree in computer science from Institute of Software, Chinese Academy of Sciences, Beijing, in 2011. He is currently a research assistant professor at University of Houston, Houston. From 2011 to 2013, he worked as a research engineer at the Central Research Institute, Huawei Technologies Co. Ltd., Beijing. His research interests include blockchain, cloud security, and applied cryptography.



Lin Chen received his Ph.D. degree in computer science from Zhejiang University, Hangzhou, in 2013. From 2013 to 2016, he worked as a post-doctoral fellow at Technical University of Berlin, Berlin, and then Hungarian Academy of Science, Budapest. He is currently a research assistant professor at University of Houston, Houston. His research interests include blockchain, stochastic optimization, parameterized algorithms and complexity.



Xi Zhao received his Ph.D. (Hons.) degree in computer science from the Ecole Centrale de Lyon, France, in 2010. After graduation, he conducted research in the fields of biometrics, face analysis, and pattern recognition, as a research assistant professor with the Department of Computer Science, University of Houston, Houston. He is currently an associate professor in School of Management, Xi'an Jiaotong University, Xi'an. His research interests include big data analytics, mobile computing, and computational social science.

Yang Lu received her M.S. degree in information technology from Southern Polytechnic State University, Atlanta. Her research is focused on computing security and blockchain. She previously worked as a software engineer at File-Vison LLC. Currently, she works as project manager in the Department of Computer Science at the University of Houston, Houston.

Weidong Shi received his Ph.D. degree in computer science from Georgia Institute of Technology, Georgia, where he did research in computer architecture and computer systems. He was previously a senior research staff engineer at Motorola Research Lab, Nokia Research Center, and co-founder of a technology startup. Currently, he is employed as an associate professor by University of Houston, Houston.