Editor: **Cesare Pautasso**
University of Lugano
c.pautasso@ieee.org

Editor: **Olaf Zimmermann**
University of Applied Sciences
of Eastern Switzerland, Rapperswil
ozimmerm@hsr.ch

Adaptable Blockchain-Based Systems

A Case Study for Product Traceability

Qinghua Lu and Xiwei Xu

From the Editors
Tracing the origin of products across complex supply chains requires a transparent, tamper-proof metadata infrastructure that's not only trusted by all the involved parties but also adaptable to changing environments and regulations. Can such advanced infrastructure be implemented in a decentralized way? Qinghua Lu and Xiwei Xu share their story of developing the originChain system, which leverages emerging blockchain technology to do so. —*Cesare Pautasso and Olaf Zimmermann*

A TRACEABILITY SYSTEM enables tracking products by providing information about them (for example, originality, components, or locations) during production and distribution. Product suppliers and retailers usually require independent *traceability service providers* who are government-certified to inspect the products throughout the supply chain. If everything satisfies the requirements, the traceability service providers issue inspection certificates that verify the products' quality and originality. To expose information and generate certificates, these service providers employ a traceability system.

In this context, security is important for accountability and forensic information. Traceability systems normally store information in conventional databases controlled by the service providers. Such centralized data storage becomes a potential single point of failure and runs the risk of tampering.

One of our partners is an independent third-party traceability service provider certified by the Chinese government. Its system provides traceability information for products imported to China. This system has been integrated with several big Chinese e-commerce websites (for example, JD.COM). Hundreds of product suppliers and retailers use its traceability services to manage their products' traceability information, and millions of product consumers use it to access the information.

Here, we share our experience of building originChain. It restructures the service provider's current traceability system by replacing the central database with a blockchain. (For more on blockchains, see the two sidebars.) OriginChain provides transparent tamper-proof traceability data, enhances the data's availability, and automates regulatory-compliance checking. We implemented originChain and tested it under realistic conditions

0740-7459/17/$33.00 © 2017 IEEE

employing the users' traceability information. We're planning how to replace the existing system with the restructured one.

BLOCKCHAINS AND SMART CONTRACTS

As a data structure, a blockchain is an ordered list of blocks that contain transactions such as monetary transfers and smart-contract creation and invocation.[1] Each block contains a hash of the previous block's representation, thus creating the chain. So, historical transactions in the blockchain can't be deleted or altered without invalidating the chain of hashes. Combined with computational constraints and incentive schemes for block creation, this can prevent the tampering with and revision of information in the blockchain.

Blockchains also provide a general-purpose programmable infrastructure. Programs can be deployed and run on a blockchain; such programs are called *smart contracts*.[2] The result of a smart-contract invocation is stored in public data storage. Smart contracts can express triggers, conditions, and business logic to enable more complex programmable transactions. Hence, smart contracts differ from service contracts in service-oriented computing, which are interfaces between services and consumers so that they can successfully interact. A common simple example of a smart-contract-enabled service is escrow, which holds funds until the obligations defined in the smart contract are fulfilled.

PROPERTIES

Any data in a committed transaction eventually becomes immutable. The immutable chain of cryptographically signed historical transactions provides nonrepudiation of the stored data. Cryptographic tools also support data integrity, the public access provides data transparency, and every participant has potentially the same ability to access and manipulate the blockchain. However, those rights can be weighted by the participants' stake or computational power. To facilitate transactions, the participants rely on the blockchain network itself instead of trusted third-party organizations.

LIMITATIONS

Blockchains lack data privacy; there are no privileged users, and, as we just mentioned, every participant can access all the information on the blockchain. In addition, public blockchains have limits on the amount of data, transaction processing rate, and data transmission latency. *Consortium blockchains*, in which the consensus process is limited to several participants, perform much better. However, developers still must consider these factors when designing systems.

Furthermore, some public blockchains use a *proof-of-work* consensus mechanism that "wastes" significant electricity because it doesn't lead directly to a successful solution. Researchers are developing alternative consensus mechanisms for public blockchains. One example is the *proof-of-stake* mechanism, which isn't computationally expensive. Consortium and private blockchains also often use consensus mechanisms that don't rely on proof of work.

References

1. M. Swan, *Blockchain: Blueprint for a New Economy*, O'Reilly Media, 2015.
2. S. Omohundro, "Cryptocurrencies, Smart Contracts, and Artificial Intelligence," *AI Matters*, vol. 1, no. 2, 2014, pp. 19–21.

Dealing with Traceability

Product suppliers and retailers apply for traceability services for different purposes. Suppliers want to receive certificates to show their products' origin and quality to consumers and to comply with regulations. Retailers want verification of the products' origin and quality.

Each product supplier that uses our partner's services has on average 20 products to be traced. The traceability information's granularity is rather large because it corresponds to product packages rather than individual products. This information's size isn't easy to estimate because many documents currently aren't digitized, such as certificates issued by traceability service providers.

Traceability is flexible. Figure 1 shows a simplified possible process in BPMN (Business Process Model and Notation).[1] In the real world, the sequence of some activities in Figure 1 (such as "examine factory" and "test sample") is dynamic owing to customization of the quality control and inspection processes. The labs that test samples must adapt to the labs' availability and the characteristics of the products, such as powdered milk or clothing materials.

Furthermore, regulatory-compliance checking can change because of new regulations. For example, China's

BLOCKCHAINS IN THE SUPPLY CHAIN

The supply chain is a promising area for applying blockchains.[1] There are blockchain startups in this field. For example, Everledger (www.everledger.io) uses blockchains to track diamonds' features, such as cut and quality, and to help reduce risk and fraud for banks, insurers, and open marketplaces.

Big enterprises are also applying blockchains in supply chains for different domains. For example, in January 2017, Microsoft started the Manifest project through a partnership with Mojix to leverage an Internet-of-Things platform with a blockchain to help factories, distribution centers, and retailers track goods using RFID devices.[2] In May 2017, Manifest grew to 13 partners.[3] BHP Billiton has been exploring blockchain technology to track movements of wellbore rock and fluid samples and secure the real-time data generated during the samples' delivery.[4]

References

1. M. Staples et al., *Risks and Opportunities for Systems Using Blockchain and Smart Contracts*, Commonwealth Scientific and Industrial Research Org., 2017.
2. M. del Castillo, "Microsoft Unveils Project Manifest, a Plan for Blockchain Product Tracking," CoinDesk, 25 Jan. 2017; www.coindesk.com/microsoft-unveils-project-manifest-a-plan-for-product-tracking-via-blockchain.
3. M. del Castillo, "Microsoft's Blockchain Supply Chain Project Grows to 13 Partners," CoinDesk, 3 May 2017; www.coindesk.com/microsofts-blockchain-supply-chain-project-grows-to-13-partners.
4. P. Rizzo, "World's Largest Mining Company to Use Blockchain for Supply Chain," CoinDesk, 23 Sept. 2016; www.coindesk.com/bhp-billiton-blockchain-mining-company-supply-chain.

Food Safety Law, which took effect in October 2015, set out new requirements for formulating national food safety standards and traceability systems (specifying what information should be provided). So, adaptability was one of our main concerns when we designed originChain.

OriginChain's Architecture

Figure 2 illustrates originChain's architecture. OriginChain currently employs a geographically distributed private blockchain at the traceability service provider, which has branch offices in three countries. The plan is to establish a trustworthy traceability platform that covers other organizations, including government-certified labs, big suppliers, and retailers that have long-term relationships with the company (such as e-commerce companies that have already built their reputation among customers). Compared to a public blockchain, such a *consortium blockchain* can perform better and cost less.

Blockchains grow continually because the data and code on them are immutable. So, a major design decision is to choose what data and computation to keep on-chain and off-chain. We discuss this in more detail later.

As Figure 2 shows, product suppliers or retailers manage product or enterprise information through the *product and enterprise management* module. They access the information on the blockchain through a webserver hosted by originChain. In the future, suppliers and big retailers that host a node by themselves will be able to access their own nodes to obtain information on the blockchain.

After the traceability service provider validates an application from a product supplier or retailer on the basis of paperwork (see Figure 1), the two parties sign a legal agreement about which traceability services are covered. OriginChain generates a *smart contract* that represents the legal agreement. (For more information on smart contracts, see the sidebar "Blockchains and Smart Contracts.") The smart contract codifies the combination of services and other conditions defined in the agreement. So, the smart contract can automatically check and enforce these conditions. It also checks whether all the information required by regulation is provided, to enable automated regulatory-compliance checking.

The traceability service provider manages traceability information, certificates, and onsite photos using the *traceability management* module. Because of the blockchain's limited data storage, originChain stores two types of data on-chain as variables of smart contracts:

- the hash of traceability certificates or photos and

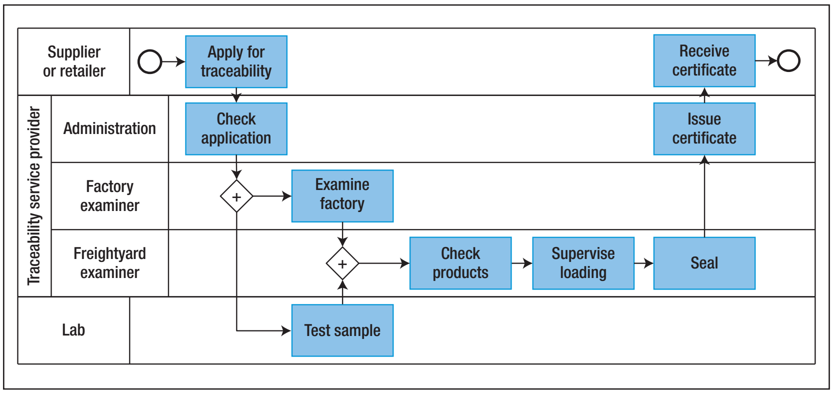

FIGURE 1. The dynamism of the traceability process. The sequence of some activities (such as “examine factory” and “test sample”) is dynamic owing to customization of the quality control and inspection processes.

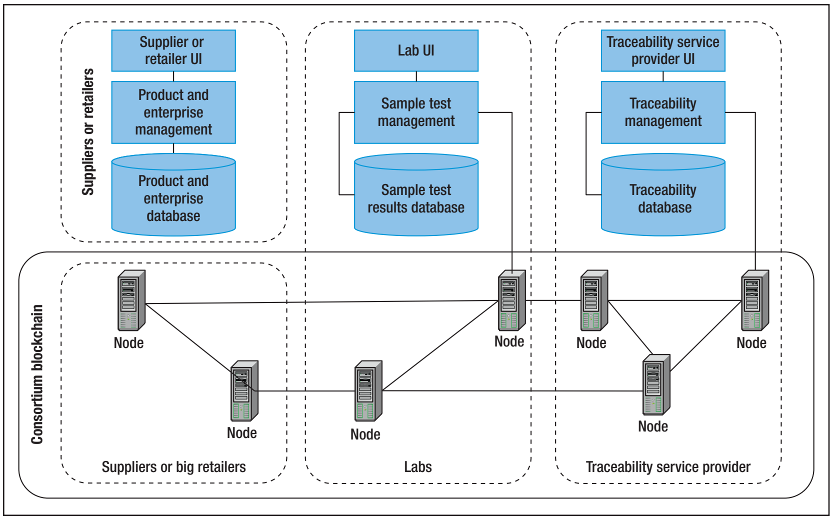
Suppliers or retailers
Supplier or retailer UI
Product and enterprise management
Product and enterprise database
Lab UI
Sample test management
Sample test results database
Traceability service provider UI
Traceability management
Traceability database
Consortium blockchain
Node
Node
Node
Node
Node
Node
Node
Suppliers or big retailers
Labs
Traceability service provider

FIGURE 2. OriginChain’s architecture. OriginChain currently employs a geographically distributed private blockchain at the traceability service provider company. The plan is to establish a consortium blockchain, which will include other organizations.

- the small amount of traceability information required by the traceability regulation, such as the batch number, traceability results, place of origin, and inspection date.

The raw files of traceability certificates and photos (.pdf or .jpg) and the addresses of the smart contracts are off-chain in a centralized MySQL database hosted by originChain. Other partners can still maintain their own database of product information (for the suppliers or retailers) or other numbers shown in the sample testing (for the labs).

The labs manage sample-testing results through the *sample test management* module. A blockchain’s execution environment is self-contained. So, a smart contract can access only information stored in the blockchain. It can’t directly access the states of external systems (for example, testing results and product geolocations). Thus, the labs periodically inject the result of sample testing from the external world into the blockchain.

The information of blockchain-layer permission control (for example, permission for content management, for writing smart contracts, or for joining a consortium blockchain) can be on-chain or off-chain. However, an off-chain centralized permission management module could become a single point of failure from both an operational and a management perspective. So, originChain stores the control information, such as permission to join the blockchain network (to own a copy of all the historical transactions). On-chain permission management leverages the blockchain’s decentralized nature so that all the participants can access the blockchain.

Figure 3 shows how smart contracts are designed using our blockchain; Figure 4 shows part of the related pseudocode. In originChain, a *factory contract* creates smart contracts. This reduces the complexity of creating customized smart contracts. The factory contract contains code fragments representing different traceability services. The generation of smart contracts requires authority from both the traceability service provider and the supplier or retailer.

When the factory contract is called, it creates two kinds of smart contract: a *registry contract* and *service contract*. The registry contract represents the legal agreement and contains the address of the service contract, which codifies the legal agreement. The service contract could be updated by replacing its address stored in the registry contract with the address of a new version. Possible updates include adding or removing services from the legal agreement after the initial legal agreement is signed, or selecting labs for sample testing on the basis of their availability. The registry contract specifies a list of addresses allowed to update the registry contracts, and a threshold of the minimum number of addresses required to authorize an update.

If a testing sample involves multiple labs for crosschecking, signatures from all the labs are required before the traceability application can undergo further processing. To enable more dynamic lab selection, users can employ an *M*-of-*N* multisignature to define that *M* out of *N* labs are required to authorize the testing results.

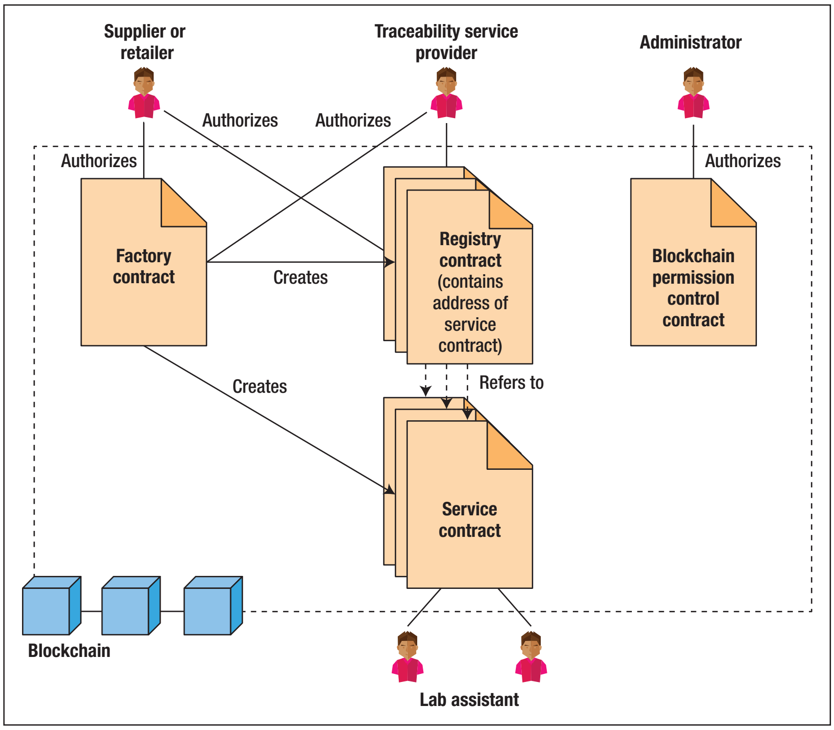

FIGURE 3. Designing smart contracts in originChain. The factory contract reduces the complexity of creating customized smart contracts.

Lessons Learned

Our experiences with originChain led to the following insights.

The Design of Blockchain-Based Systems

Owing to blockchains' unique properties, some design considerations are specific to blockchain-based applications—for example, the consideration of on-chain and off-chain. On the other hand, because smart contracts are programs running on a blockchain, some existing architectural patterns might be applicable to them. From the business process perspective, approaches such as model-driven development[2] and behavior-driven development[1] are also applicable.

The smart contract's structural design has a large impact on the cost if the blockchain is public. The contract's deployment cost depends on its size because the code is stored in the blockchain, which entails data storage fees proportional to the contract's size. So, a structural design with more lines of code costs more money.

A consortium blockchain doesn't have to have a token or currency, so monetary cost isn't an issue there. However, the blockchain's size is still a design concern because it grows with every transaction and because every participant has a replica of the whole blockchain. In addition, a more structural design might affect performance because it might require more transactions.

On-Chain vs. Off-Chain

Regarding what should be on-chain and off-chain, two factors are important: performance and privacy. Performance depends highly on the blockchain's deployment. For example, a consortium blockchain can be configured to perform much better than a public blockchain.

With originChain, owing to the current traceability system's characteristics (for example, it has low writing throughput because of the large granularity of traceability information), a blockchain's limited throughput isn't the main concern. However, as we mentioned before, all the participants can access the data on the blockchain. So, private data (for example, customer information) shouldn't be on-chain. Regarding traceability, large sensitive raw data (for example, traceability certificates and photos) must be immutable. Thus, the raw data is off-chain, whereas its hash is on-chain.

```
contract FactoryContract {
     address[] registryContracts;
     address[] serviceContracts;
     // deploy a new registry contract
     function newRegistryContract() returns(address
newRegistryContract){
          RegistryContract r = new RegistryContract();
          registryContracts.push(r);
          return r;
     }
     // deploy a new service contract
     function newServiceContract() returns(address
newServiceContract){
          ServiceContract s = new ServiceContract();
          serviceContracts.push(s);
          return s;
     }
}

contract RegistryContract{
     address[] serviceContracts;
     address[] authorities;
     uint threshold;
     uint authorityNum;
     function authorizeVotingRight(address authority) {
          authorities.push(authority);
     }
     function setThreshold(uint threshold) {
          threshold = threshold;
     }
     function vote(address authority) {
          authorityNum++;
     }
     function update() returns(address newServiceContract){
          // If there are enough authorities then
          …
          If(authorityNum >= threshold){
               ServiceContract s = new ServiceContract();
               serviceContracts.push(s);
               return s;
          }
     }
}

contract ServiceContract {
     bool testsampleselected;
     bool examinefactoryselected;
     bool superviseloadingselected;
     function ServiceContract (bool testsample, bool examinefactory,
bool superviseloading){…}
     function testSample(){…}
     function examineFactory(){…}
     function superviseLoading(){…}
```

FIGURE 4. Pseudocode for the design of a smart contract in originChain. The generation of smart contracts requires authority from both the traceability service provider and the supplier or retailer.

The Adaptability of Blockchain-Based Systems

Adaptability is a quality attribute required by many industrial projects that are inherently dynamic. For example, changes to the legal agreement or new regulations could necessitate adaptation. Adaptation here means that the smart contract could be updated by a number of authorities above the threshold defined in the factory contract. However, research on blockchain-based systems rarely discusses adaptability.

We view the blockchain as a component of a larger distributed system. In originChain, we implement some of the business logic on-chain as smart contracts. Thus, smart contracts' structural design also affects their updatability and the whole system's adaptability (for example, separation between data and control).

However, if the blockchain is for data storage only, not much can be done to affect the whole system's adaptability. Moving some logic to the blockchain can leverage the trustworthiness (and the interoperability and the transparency of data and operations) that the blockchain provides as a computational platform. In addition, the data in smart contracts is easier to query (directly on the blockchain) than is the data in transactions.[3]

Access Control for Smart Contracts

Smart contracts running on a blockchain can be accessed and called by all the participants. A smart contract, by default, has no owner; once it's deployed, its author no longer has any special privileges on it. Unauthorized users could accidentally trigger a permissionless function. So, smart contracts should have an embedded permission control mechanism to check permission for every caller that triggers the functions defined in the contracts.

Traceability processes in supply chain management are complex and dynamic because they involve multiple parties. A blockchain provides neutral ground that should help integrate the disparate participants into those processes. Also, the integrity and audit trail in a blockchain ledger should improve transparency and confidence across the processes.

Although joining a consortium blockchain benefits all the relevant stakeholders, adopting a new technique such as a blockchain is always a challenge to traditional industries because of the learning curve and the cost of integrating the blockchain into the existing systems. Negotiating the business details also takes time. In addition, the development of smart contracts must take into account quality attributes such as adaptability.

Data transparency and sharing data with others are main concerns for most companies that provide intermediary services in industries. Overall, blockchains are a good option for providing traceability in supply chain management. Nevertheless, industry needs to take the time to understand their risks and opportunities.

ABOUT THE AUTHORS

QINGHUA LU is an associate professor in the College of Computer and Communication Engineering at China University of Petroleum (East China). Contact her at qinghualu@upc.edu.cn.

XIWEI XU is a research scientist in Data61 at the Commonwealth Scientific and Industrial Research Organisation (CSIRO). Contact her at xiwei.xu@data61.csiro.au.

Acknowledgments

The National Natural Science Foundation of China (grant 61402533) and the Fundamental Research Funds for the Central Universities (grant 16CX02047A) supported this project.

References

1. D. Lübke and T. van Lessen, “Modeling Test Cases in BPMN for Behavior-Driven Development,” *IEEE Software*, vol. 33, no. 5, 2016, pp. 15–21.
2. I. Weber et al., “Untrusted Business Process Monitoring and Execution Using Blockchain,” *Business Process Management*, LNCS 9850, Springer, 2016, pp. 329–347.
3. X. Xu et al., “A Taxonomy of Blockchain-Based Systems for Architecture Design,” *Proc. 2017 IEEE Int’l Conf. Software Architecture* (ICSA 17), 2017, pp. 243–252.