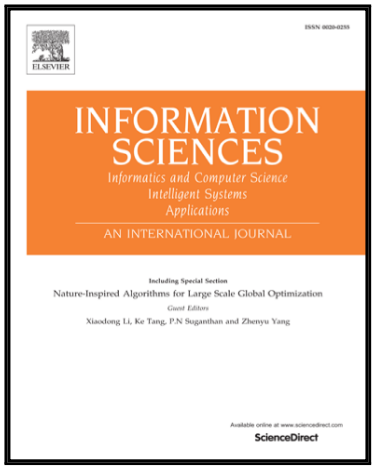
Journal Pre-proof



Blockchain-based fair payment smart contract for public cloud storage auditing

Hao Wang, Hong Qin, Minghao Zhao, Xiaochao Wei, Hua Shen, Willy Susilo

PII: S0020-0255(20)30062-1
DOI: https://doi.org/10.1016/j.ins.2020.01.051
Reference: INS 15182

To appear in: *Information Sciences*

Received date: 28 August 2019
Revised date: 27 January 2020
Accepted date: 29 January 2020

Please cite this article as: Hao Wang, Hong Qin, Minghao Zhao, Xiaochao Wei, Hua Shen, Willy Susilo, Blockchain-based fair payment smart contract for public cloud storage auditing, *Information Sciences* (2020), doi: https://doi.org/10.1016/j.ins.2020.01.051

Blockchain-based fair payment smart contract for public cloud storage auditing

Hao Wang[a,b], Hong Qin[a], Minghao Zhao[c], Xiaochao Wei[a], Hua Shen[d], Willy Susilo[b]

[a]*School of Information Science and Engineering, Shandong Normal University, China*
[b]*School of Computing and Information Technology, University of Wollongong, Australia*
[c]*School of Software, Tsinghua University, China*
[d]*School of Computers, Hubei University of Technology, China*

Abstract

Cloud storage plays an important role in today's cloud ecosystem. Increasingly clients tend to outsource their data to the cloud. In spite of its copious advantages, integrity has always been a significant issue. The audit method is commonly used to ensure integrity in cloud scenarios. However, traditional auditing schemes expect a third-party auditor (TPA), which is not always available in the real world. Also, the former scheme implies a limited pay-as-you-go service, as it requires the client to pay for the service in advance.

In this paper, we aim to address the aforementioned drawback by adopting blockchain to replace TPA and designing a blockchain-based fair payment smart contract for public cloud storage auditing. In our system, data owner and cloud service provider (CSP) will run a blockchain-based smart contract. The contract ensures that the CSP is required to submit data possession proof regularly. The CSP gets paid only if the verification is passed; otherwise, it gets no remuneration but has to pay the penalties. To reduce the number of interactions in the execution of contract, we present the notion of non-interactive public provable data possession and design a blockchain-based smart contract for public cloud storage auditing based on this primitive.

Keywords: blockchain, smart contract, public auditing, fair payment, cloud storage

1. Introduction

In cloud storage services, the cloud service provider (CSP) offers the clients with on-demand storage services, either in the form of IssS (*e.g.,* Amazon AWS S3 and Google Cloud Storage) or SaaS (*e.g.,* iCloud and Dropbox). Until now, many pieces of research have been devoted to efficiency, reliability, and user-friendliness aspects of cloud storage services (*e.g.,* [44, 37, 42]). Recently, cloud storage service has become a big industry, and it is estimated that by 2022, the size of the cloud storage market will grow from $23.48 billion in 2016 to $88.91 billion [1]. Also, with the rapid prevalence of Internet-of-Things (IoT) Devices, growth on the diversity of computation-intensive service (*e.g.,* serverless computing and machine learning services), and the increasing need for enterprise mobility, the popularity trend of cloud storage is still thriving.

In spite of its numerous advantages, security, reliability and privacy of cloud storage has always been a severe issue [16, 43, 17]. For the cloud service providers, their storage data centers are normally built as distributed systems with commodity hardware. Accordingly, they are susceptible to both independent and correlated failures [20, 22]. Although many hardware (*e.g.,* RAID facilities and ECC Memory) and software (*e.g.,* replication and erasure codes) techniques have been used to prevent data corruption and ensure security and reliability, data corruption accidents still happen occasionally (*e.g.,* [5, 8, 6, 7]). In this case, the clients hope to be ensured that, their data is safe, reliable and unmodified stored on the cloud.

However, traditional integrate insurance methods, such as hash function and signature, cannot be applied to or is not effective in cloud storage scenarios, as these tools require the client to have full copies of data. Accordingly, it is required to construct specific methods to check the data reliability and integrity in cloud storage cases [10, 23, 13]. Cloud audit protocols, especially the Provable Data Possession (PDP) schemes, have been proposed to fulfill this requirement.

In a typical cloud audit scheme, there exists an auditor (normally referred to as the *third-party auditor, TPA*), which uses a spot-checking technique (instead of accessing the whole data stored on the cloud) to check the data integrity. This is also known as *public auditing.* Recently, public auditing systems have been widely studied [25, 35, 10, 31, 33, 34]. A secure public-auditing system should be able to resist a forge attack, replacing attack, and replay attack from CSP. Moreover, an ideal public-auditing scheme should

also take some desirable proprieties, such as privacy preservation, auditing of dynamic data, batch auditing, auditing of multiple replicas, auditing of shared data and lightweight overheads [32].

However, a suitable TPA may not always exist. In addition, in the use of a cloud storage system, data owners must pay for the cloud storage service in advance. Accordingly, once the data is lost or damaged by CSP, it is hard for data owners to protect their rights. Although the TPA can provide relevant evidence, data owners have to use legal means to defend their rights, which actually requires high additional cost. What is more worrying is that the current law system involving cloud computing is not yet sound, and some related legal issues are difficult to define. Especially, for a cloud storage provider, his storage servers are deployed and distributed around the world. This raises a lot of questions of legal governance over the data. In case a conflict arises between the cloud vendor and the customer, which country's legal system will settle the dispute is still an unresolved question. To resolve the aforementioned problem, we replace the traditional TPA with a smart contract, which is an executable code that runs on the blockchain. Using this technology, we design a fair payment smart contract for the cloud storage system. When data is lost or damaged, the user will no longer have to pay the rent of the cloud service, and will be compensated automatically.

1.1. Our Contribution

Taking advantage of decentralization and automatical triggering of the blockchain, we design a blockchain-based fair payment smart contract for public cloud storage auditing. In our system, data owners and CSP will run a smart contract based on blockchain. The contract ensures that the CSP is required to submit data possession proof regularly. Only if the verification is passed, the CSP will be paid, otherwise the CSP will not only receive no remuneration, but will pay the penalties.

When using the traditional public auditing protocol, the verifier needs to interact with the CSP. In this process, a verifier usually generates a random challenge and CSP returns a data possession proof based on this challenge. This kind of interactive proof is not suitable for executing on a smart contract platform, because each consensus node (as a verifier) has to interact with the CSP, the complexity of system communications and the computing cost of CSP will be unacceptable. In order to avoid the interaction between smart contract platforms and CSP in the execution of a contract, we present the

notion of non-interactive public provable data possession (NI-PPDP) and design a blockchain-based fair payment smart contract for cloud storage based on this primitive. Concretely, we construct an efficient NI-PPDP scheme by non-trivially extending the Wang et al.'s interactive public auditing scheme [33]. Specifically, the contributions of this paper mainly includes the following three aspects:

- Present the notion of non-interactive public provable data possession (NI-PPDP);
- Construct an efficient NI-PPDP scheme, and give formal proof in the random oracle model;
- Design a blockchain-based fair payment smart contract for cloud storage based on NI-PPDP.

1.2. Organization

We introduce the background and preliminaries in Section 2, and give the formal definition of non-interactive public provable data possession (NI-PPDP) in Section 3. Then, we describe the designs of blockchain-based fair payment contract for cloud storage in Section 4. In Section 5, we present a specific NI-PPDP scheme, and analyze that our NI-PPDP scheme meets all the design goals in Section 6. In Section 7, we analyse the performance of our NI-PPDP scheme. Finally, we give a conclusion in Section 8.

1.3. Related Work

1.3.1. Provable Data Possession

Data owners will lose the physical control of their data, when they use the cloud storage service. How to ensure the integrity of remote data is the most important security issues. In 2007, Ateniese et al. [10] introduced a notion of provable data possession (PDP) that allows users to check the remote data without downloading it. Then, Erway et al. [19] proposed the concept of dynamic PDP, which supports data updating. Almost in the same period, Sebé et al. [30] did similar work. In 2008, Shacham and Waters [31] introduced the first fully secure proof-of-retrievability scheme in the Juels-Kaliski model. Since then, a lot of research has been done in this area [39, 14, 36, 26]. However, most of schemes of that period suffer from efficiency problems. In 2015, Liu et al. [27] gave an efficient public auditing scheme based on the Merkle hash tree. This scheme greatly reduces the communication overhead

and improves the verification efficiency. Furthermore, considering the case of identity-based cryptosystem and certificateless cryptosystem, Wang [34] proposed an identity-based PDP scheme, and He et al. [24] proposed a certificateless PDP scheme.

1.4. Blockchain and Smart Contract

In 2008, Satoshi Nakamoto [29] proposed the concept of bitcoin. By combining decentralized consensus technique and append-only data structure, a type of cryptocurrency without the existence of a trusted party is designed. The framework used by bitcoin is generally called *blockchain*. Inspired by bitcoin, some cryptocurrencies and intelligent applications based on blockchain have been proposed one after another, such as Litecoin[3], Zcash[11], Monero[4], Ethereum[2]. These systems have common features: (1) using a consensus protocol to achieve data consensus; (2) using hash chain structure to store data. In general, blockchains are considered to be a linked list of data blocks, where each block is linked by a hash pointer. As shown in Figure 1, the hash value of the previous block is recorded on the head of the next block. Each block includes a collection of data. Once a block is appended to the blockchain, any change to that block will cause a series of changes in the subsequent blocks. In a blockchain system, the distributed nodes update the hash chain synchronously by running a consensus protocol.

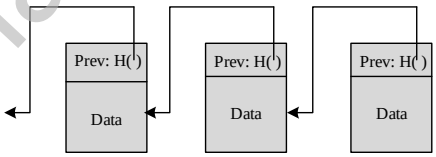


Figure 1: Blockchain

A smart contract is an executable code that runs automatically on the blockchain by consensus nodes without any trusted third party. A smart contract can perform specified operations once pre-defined rules have been met [12]. For instance, using a smart contract, Bob cloud receive x currency units from Alice, if he sends correct calculation results to Alice.

In the smart contract system (Figure 2), each contract has a unique address and cannot be changed after being deployed into the blockchain. When

users execute a contract, they only need to send the transaction to the address stated on that contract. Then, every active consensus node will execute this transaction in the smart contract system to get a consensual result. At present, researchers are trying to use smart contracts to solve various problems in a variety of areas, such as Insurance [21], Medical Care [38], e-Voting [28], Cloud Computing [18] and IoT [15].

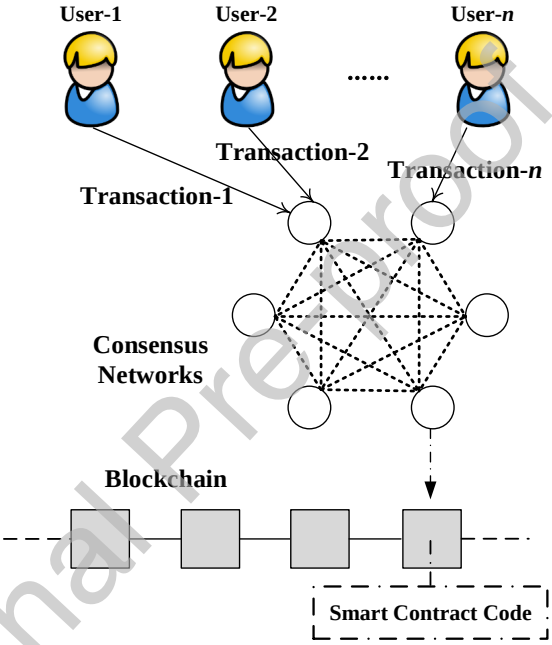


Figure 2: Smart Contract

Recently, Zhang et al. [40, 41] studied the fair payment issues for cloud storage based on blockchain. In their works, PDP is still implemented in a traditional challenge-response way between users and servers. They used bitcoin-based timed commitment technology [9] to achieve fair payment. Our work uses a different approach. After deploying smart contracts, there is no need for challenge-response interaction among users, CSP and any smart contract platform in our system. This will facilitate the implementation of smart contracts by consensus nodes in the public blockchain.

2. Preliminaries

2.1. Bilinear Pairings

Let $\mathbb{G}$ and $\mathbb{G}_T$ be multiplicative cyclic groups with prime order p, and $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ be a function from $\mathbb{G} \times \mathbb{G}$ to $\mathbb{G}_T$. $\mathbb{G}$ and $\mathbb{G}_T$ are called bilinear groups, if

- (Bilinear) $\forall g_1, g_2 \in \mathbb{G}, x, y \in Z_p$, $e({g_1}^x, {g_2}^y) = e(g_1, g_2)^{xy}$.
- (Non-degenerate) $\exists h_1, h_2 \in \mathbb{G}_1$, $e(h_1, h_2)$ is a generator of $\mathbb{G}_T$.

Furthermore, all group operations and function e should be computable.

2.2. Computational Diffie-Hellman (CDH) Assumption

Definition 1. *Suppose $\mathbb{G}$ is a q-order cyclic group, g is a random generator of $\mathbb{G}$. For $\forall x, y \in \{0, \ldots, q-1\}$, given (g, g^x, g^y), it is computationally intractable to compute g^{xy}.*

2.3. Review Wang et al's Interactive Public Auditing Scheme [33]

We only review the basic construction of Wang et al's scheme. For a description of the system model and security model, please refer to [33]. In their scheme, there are two phases, (1) Setup and (2) Audit:

- Setup Phase:

  KeyGen$(1^\lambda) \to pk, sk$: Let g be a generator of bilinear group $\mathbb{G}$. The data owner chooses two hash functions, $H(\cdot)$: $\{0,1\}^* \to \mathbb{G}$, $h(\cdot)$: $\mathbb{G}_T \to \mathbb{Z}_p$, and generates public/private key pairs (spk, ssk) for a digital signature algorithm. Then, it chooses $x \leftarrow \mathbb{Z}_p$, $u \leftarrow \mathbb{G}$ randomly, and calculates $v \leftarrow g^x$. The secret key is $sk = (ssk, x)$ and the public key is pk=(spk, g, u, v, $e(u, v)$, $H(\cdot)$, $h(\cdot)$).

  TagGen$(F, pk, sk) \to \Phi$: We suppose that the data file can be expressed as $F = \{m_i\}_{1 \le i \le n}$, where $m_i \in \mathbb{Z}_p$. The data owner computes authenticators $\sigma_i \leftarrow (H(W_i) \cdot u^{m_i})^x \in \mathbb{G}$ for $i \in [1, n]$, where $W_i = name||i$, and $name \leftarrow \mathbb{Z}_p$ is chosen by the data owner randomly as the identifier of file F. Let $\Psi = \{\sigma_i\}_{1 \le i \le n}$.

  To ensure the correctness of the file identifier $name$, it runs a signing algorithm Sig on the $name$ under ssk, and sets $t = name \,||\, Sig_{ssk}(name)$ as the file identifier for F. The data owner then uploads data file F and corresponding data tags $\Phi = (\Psi, t)$ to CSP.

- Audit Phase:

  The verifier first retrieves the file identifier t, and verifies the signature $Sig_{ssk}(name)$ via spk, and aborts by outputing $\perp$ if verification fails. Otherwise, the verifier recovers the $name$. Then, the verifier picks a random c-element subset $I = \{s_1, ..., s_c\}$ of set $[1, n]$. For each index $i \in I$, the verifier chooses $v_j \leftarrow Z_p^*$ randomly. The verifier sends $chal = \{(i, v_i)\}_{i \in I}$ to the server.

  ProofGen$(pk, \Phi, F, chal) \rightarrow \Sigma$: It computes

$$\sigma = \prod_{j \in I} \sigma_j^{v_j},$$

  and

$$\mu' = \sum_{j \in I} v_j \cdot m_j.$$

  The CSP chooses $s \leftarrow \mathbb{Z}_p$ randomly, and calculates $T = e(u, v)^s \in \mathbb{G}_T$. Then, it computes: $\mu = s + \gamma\mu' \ mod \ p$, where $\gamma = h(T) \in Z_p$. It sends $\Sigma = \{\mu, \sigma, T\}$ as the data possession proof to the verifier.

  Verify(pk, Σ): The verifier computes $\gamma = h(T)$ and outputs 1 or 0, according to the correctness of equation below

$$T \cdot e(\sigma^\gamma, g) \stackrel{?}{=} e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^\gamma \cdot u^\mu, v).$$

3. Non-Interactive Public Provable Data Possession Scheme

3.1. System Model

In a non-interactive public provable data possession (NI-PPDP) scheme, there are three types of entities, that is, data owner, CSP and verifier (as shown in Figure 3). Different from the traditional interactive public provable data possession scheme, CSP and the verifier do not need to interact during auditing. An NI-PPDP scheme consists of 4 algorithms, which are divided into two phases:

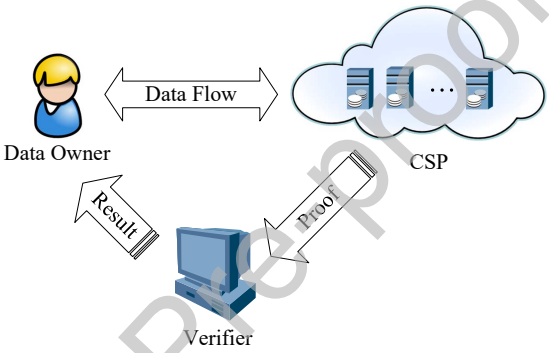


Figure 3: System Model

- Setup Phase: In this phase, data owners generate the data tags Φ corresponding to their data file F and store F along with Φ on the cloud storage service. They will run the key generation algorithm and tag generation algorithm as follows.

  $\mathbf{KeyGen}(1^{\lambda}) \rightarrow pk, sk$: The key generation algorithm is run by data owner. It takes security parameter λ as input, and outputs public key pk, secret key sk.

  $\mathbf{TagGen}(F, pk, sk) \rightarrow \Phi$: The tag generation algorithm is run by the data owner. It takes data file F, public key pk, secret key sk as input, and outputs the corresponding data tags Φ. The data owner then uploads F and Φ to the CSP.

- Audit Phase: In this phase, CSP will prove that it stores complete data. It gives proof based on data tags Φ, data file F and current state τ,

by running a proof generation algorithm. Everybody could verify the proof publicly, by running a verifying algorithm. Usually, the verifier is a third-party auditor (TPA), who has a higher computing capability than the data owner, and can check data integrity for data users.

ProofGen$(pk, \Phi, F, \tau) \rightarrow \Sigma$: The proof generation algorithm is run by CSP. It takes public key pk, data tags Φ, data file F and current state τ as input, and outputs a proof Σ. We suppose the current state τ is some time-varying public information, which cannot be controlled by CSP.

Verify$(pk, \Sigma) \rightarrow 0, 1$: This is a publicly verifiable algorithm, that can be executed by anyone in this system. It takes public key pk, proof Σ as input and outputs 1 or 0 based on the correctness of Σ.

3.2. Threat Model

- We assume that CSP has no incentives to reveal its hosted data to external parties and also has no incentives to drop its hosted data. However, due to some uncontrollable factors, such as, software bugs, hardware failures, bugs in the network path, economically motivated hackers, malicious or accidental management errors, the integrity of users' data might be destroyed. Moreover, for its own benefits, CSP might even decide to hide this data corruption incident to data owners.
- The verifier can verify the integrity of data for data owners according to the proof provided by CSP. However, it may harm the data owners if the verifier could learn related information of the outsourced data from the proof.
- We assume that CSP will not collude with any verifier.

3.3. Design Goals

The NI-PPDP scheme should achieve:

- Correctness: For all keypairs $(pk, sk) \leftarrow$**KeyGen**(1^λ), for all data files F, and for all states τ, the verification algorithm always outputs

$$1 \leftarrow \textbf{Verify}(pk, \textbf{ProofGen}(\textbf{TagGen}(F, sk), F, \tau)).$$

- Soundness(Data integrity): to ensure that the verification can be passed only if the integrity of data is achieved.

- Privacy preserving: to ensure that auditing process does not disclose any information data.
- Non-interactive: to ensure that the CSP and verifier do not need to interact during auditing.
- Public auditability: to ensure that anyone can verify the integrity of the remote data only depending on the data possession proof given by the cloud storage provider and the public key of data owners.

3.4. Formal Security Definition

Among the above design goals, data integrity and privacy preserving are the key security features of NI-PPDP. Therefore, we give the formal definition as follows.

3.4.1. Data Integrity (Soundness)

We use the following game between adversary $\mathcal{A}$ and challenger $\mathcal{C}$ to define the soundness of data integrity:

1. $\mathcal{C}$ calls key generation algorithm $\mathbf{KeyGen}(1^{\lambda})$ to generate keypair (pk, sk), and gives pk to $\mathcal{A}$.
2. $\mathcal{A}$ can interact with $\mathcal{C}$ repeatedly and make queries for some file F. Then, $\mathcal{C}$ returns $\Phi \leftarrow \mathbf{TagGen}(F, pk, sk)$ to $\mathcal{A}$.
3. Finally, $\mathcal{A}$ outputs Σ for some data file F and data tag Φ on state τ.

Define the advantage of $\mathcal{A}$ is $Adv_{\mathcal{A}} = Pr[\mathbf{Verify}(pk, \Sigma) = 1]$. We say the adversary wins the above game, if $Adv_{\mathcal{A}}$ is non-negligible.

Definition 2. *A non-interactive public provable data possession scheme is sound if exists an efficient extraction algorithm* ***Extr*** *such that, for every adversary $\mathcal{A}$, who outputs Σ for some data file F and data tag Φ on state τ and wins above game, the extraction algorithm recovers file F from Φ and Σ, i.e., $\boldsymbol{Extr}(pk, \Phi, \Sigma) = F$.*

3.4.2. Privacy Preserving

We use the following game between adversary $\mathcal{A}$ and challenger $\mathcal{C}$ to define privacy preserving:

1. $\mathcal{C}$ calls key generation algorithm $\mathbf{KeyGen}(1^{\lambda})$ to generate keypair (pk, sk), and gives pk to $\mathcal{A}$.
2. A can interact with $\mathcal{C}$ repeatedly and make queries for some file F. Then, $\mathcal{C}$ returns $\Phi \leftarrow \mathbf{TagGen}(F, pk, sk)$ to $\mathcal{A}$.
3. At some point, $\mathcal{A}$ submits two files F_0^* and F_1^* to $\mathcal{C}$. Then, C selects $b \in \{0, 1\}$ randomly, and returns $\mathbf{ProofGen}(\mathbf{TagGen}(F_b^*, sk), F_b^*, \tau)$ to A.
4. Finally, $\mathcal{A}$ outputs a guess bit b'.

Defining the advantage of $\mathcal{A}$ as $Adv_{\mathcal{A}} = Pr[b' = b] - 1/2$. We say the adversary wins the above game, if $Adv_{\mathcal{A}}$ is non-negligible.

Definition 3. *A non-interactive public provable data possession scheme is privacy preserved if for any probability polynomial time adversary $\mathcal{A}$, advantage of $\mathcal{A}$ in above game is negligible.*

4. Blockchain-Based Fair Payment Smart Contract for Cloud Storage

In a traditional cloud storage system, data owners must pay the rent before using cloud storage. Once the data is lost or damaged by cloud storage service provider, it is hard for data owners to restore economic losses. To solve this problem, we introduce a novel cloud storage payment model, in which the data owners will pay the fees according to the service quality after enjoying the service. In order to protect the rights of both the data owners and cloud storage service provider, we use a blockchain-based smart contract platform and non-interactive public provable data possession scheme in this system.

As shown in Figure 4, the data owner runs key generation algorithms and tag generation algorithms of the NI-PPDP scheme, uploads file F and the data tags Φ to the CSP. At the same time, it submits the contract T_0 (Figure 5) to the smart contract platform. T_0 includes file name, file size, file hash,

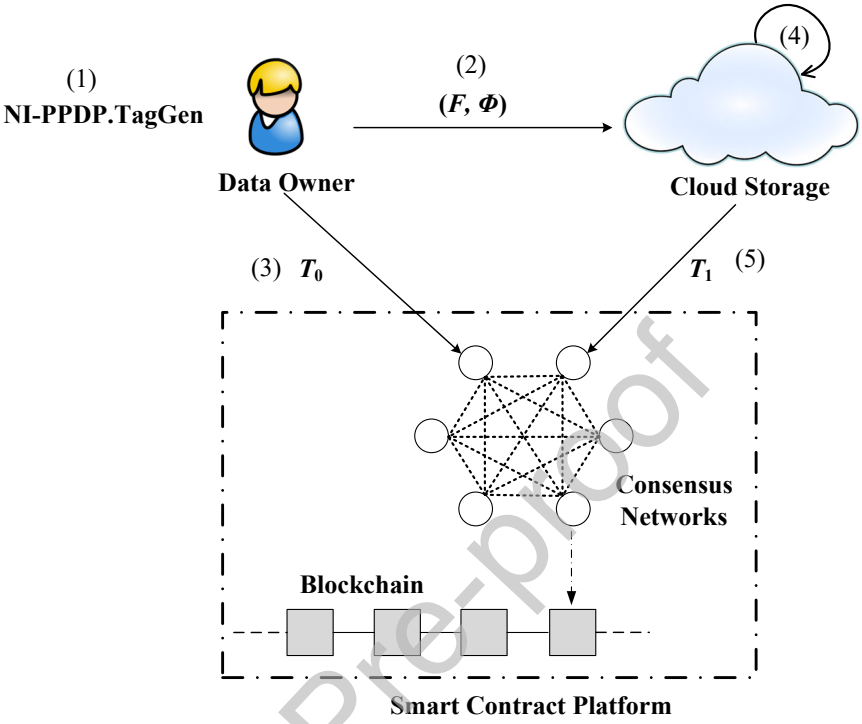


Figure 4: Data Storage Process

upload time, storage period, service charges, data owner account (cryptocurrency), cloud storage account (cryptocurrency), data owner's public key, data owner's signature and code of smart contract. This contract ensures that if cloud storage service provider could submit correct data possession proof (using NI-PPDP scheme) on time, the data owner will pay the service fees on time.

After receiving the file F and data tags Φ, the cloud storage provider will check the integrity of data, and the authentication of source. If all verification checks are passed, cloud storage server will submit the contract T_1 (Figure 6) to the smart contract platform. T_1 confirms that file F has been received by CSP, and ensures that if the data possession proof is not passed, CPS will pay the compensation to the data owner. That is T_1 has two functions: (1) to confirm the receipt of F, (2) to make a promise of compensation. Note that compensation is not necessary, but compensation reflects the reputation of the cloud storage provider.

Contract T_0
File Name: FN File Size: FS File Hash: FH Upload time: UT Storage Period: CP Service Charges: SC Data owner Account: DOA Cloud Storage Account: CSA Data owner's public key: pk_D Data owner's signature: sig_D

Contract Content:
promise
{
if (NI-PPDP.Verify(pk_D, Σ)==1)
pay SC from DOA to CSA;
}

Figure 5: Contract T_0

The specific workflow can be described as:

1. Data owner employs the NI-PPDP scheme, and runs its **TagGen** algorithms on file F to obtain the corresponding data tags Φ.
2. Date owner uploads the file F and the data tags Φ to CSP.
3. Date owner submits the contract T_0 (Figure 5) to the smart contract platform.
4. CSP checks the integrity of data and the authentication of source.
5. CSP submits the contract T_1 (Figure 6) to the smart contract platform.

As shown in Figure 7, in order to get service charges, the cloud server periodically submits a contract T_2 (Figure 8), which contains a non-interactive data possession proof Σ. The consensus network will verify this data possession proof in the T_2 and activate T_0 or T_1 based on the validation result. If validation is successful, the T_0 will be activated and the data owner pays the service fees to the cloud server, else the T_1 will be activated, and the CPS will pay compensation to the data owner.

Contract T_1
File Name: FN File Size: FS File Hash: FH Receiving Time: RT Storage Period: CP Penalty: Pen Data owner Account: DOA Cloud Storage Account: CSA CSP's public key: pk_C CSP's signature: sig_C

Contract Content:
confirm F; /*T_0 takes effect*/
promise
{
if (NI-PPDP.Verify(pk_D, Σ)==0)
pay Pen from CSA to DOA;
}

Figure 6: Contract T_1

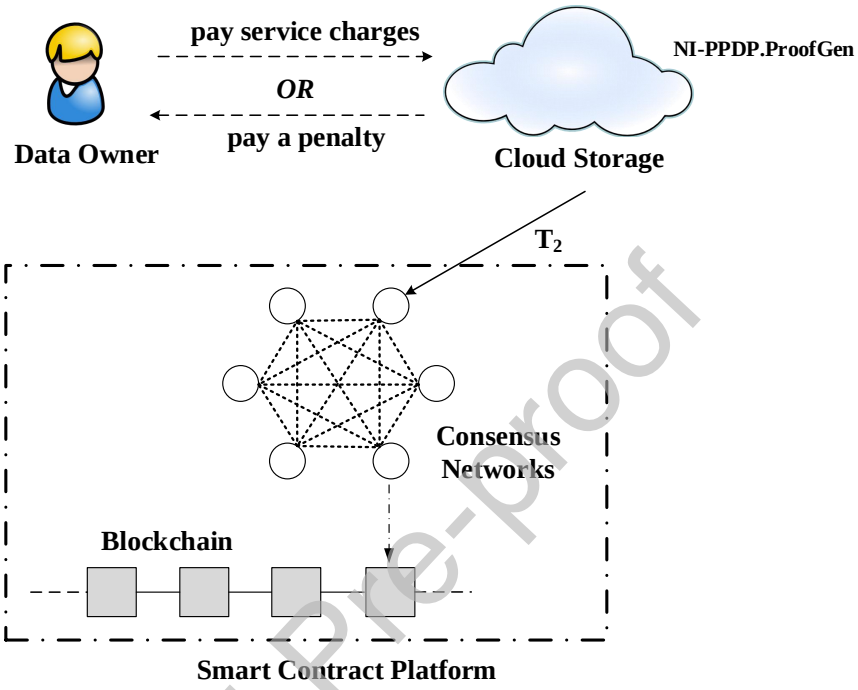


Figure 7: Data Validation Process

Contract T_2
Payment contract: T_0 Compensation contract: T_1 Data possession proof: Σ
Contract Content: activate T_0 or T_1 depended on the validation result of Σ;

Figure 8: Contract T_2

5. A Specific NI-PPDP Scheme

In the following, we present a specific construction of non-interactive public provable data possession scheme. Our construction is achieved by extending interactive public auditing schemes introduced by Shacham and Waters [31] and Wang et al. [33].

5.1. Our Construction

Our NI-PPDP scheme is constructed, based on Wang et al.'s public auditing scheme [33]. The main difference is that there is no interaction between the verifier and CSP in the Audit Phase. In order to simulate the challenge process, we use the pseudorandom function on the input of current state. There are also two phases in our scheme.

- Setup Phase:

  KeyGen$(1^\lambda) \rightarrow pk, sk$: Let $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$, select g as a generator of group $\mathbb{G}$, select two cryptographic hash functions $H(\cdot) : \{0,1\}^* \rightarrow \mathbb{G}$, $h(\cdot) : \mathbb{G}_T \rightarrow \mathbb{Z}_p$, and select pseudorandom function $\mathcal{F}(\cdot)$: $\{0,1\}^* \rightarrow [1,n]$, which maps arbitrary values uniformly to an integer range $[1,n]$.

  The data owner generates public/private key pairs (spk, ssk) for a digital signature algorithm. Then, it chooses $x \leftarrow \mathbb{Z}_p$, $u \leftarrow \mathbb{G}$ randomly, and calculates $v \leftarrow g^x$. The secret key is $sk = (ssk, x)$ and the public key is pk=(spk, g, u, v, e, $H(\cdot)$, $h(\cdot)$, $\mathcal{F}(\cdot)$).

  TagGen$(F, pk, sk) \rightarrow \Phi$: Suppose that the data file can be expressed as $F = \{m_i\}_{1 \leq i \leq n}$, where $m_i \in \mathbb{Z}_p$. The data owner computes authenticators $\sigma_i \leftarrow (H(W_i) \cdot u^{m_i})^x$ for $i \in [1,n]$, where $W_i = name||i$, and $name \leftarrow \mathbb{Z}_p$ is chosen by the data owner randomly as the identifier of file F. Let $\Psi = \{\sigma_i\}_{1 \leq i \leq n}$.

  To ensure the correctness of the file identifier $name$, it runs signature algorithm Sig on $name$ under ssk, and sets $t = name \,||\, Sig_{ssk}(name)$ as the file identifier for F. The data owner then uploads the data file F and corresponding data tags $\Phi = (\Psi, t)$ to CSP.

- Audit Phase:

  In this phase, the verifier does not need to choose the challenge set. The CSP uses the current state as the input of pseudorandom function $\mathcal{F}(\cdot)$, to simulate the challenge process.

ProofGen$(pk, \Phi, F, \tau) \rightarrow \Sigma$: In input, τ represents the current public status information, which contains current time and some other public information and cannot be controlled by CSP. We assume that τ will change in each running of **ProofGen** algorithm. In our blockchain-based fair payment model, τ should also include the header information of the current block of blockchain, which can not be controlled by CSP.

First of all, CSP choose an appropriate number $c < n$. For $i \in [1, c]$, CSP computes

$$s_i \leftarrow \mathcal{F}(\tau||i).$$

$I = \{s_1, s_2, ..., s_c\}$ is a c-element multiset, $\forall s_i \in [1, n]$. Note, the multiset I is allowed to contain repeated elements.

For $j \in I$, the CSP computes

$$v_j \leftarrow h(\tau||j).$$

Then, it computes

$$\sigma = \prod_{j \in I} \sigma_j^{v_j},$$

and

$$\mu' = \sum_{j \in I} v_j \cdot m_j.$$

The CSP chooses $s \leftarrow \mathbb{Z}_p$ randomly, and calculates $T = e(u, v)^s \in \mathbb{G}_T$. Then, it computes: $\mu = s + \gamma\mu' \ mod \ p$, where $\gamma = h(T) \in Z_p$. It sends $\Sigma = \{\mu, \sigma, T, \tau, c\}$ as the data possession proof to the verifier.

Verify(pk, Σ): The verifier runs the verification algorithm of $Ver(spk, id, Sig_{ssk}(name))$ to verify integrity of id and verifies the authenticity of state information τ via blockchain. It aborts, if any verification fails. Otherwise, the verifier recovers the $name$. Then, the verifier computes $I = \{\mathcal{F}(\tau||1), \mathcal{F}(\tau||2), ..., \mathcal{F}(\tau||c)\}$, $\{v_j = h(\tau||j)\}_{j \in I}$, $\{h(N_j)\}_{j \in I}$, $\gamma = h(T)$ and then outputs 1 or 0, according to the correctness of equation below

$$T \cdot e(\sigma^\gamma, g) \stackrel{?}{=} e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^\gamma \cdot u^\mu, v).$$

5.2. Correctness

$$
\begin{aligned}
T \cdot e(\sigma^{\gamma}, g) =& e(u, v)^{s} \cdot e((\prod_{i=s_1}^{s_c} (H(W_i) \cdot u^{m_i})^{x \cdot v_i})^{\gamma}, g) \\
=& e(u^{s}, v) \cdot e((\prod_{i=s_1}^{s_c} (H(W_i)^{v_i} \cdot u^{m_i v_i}))^{\gamma}, g)^{x} \\
=& e(u^{s}, v) \cdot e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma} \cdot u^{\mu' \gamma}, v) \\
=& e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma} \cdot u^{\mu' \gamma + s}, v) \\
=& e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma} \cdot u^{\mu}, v)
\end{aligned}
$$

6. Design Goals Analysis

6.1. Data Integrity (Soundness)

We use the hybrid argument technique to prove soundness as in [31]. First of all, we define the following games:

Game-0. Game-0 is the original game defined in Section 3.4.1.

Game-1. Game-1 is the same as Game-0, except that the challenger $\mathcal{C}$ records all the tags it signed in a local list. If adversary $\mathcal{A}$ ever submits a tag Φ, that (1) has a valid signature under ssk but (2) is not signed by $\mathcal{C}$, then $\mathcal{C}$ announces failure and aborts.

Game-2. Game-2 is the same as Game-1, except that $\mathcal{C}$ records all the responses to **TagGen** queries from $\mathcal{A}$. If $\mathcal{A}$ is successful (i.e., **Verify** output 1) but $\mathcal{A}$'s aggregate signature σ is not equal to $\prod_{j \in I} \sigma_j^{v_j}$, then the challenger $\mathcal{C}$ announces failure and aborts.

Game-3. Game-3 is the same as Game-2, except that challenger $\mathcal{C}$ announces failure and aborts, if at least one of the aggregate messages μ' is not equal to $= \sum_{j \in I} v_j \cdot m_j$.

Lemma 1. *If there is an algorithm $\mathcal{A}$ can distinguish between **Game-0** and **Game-1** with the non-negligible probability, then we can construct an algorithm $\mathcal{B}$ that has a non-negligible advantage to break the existentially unforgeability.*

Analysis. If $\mathcal{A}$ causes $\mathcal{C}$ to abort in Game-1, then we can use $\mathcal{A}$ to construct an algorithm $\mathcal{B}$ against the existentially unforgeability of the signature scheme.

Lemma 2. *If there is an algorithm $\mathcal{A}$ can distinguish between **Game-1** and **Game-2** with the non-negligible probability, then we can construct an algorithm $\mathcal{B}$ that has non-negligible advantage to break the computation Diffie-Hellman assumption.*

Analysis. Suppose g^x and g^y are the elements of CDH problem, we set $v = g^x$, $u = g^y$. Suppose $\mathcal{A}$ can respond a signature σ', which is different from the expected signature σ. We can calculate

$$e(\sigma'/\sigma, g) = e(\prod_{j\in I} u^{\Delta\mu_j}, v) = e(g^{\sum_{j\in I}\Delta\mu_j \cdot x\cdot y}, g)$$

Therefore, we can calculate $g^{x\cdot y} = (\sigma'/\sigma)^{\frac{1}{\sum_{j\in I}\Delta\mu_j}}$

Lemma 3. *If there is an algorithm $\mathcal{A}$ that can distinguish between **Game-2** and **Game-3** with the non-negligible probability, then we can construct an algorithm $\mathcal{B}$ that has a non-negligible advantage to break the computation Diffie-Hellman assumption.*

Analysis. We only introduce the main ideas. We suppose that $h(\cdot)$ is a random oracle controlled by an extractor, who answers the hash query asked by the adversary (CSP). For $\gamma = h(T)$ from extractor, the adversary outputs $\{\mu, \sigma, T, t, \tau, c\}$ makes:

$$T \cdot e(\sigma^\gamma, g) = e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^\gamma \cdot u^\mu, v).$$

Then, the extractor rewinds $h(T)$ to be $\gamma^* \neq \gamma$. The adversary outputs $\{\mu^*, \sigma, T, t, \tau, c\}$ makes:

$$T \cdot e(\sigma^{\gamma^*}, g) = e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma^*} \cdot u^{\mu^*}, v).$$

Divide above two equations, we have

$$\begin{aligned}
e(\sigma^{\gamma-\gamma^*}, g) &= e(u^{\sum_{j\in I}(h(N_j)v_j)(\gamma-\gamma^*)} \cdot u^{\mu-\mu^*}, v) \\
e(\sigma^{\gamma-\gamma^*}, g) &= e(u^{\sum_{j\in I}(h(N_j)v_j)(\gamma-\gamma^*)} \cdot u^{\mu-\mu^*}, g^x) \\
\sigma^{\gamma-\gamma^*} &= u^{\sum_{j\in I}(h(N_j)v_j)x(\gamma-\gamma^*)} \cdot u^{x(\mu-\mu^*)} \\
(\Pi_{j\in I}\sigma_j^{v_j})^{\gamma-\gamma^*} &= (\Pi_{j\in I} u^{h(N_j)v_j x(\gamma-\gamma^*)} \cdot u^{x(\mu-\mu^*)} \\
u^{x(\mu-\mu^*)} &= (\Pi_{j\in I}(\sigma_j / u^{h(N_j)x})^{v_j})^{\gamma-\gamma^*} \\
u^{x(\mu-\mu^*)} &= (\Pi_{j\in I}(u^{xm_j})^{v_j})^{\gamma-\gamma^*} \\
\mu - \mu^* &= (\sum_{j\in I} m_j v_j) \cdot (\gamma - \gamma^*) \\
\sum_{j\in I} m_j v_j &= (\gamma - \gamma^*)/(\mu - \mu^*)
\end{aligned}$$

Finally, $\{\sigma, \mu' = (\mu - \mu^*)/(\gamma - \gamma^*)\}$ can be treated as a response for the extractor.

Theorem 1. *If the signature scheme is existentially unforgeable and computational Diffie-Hellman assumption holds in bilinear groups, then no probabilistic polynomial time adversary can break the soundness of our NI-PPDP scheme with non-negligible probability.*

Proof. Any adversary's advantage in **Game 3** must be 0 since the challenger always announces failure and aborts if there is no integral file F, i.e. at least one of the aggregate messages μ' is not equal to $= \sum_{j\in I} v_j \cdot m_j$. By the games sequence and Lemmas 1-3, an adversary's advantage in the original game **Game 0** must be negligibly close to 0.

6.2. Privacy Preserving

This theorem shows that the verification process do not reveal any information of the users' data.

Theorem 2. *Our NI-PPDP scheme is privacy preserved.*

Proof. This proof follows from [31] and [33]. We only introduce the main ideas, i.e. the data possession proof $\Sigma = \{\mu, \sigma, T, t, \tau, c\}$ can not reveal any information about μ'. In the random oracle model, the simulator can construct the response without knowing μ'. It randomly chooses γ, μ from Z_p, and sets

$$T \leftarrow e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma} \cdot u^{\mu}, v)/e(\sigma^{\gamma}, g).$$

Then, the simulator sets random oracle $h(\cdot)$, makes $\gamma = h(T)$.

6.3. Non-Interactive

Compared with the traditional interactive public provable data possession scheme, there is no challenge stage in our scheme. CSP does not have to interact with the verifier when it makes proof. To achieve this goal, we use the pseudorandom function, $\mathcal{F}(\cdot)$, to generate the challenge set. Since the input of the $\mathcal{F}(\cdot)$ contains the current state information, this information is related to the current time and the current block chain state, so it cannot be controlled by the CPS. Due to the pseudo randomness, the challenge set generated in this way is indistinguishable to the random choice of verifier.

6.4. Public Auditability

It is clear that our scheme has achieved public auditability. The validation algorithm does not depend on any secret inputs.

7. Efficiency Analysis

7.1. Theoretical Analysis

In the pairing-based cryptography scheme, the computation overhead mainly comes from pairing, as well as exponentiation and multiplication in the group $\mathbb{G}$.

The specific computational analysis is given in Table 1. In the TagGen phase, the main computation overhead comes from the calculation of $\{\sigma_i \leftarrow (H(W_i) \cdot u^{m_i})^x\}_{1 \le i \le n}$, which contains c multiplications and n exponentiations on the group $\mathbb{G}$. In the ProofGen phase, the main computation overhead comes from the calculation of $\sigma = \prod_{j \in I} \sigma_j^{v_j}$, which contains c exponentiations and c multiplications on group $\mathbb{G}$. In the Verify phase, the main computation overhead comes from the calculation of $T \cdot e(\sigma^{\gamma}, g) \stackrel{?}{=} e((\prod_{i=s_1}^{s_c} H(W_i)^{v_i})^{\gamma} \cdot u^{\mu}, v)$, which contains $c+1$ multiplications on $\mathbb{Z}_p$, $c+3$ exponentiations on $\mathbb{G}$, and 2 pairings.

Table 1: Computational Analysis

	Multiplication	Exponentiation	Pairing
TagGen	n	$n+1$	0
ProofGen	c	c	0
Verify	$c+1$	$c+3$	2

7.2. Experimental Evacuation

In order to get an in-depth evaluation for the performance of our scheme, we conducted an experiment by implementing a prototype of our scheme with the C Programming language. We adopted the GMP[1] and PBC[2] library for big integer and pairing operation, and adopted OpenSSL[3] for basis cryptographic primitives (e.g., pseudorandom function). We choose the Type-A elliptic curve with order of 160-bit. Both of the programs for the Cloud Service Provider and the Client side are compiled with clang of vision 900.0.39.2, and run on MacBook Pro with 2.7 GHz Intel Core i5 CPU and 8 GB 1867 MHz DDR3 memory.

In our experiment, we mainly focus on the computational overhead (i.e., The operation time of TagGen, ProofGen and Verify), whereas do not take the Round-Trip Time (RTT) into consideration. This is mainly because the RTT is heavily dependent on the network condition, instead of the proposed scheme and sometimes the RTT may dwarf the operation times, which will make the evaluation unaccurate.Similarly, we also do not account the I/O latency, because the I/O latency is generally determined by type of external storage facilities (e.g, SSD or HDD), the disk scheduling algorithms used in the operating system, data transmission rate and disk interface type. These factors are independent of the proposed scheme, and these uncertainties mentioned above will involve turbulence in our evaluation. Thus, the time overhead of these operations is measured as the time duration between the data is loaded to the memory until the operation is finished.

In terms of TagGen, the client generates the tags for a file, which has been split into several segment (blocks). Figure 9 and Figure 10 demonstrate the time overhead of TagGen (i.e., the Y-ray) with different number of file

[1]The GNU MP Bignum Library, https://gmplib.org/
[2]PBC Library - Pairing-Based Cryptography, https://crypto.stanford.edu/pbc/
[3]Cryptography and SSL/TLS Toolkit, https://www.openssl.org/

blocks. Specifically, in order to observe the variation details, in Figure 9 we depict the operation time of TagGen as the number of file blocks ranging from 100 to 100 (with the resolution of 100); whereas for the perpose of getting a general overview of the variation tendency, in Figure 10 we test it with file blocks ranging from 1000 to 10000 (with resolution of 1000). It is manifest that the operation time of TagGen grows linearly with the increase of file blocks and the stable linear-incremental tendency not only exists in small-scale file blocks, fine-grained interval settings (i.e., depicted in Figure 9), but also appears as a general tide (i.e., Figure 10). As is shown in these figures, when the file consists 10000 blocks, generating the tags for it costs less than 0.02 second and when the files grow even larger, we can infer that the time overheard for tag generation increase proportionally with block numbers. Thus, if the file size extends to Gigabyte scale (which will result in $\frac{1000\times1024\times1024\times8}{160} = 52,428,800$ blocks), based on the pattern observed above, the Tag Generation can be finished within $52,428,800 \times \frac{0.02}{10000} \approx 2$ minutes. As this procedure is executed by the client in an offline manner and the file size is seldom as large as the gigabit scale, the overhead for TagGen is acceptable.

The ProofGen is executed by cloud service provider after receiving the challenge request (indicating the number of proof should be generated), and it will generate the proofs of the file under direction of the challenge. In Verify, the client checks the validity of all the proofs generated by the cloud service provider. In our experiment, we measure the operation overhead of operations ProofGen and Verify in the scene of proving and verifying a file with 10,000 blocks. As depicted in Figure 11, when the challenge blocks increases for 100 to 1000 (i.e., the X-ray), the operation time of ProofGen and Verify also increase accordingly. Specifically the growth rate of ProofGen becomes sharper as the raise of challenge blocks; whereas the Verify, whose operation overhead is much lower than ProofGen, maintains a rigorous linear growth trend. As shown in Figure 11, the overhead of ProofGen and Verify is approximately 14 ms and 11 ms respectively, when the number challenge blocks reaches 1000. Thus, our scheme is capable for most of the real world settings.

8. Conclusion

In this paper, we design a novel blockchain-based fair payment smart contract for cloud storage. The contract ensures that the CSP is required to

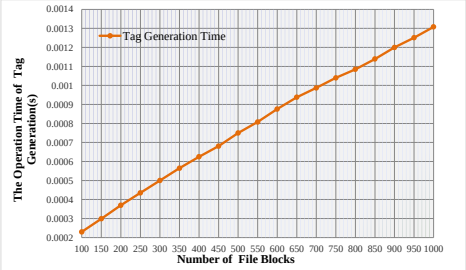


Figure 9: Operation Time for Tag Generation with File Blocsk from 100 to 1000.

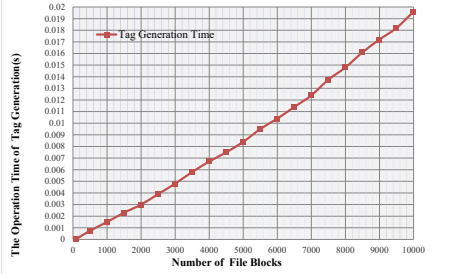


Figure 10: Operation Time for Tag Generation with File Blocks from 1000 to 10000.

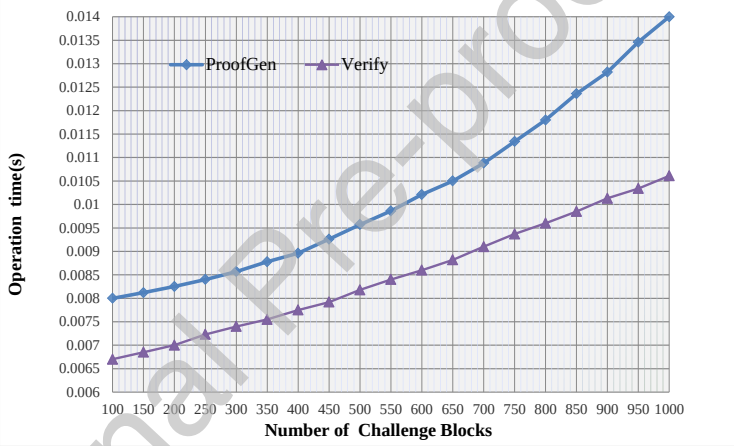


Figure 11: Operation Time for ProofGen and Verify

submit data possession proof regularly. Only if the verification is passed, the CSP will be paid, otherwise the CSP will not only receive no remuneration, but also will be responsible to pay the penalties. In order to avoid the interaction between smart contract platform and CSP in the execution of contract, we introduce the non-interactive public provable data possession scheme and design an efficient construction.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No. 61602287, No. 61802235, No. 61672330, and No. 61702168), the

Primary Research & Development Plan of Shandong Province (No. 2018GGX101037), the Major Scientific and Technological Innovation Project of Shandong Province (No. 2018CXGC0702), and the Development and Construction Funds Project of National Independent Innovation Demonstration Zone in Shandong Peninsula (No. S190101010001).

REFERENCES

References

[1] Cloud storage market report. https://www.marketsandmarkets.com/Market-Reports/cloud-storage-market-902.html. Accessed July 16, 2019.

[2] Ethereum. https://ethereum.org/. Accessed July 16, 2019.

[3] Litecoin. https://litecoin.com. Accessed July 16, 2019.

[4] Monero. https://monero.org/. Accessed July 16, 2019.

[5] Cloud Storage Often Results in Data Loss, 2011. https://www.businessnewsdaily.com/1543-cloud-data-storage-problems.html.

[6] What is the risk of losing data stored in the cloud?, 2018. https://www.quora.com/What-is-the-risk-of-losing-data-stored-in-the-cloud.

[7] OOPS: Google "loses" your cloud data (sky falling; film at 11), August 20, 2015. https://www.computerworld.com/article/2973600/cloud-computing/google-cloud-loses-data-belgium-itbwcw.html.

[8] What happens when data gets lost from the cloud?, January 26, 2015. https://www.cloudcomputing-news.net/news/2015/jan/26/what-happens-when-data-gets-lost-cloud/.

[9] Marcin Andrychowicz, Stefan Dziembowski, Daniel Malinowski, and Lukasz Mazurek. Secure multiparty computations on bitcoin. In *2014 IEEE Symposium on Security and Privacy, SP 2014, Berkeley, CA, USA, May 18-21, 2014*, pages 443–458, 2014.

[10] Giuseppe Ateniese, Randal C. Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary N. J. Peterson, and Dawn Xiaodong Song. Provable data possession at untrusted stores. In *Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007*, pages 598–609, 2007.

[11] Eli Ben-Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from bitcoin. In *2014 IEEE Symposium on Security and Privacy, SP 2014, Berkeley, CA, USA, May 18-21, 2014*, pages 459–474, 2014.

[12] Vitalik Buterin. A next-generation smart contract and decentralized application platform. 2014.

[13] David Cash, Alptekin Küpçü, and Daniel Wichs. Dynamic proofs of retrievability via oblivious ram. *Journal of Cryptology*, 30(1):22–57, 2017.

[14] Ee-Chien Chang and Jia Xu. Remote integrity check with dishonest storage server. In *Computer Security - ESORICS 2008, 13th European Symposium on Research in Computer Security, Málaga, Spain, October 6-8, 2008. Proceedings*, pages 223–237, 2008.

[15] Konstantinos Christidis and Michael Devetsikiotis. Blockchains and smart contracts for the internet of things. *IEEE Access*, 4:2292–2303, 2016.

[16] Cheng-Kang Chu, Wen-Tao Zhu, Jin Han, Joseph K Liu, Jia Xu, and Jianying Zhou. Security concerns in popular cloud storage services. *IEEE Pervasive Computing*, 12(4):50–57, 2013.

[17] Hyunji Chung, Jungheum Park, Sangjin Lee, and Cheulhoon Kang. Digital forensic investigation of cloud storage services. *Digital investigation*, 9(2):81–95, 2012.

[18] Changyu Dong, Yilei Wang, Amjad Aldweesh, Patrick McCorry, and Aad van Moorsel. Betrayal, distrust, and rationality: Smart counter-collusion contracts for verifiable cloud computing. In *Proceedings of*

the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017, pages 211–227, 2017.

[19] C. Christopher Erway, Alptekin Küpçü, Charalampos Papamanthou, and Roberto Tamassia. Dynamic provable data possession. In *Proceedings of the 2009 ACM Conference on Computer and Communications Security, CCS 2009, Chicago, Illinois, USA, November 9-13, 2009*, pages 213–222, 2009.

[20] Daniel Ford, François Labelle, Florentina I Popovici, Murray Stokely, Van-Anh Truong, Luiz Barroso, Carrie Grimes, and Sean Quinlan. Availability in Globally Distributed Storage Systems. In *Proceedings of the 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, pages 61–74. USENIX, 2010.

[21] Valentina Gatteschi, Fabrizio Lamberti, Claudio Demartini, Chiara Pranteda, and Victor Santamaria. Blockchain and smart contracts for insurance: Is the technology mature enough? *Future Internet*, 10(2):20, 2018.

[22] Phillipa Gill, Navendu Jain, and Nachiappan Nagappan. Understanding Network Failures in Data Centers: Measurement, Analysis, and Implications. In *Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)*, pages 350–361. ACM, 2011.

[23] Jing Han, Yanping Li, Jianqing Liu, and Minghao Zhao. An efficient lucas sequence-based batch auditing scheme for the internet of medical things. *IEEE Access*, 7:10077–10092, 2018.

[24] Debiao He, Neeraj Kumar, Huaqun Wang, Lina Wang, and Kim-Kwang Raymond Choo. Privacy-preserving certificateless provable data possession scheme for big data storage on cloud. *Applied Mathematics and Computation*, 314:31–43, 2017.

[25] Ari Juels and Burton S. Kaliski Jr. Pors: proofs of retrievability for large files. In *Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007*, pages 584–597, 2007.

[26] Jingwei Li, Jin Li, Dongqing Xie, and Zhang Cai. Secure auditing and deduplicating data in cloud. *IEEE Trans. Computers*, 65(8):2386–2396, 2016.

[27] Chang Liu, Rajiv Ranjan, Chi Yang, Xuyun Zhang, Lizhe Wang, and Jinjun Chen. Mur-dpa: Top-down levelled multi-replica merkle hash tree based secure public auditing for dynamic big data storage on cloud. *IEEE Trans. Computers*, 64(9):2609–2622, 2015.

[28] Patrick McCorry, Siamak F. Shahandashti, and Feng Hao. A smart contract for boardroom voting with maximum voter privacy. In *Financial Cryptography and Data Security - 21st International Conference, FC 2017, Sliema, Malta, April 3-7, 2017, Revised Selected Papers*, pages 357–375, 2017.

[29] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[30] Francesc Sebé, Josep Domingo-Ferrer, Antoni Martínez-Ballesté, Yves Deswarte, and Jean-Jacques Quisquater. Efficient remote data possession checking in critical information infrastructures. *IEEE Trans. Knowl. Data Eng.*, 20(8):1034–1038, 2008.

[31] Hovav Shacham and Brent Waters. Compact proofs of retrievability. In *Advances in Cryptology - ASIACRYPT 2008, 14th International Conference on the Theory and Application of Cryptology and Information Security, Melbourne, Australia, December 7-11, 2008. Proceedings*, pages 90–107, 2008.

[32] Hui Tian, Yuxiang Chen, Hong Jiang, Yongfeng Huang, Fulin Nan, and Yonghong Chen. Public auditing for trusted cloud storage services. *IEEE Security & Privacy*, 17(1):10–22, 2019.

[33] Cong Wang, Sherman S. M. Chow, Qian Wang, Kui Ren, and Wenjing Lou. Privacy-preserving public auditing for secure cloud storage. *IEEE Trans. Computers*, 62(2):362–375, 2013.

[34] Huaqun Wang. Identity-based distributed provable data possession in multicloud storage. *IEEE Trans. Services Computing*, 8(2):328–340, 2015.

[35] Qian Wang, Cong Wang, Kui Ren, Wenjing Lou, and Jin Li. Enabling public auditability and data dynamics for storage security in cloud computing. *IEEE Trans. Parallel Distrib. Syst.*, 22(5):847–859, 2011.

[36] Lifei Wei, Haojin Zhu, Zhenfu Cao, Weiwei Jia, and Athanasios V. Vasilakos. Seccloud: Bridging secure storage and computation in cloud. In *30th IEEE International Conference on Distributed Computing Systems Workshops (ICDCS 2010 Workshops), 21-25 June 2010, Genova, Italy*, pages 52–61, 2010.

[37] Zhe Wu, Curtis Yu, and Harsha V Madhyastha. Costlo: Cost-effective redundancy for lower latency variance on cloud storage services. In *Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 543–557, 2015.

[38] Xiao Yue, Huiju Wang, Dawei Jin, Mingqiang Li, and Wei Jiang. Healthcare data gateways: Found healthcare intelligence on blockchain with novel privacy risk control. *J. Medical Systems*, 40(10):218:1–218:8, 2016.

[39] Ke Zeng. Publicly verifiable remote data integrity. In *Information and Communications Security, 10th International Conference, ICICS 2008, Birmingham, UK, October 20-22, 2008, Proceedings*, pages 419–434, 2008.

[40] Yinghui Zhang, Robert H. Deng, Ximeng Liu, and Dong Zheng. Blockchain based efficient and robust fair payment for outsourcing services in cloud computing. *Inf. Sci.*, 462:262–277, 2018.

[41] Yinghui Zhang, Robert H. Deng, Ximeng Liu, and Dong Zheng. Outsourcing service fair payment based on blockchain and its applications in cloud computing. *IEEE Transactions on Services Computing*, pages 1–1, 2018.

[42] Yupu Zhang, Chris Dragga, Andrea C Arpaci-Dusseau, and Remzi H Arpaci-Dusseau. Viewbox: Integrating local file systems with cloud storage services. In *Proceedings of the 12th USENIX Conference on File and Storage Technologies (FAST)*, pages 119–132, 2014.

[43] Minghao Zhao, Chengyu Hu, Xiangfu Song, and Chuan Zhao. Towards dependable and trustworthy outsourced computing: A comprehensive

survey and tutorial. *Journal of Network and Computer Applications*, 131:55–65, 2019.

[44] Minghao Zhao, Zhenhua Li, Ennan Zhai, Gareth Tyson, Chen Qian, Zhenyu Li, and Leiyu Zhao. H2cloud: Maintaining the whole filesystem in an object storage cloud. In *Proceedings of the 47th International Conference on Parallel Processing*, page 68. ACM, 2018.

CRediT Author Statement

Hao Wang: Conceptualization, Methodology, Formal analysis, Writing- Reviewing and Editing, Funding acquisition.

Hong Qin: Investigation, Validation, Writing- Original draft preparation.

Minghao Zhao: Software.

Xiaochao Wei: Formal analysis.

Hua Shen: Visualization.

Willy Susilo: Supervision.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: