Information Sciences 258 (2014) 371–386

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Security and privacy for storage and computation in cloud computing

Lifei Wei [a], Haojin Zhu [a], Zhenfu Cao [a,*], Xiaolei Dong [a], Weiwei Jia [a], Yunlu Chen [a], Athanasios V. Vasilakos [b]

[a] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Department of Computer and Telecommunications Engineering, University of Western Macedonia, Kozani, Greece

ARTICLE INFO

Article history:
Available online 27 April 2013

Keywords:
Secure computation auditing
Secure storage
Privacy-cheating discouragement
Designated verifier signature
Batch verification
Cloud computing

ABSTRACT

Cloud computing emerges as a new computing paradigm that aims to provide reliable, customized and quality of service guaranteed computation environments for cloud users. Applications and databases are moved to the large centralized data centers, called *cloud*. Due to resource virtualization, global replication and migration, the physical absence of data and machine in the cloud, the stored data in the cloud and the computation results may not be well managed and fully trusted by the cloud users. Most of the previous work on the cloud security focuses on the storage security rather than taking the computation security into consideration together. In this paper, we propose a privacy cheating discouragement and secure computation auditing protocol, or *SecCloud*, which is a first protocol bridging secure storage and secure computation auditing in cloud and achieving privacy cheating discouragement by designated verifier signature, batch verification and probabilistic sampling techniques. The detailed analysis is given to obtain an optimal sampling size to minimize the cost. Another major contribution of this paper is that we build a practical secure-aware cloud computing experimental environment, or *SecHDFS*, as a test bed to implement SecCloud. Further experimental results have demonstrated the effectiveness and efficiency of the proposed SecCloud.

1. Introduction

The recent development of cloud computing has shown its potential to reshape the current way that IT hardware is designed and purchased. Among numerous benefits, cloud computing offers customers a more flexible way to obtain computation and storage resources on demand. Rather than owning (and maintaining) a large and expensive IT infrastructure, customers can now rent the necessary resources as soon as, and as long as, they need [1]. Thus, customers cannot only avoid a potentially large up-front investment (which is particularly attractive for small companies and startups), they may also be able to reduce their costs through economies of scale and by paying only for the resources they actually use.

Even though cloud computing is envisioned as a promising service platform for the Next Generation Internet [14], security and privacy are the major challenges which inhibit the cloud computing wide acceptance in practice [31]. Different from the traditional computing model in which users have full control of data storage and computation, cloud computing entails that the managements of physical data and machines are delegated to the cloud service providers while the users only retain some control over the virtual machines. Thus, the correctness of data storage and computation might be compromised due to

* Corresponding author. Tel.: +86 21 34204642.
E-mail address: zfcao@cs.sjtu.edu.cn (Z. Cao).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2013.04.028

the lack of the control of data security for data owners. In this study, we further classify cloud computing security into two major classes: *Cloud Storage Security* and *Cloud Computation Security*, where the former is referred to ensuring the integrity of outsourced data stored at untrustworthy cloud servers while the latter refers to checking the correctness of the outsourced computation performed by untrustworthy cloud servers.

Most of the current researches on secure cloud computing still focus on the cloud storage security. However, the outsourced computation security receives less attention. For sake of saving computation resources, the cloud servers may not perform the necessary computations but claim to have done so. Additionally, the centralized architectures emphasize the fact that the cloud servers can represent a single point of failure, as proven by the recent meltdown of Google's Gmail systems [25]. Under Byzantine [11] failure or even external attacks, the cloud may perform unreliable computation operations while choosing to hide the computations. This cheating behavior of the cloud servers, if undetected, may render the results useless. Even from the point of accountability, some secure computation mechanisms should be in place to meet the needs of deciding whether the cloud servers or the users should be responsible for it once there is any problem taking place. Note that, it is quite natural for the servers to initially suspect a problem with the customer's software, and vice versa [18].

Generally, due to the limitation of the computation and communication resources, the cloud users cannot afford the cost incurred by result auditing or verification. One promising approach to prevent the cloud users from incurring expensive verification costs is to introduce a trusted auditor who conducts cloud auditing on behalf of the users. Even though public auditability of secure storage in cloud [33,34] has been proposed in the context, public auditability in secure computation receives less attentions. More closely related references are secure remote computation in distributed system [24]. However, few of those proposed schemes target at secure cloud computation. Furthermore, privacy preserving is a critical issue for secure cloud computing while only a few of the existing researches [20,29] have taken it into consideration.

To achieve secure computing auditing in cloud, one straightforward method is to double-check each result. The cloud providers may give the inputs and overall computing result to the auditor, which will follow an identical procedure to compute the result and then compare it with the one provided by the cloud providers. However, these schemes may lead to a waste of I/O and computation resources. Note that the data transferring bottlenecks rank in the top of the ten obstacles which may prevent the overall success of the cloud computing [1]. In [13], a Commitment-Based Sampling (CBS) technique is introduced in the conventional grid computing however it does not take the privacy issue into consideration. In this paper, we introduce a novel technique by integrating CBS with the designated verification technique.

The contributions of this paper can be summarized as follows.

- Firstly, we model the security problems in cloud computing and define the concepts: *uncheatable cloud computation* and *privacy cheating discouragement* in our cloud computing, which are our design goals.
- Secondly, we propose a basic protocol, **SecCloud**, to attain data storage security and computation auditing security as well as privacy cheating discouragement and an advanced protocol to achieve computation and communication efficiency improvement through batch verification.
- Thirdly, we analyze and prove that **SecCloud** achieves our design goals and discuss how to minimize the computation cost by choosing the optimal sampling size.
- Finally, we develop a cloud computing experimental environment **SecHDFS** and implement **SecCloud** as a test bed. Experiment results demonstrate the suitability of the proposed protocol.

The remainder of this paper is organized as follows. A brief review on the related work is given in Section 2. Section 3 describes the system architecture and security problems and presents design goals. Some necessary preliminary knowledge is given in Section 4. We propose an overview of our **SecCloud** in Section 5 and then present an advanced **SecCloud** with performance optimization in Section 6. Section 7 gives out detailed security analysis and discussion. Section 8 introduces the experiment environment **SecHDFS** and implement our **SecCloud** as a test bed. Finally, Section 9 concludes the whole paper.

2. Related work

Security and privacy issues in cloud computing has received extensive attentions recently. Generally speaking, the research work on cloud computing almost falls into the two cases: *cloud storage security* and *cloud computation security*.

Cloud storage security mainly addresses the secure outsourced storage issue. In [2], Ateniese et al. first defined a model for *provable data possession* (PDP), which allowed a client that had data stored at an untrusted server to verify that the server possessed the original data without retrieving it. They utilized RSA-based homomorphic tags for auditing outsourced data, but they did not consider the dynamic data storage. In their later work, Ateniese et al. [3] proposed a partially dynamic version of the PDP scheme using symmetric key cryptography. However, it did not support public auditability. Juels et al. [22] proposed the definition of *proof of retrievability* (PoR), which used spot-checking and error-correcting codes to ensure both possession and retrievability for data file on archive service system. Wang et al. [34] first achieved both public verifiability and dynamic data storage operations employing an Third Party Auditor and improving the proof of retrievability model by using classic Merkle Hash Tree [26] construction for BLS [8] based block tag authentication. Later, they proposed a scheme achieving privacy preserving public verifiability as well as the dynamic data storage operations in [33] by utilizing the public