Blockchain-Enabled Reengineering of Cloud Datacenters

Keke Gai
Beijing Institute of Technology

Kim-Kwang Raymond Choo
University of Texas at San Antonio

Liehuang Zhu
Beijing Institute of Technology

Editor:
Kim-Kwang Raymond Choo
raymond.choo@fulbrightmail.org

Blockchains, a decentralized storage technique, have many applications, including in reengineering cloud datacenters. This article proposes a conceptual model for fusing blockchains and cloud computing for additional value creation. The proposed model comprises three deployment modes: Cloud over Blockchain (CoB), Blockchain over Cloud (BoC), and Mixed Blockchain–Cloud (MBC). The article also highlights the potential benefits of such a fusion and outlines a number of future research directions.

Cloud computing is now a deeply entrenched phenomenon in our society, partly due to the many potential benefits such as the capability to store and access large volumes of data, perform intensive computational operations, etc., on a pay-per-use basis. One of the most popular cloud services is storage as a service (StaaS). In StaaS, owing to the centralized physical machines at the cloud service providers, additional values may be available on the basis of requirements such as data mining or information fusion.[1]

There are, however, operational challenges that need to be considered in such a centralized setting. For example, a centralized setting does not necessarily imply that user data are physically stored in a single place, as the term normally refers to a layer. Multiple service providers may participate in service delivery; thus, data transfers often take place between cloud service providers. Another known challenge is the extent of the cloud user's control over his or her data.

There have also a number of extensions to the cloud to mitigate some of the existing limitations and/or to offer new functionalities. Examples include integrating the Internet of Things (IoT) or cyber-physical systems with cloud computing (also respectively known as the Cloud of Things or cyber-physical cloud systems) and extending the cloud to the edge (also known as edge computing). However, in such extended or integrated systems, user-data-related concerns remain (e.g., security and privacy of user data and the computation of the data).

Copublished by the IEEE CS and IEEE ComSoc
2325-6095/18/$33.00 ©2018 IEEE

There have also been interest in utilizing blockchains to mitigate some of the challenges in these extended or integrated systems, owing partly to the inherent properties of blockchains (e.g., fault tolerance and tamper resistance). However, most blockchain-based solutions are not completely integrated, in the sense that each technique normally serves only one subsystem. For instance, blockchains are used to record prior transactions, such as metadata describing source addresses, while the cloud database is responsible for physical storage. In other words, blockchains are not considered a component at the control layer.

In this article, we discuss the potential fusion of blockchains and cloud computing from the perspectives of secure data transfer and transparent data usage. We then propose a three-layer model, with the aims of reengineering cloud datacenters using blockchains.

In the next section, we describe some of the challenges that motivate this research (and the design of the three-layer model).

RESEARCH MOTIVATIONS

Centralization in cloud computing is considered mainly from the cloud users' perspective, as the cloud service provider is collectively considered a single entity rather than dozens to hundreds or thousands of physical machines deployed in a distributed geographical manner. Workload offloading takes place not only between users and cloud service providers but also between providers. The virtual machine (VM) technique plays a vital role in organizing and connecting cloud resources, in which a pool of VMs form the layer of centralized computing. For an individual user, cloud services are always virtually offered by a single cloud service provider via an interface.

There are a number of challenges associated with the management of cloud datacenters, such as service outage (downtime), payment management, governance (or lack of), and multicloud integration, in addition to security and privacy issues. The latter two issues are the focus of this article.

Privacy is often cited as one of the key concerns in cloud adoption or deployment[2]—for example, when sensitive or personal information is outsourced to the cloud service vendors. A number of privacy-preserving techniques have been proposed in the literature, such as differential-privacy mechanisms. However, the lack of controls on the cloud side and interconnections and interactions between vendors may be abused by adversaries to launch an attack, such as triggering linkage attacks against differential-privacy protection methods.[3] Screening user information is effectively secure when access to the supportive database (for attack purposes) is prohibited. Otherwise, it can be challenging to avoid privacy leakage against attacks such as data-mining-based linkage attacks.

An adversary can also target communication over the wireless medium—for example, by impersonating a legitimate user and surveilling or intercepting signals.[4,5] Multichannel communication,[6] for instance, can potentially enhance the data security level, as can implementing a variety of security protocols to increase the complexity of sensitive information retrieval. However, multichannel-based solutions are hardly useful in tracing data usage. Hence, this is one challenge that needs to be addressed.

In addition, cloud service providers may be compromised or their servers breached. So, there have also been attempts to encrypt user data to minimize the impact of a compromise yet maximize utility, using solutions such as fully homomorphic encryption (FHE).[7] However, FHE does not provide resilience against tamper-related attacks.[8] Hence, this is the second research motivation for this article.

REENGINEERING CLOUD DATACENTERS

A blockchain is a chain-enabled distributed ledger that provides tamper-resistant data storage functionality, such as transactions between parties. Broadly speaking, there are three core elements to be included in a blockchain system: the block, the chain, and the activity.

Specifically, a block is the storage carrier based on a consensus agreement by all stakeholders or validators. The storage content also captures the interactions between the different parties, such as transactions in Bitcoin. Similarly to cloud computing, an activity in a blockchain system can be represented in a "service" manner. For example, a digital transaction can be deemed the service content in Bitcoin. In addition, a chain is a method of connecting all blocks in a "one-way" growing manner. The one-directional chain growth is a critical element of determining its tamper-resistant characteristic.

To achieve fusion, we need to determine the connection or connections between the two techniques—in our context, the blockchain and the cloud. At first glance, both the blockchain and cloud appear to have contradictory architectures—namely, decentralization versus centralization. There are, however, a number of complementary connections. First, cloud-computing resources or services can be deemed to be services in both blockchain and cloud systems. Second, both techniques can be used for, or facilitate, storage. Third, both techniques can complement each other, as we will demonstrate in this article.

Specifically, we present our Blockchain–Cloud Fusion (BCF) model. Because pseudo names are allowed in a blockchain system, we can leverage this feature to achieve the privacy-preserving requirement. The proposed model is also designed to ensure security over wireless communication.

Who is deployed over whom? In our proposed BCF model, there are three types of deployment: Cloud over Blockchain (CoB), Blockchain over Cloud (BoC), and Mixed Blockchain–Cloud (MBC).

CoB refers to adopting blockchain techniques in cloud computing, in the sense that we use blockchains as a functionality tool to develop a subsystem in the cloud solution. For example, blockchains can be adopted as a back-end technique (back end as a service) to produce a tamper-resistant storage function. From the perspective of the service, the blockchain is considered a type of service—blockchain as a service (BcaaS). A blockchain created for a certain purpose can be designed as a cloud-enabled service.

In BoC, the blockchain structure organizes the service architecture, where cloud is deemed a way of delivering services. The connection between the blockchain and cloud computing, in this case, is generally related to the smart-contract implementation on the blockchain. In this service model, cloud computing plays the role of offloading when a complex computation workload is required, say, by the smart contract during strategy making, data mining, optimization, and so on. While this model is executed, a complex smart contract can be processed in the blockchain without requiring a costly blockchain server.

The underlying principle in MBC is using the blockchain to record the data usage in cloud systems. This allows us to support the involvement of multiple cloud service providers and ensure the transparency of data usage to data owners. However, this deployment mode requires all service providers to be in agreement so that each data transfer or share will be recorded by a block on the basis of the consensus. The implementation of the MBC utilizes the feature of the blockchain's consensus as well as the characteristic of multiproviders. In other words, service records can be stored by a distributed ledger in the blockchain so that the recorded content cannot be changed. The identity of data users and service providers can also be validated by applying consensus over the blockchain.

CONCLUSION

The deployment modes presented in the preceding section are determined by the role of the blockchain in the actual system architecture.

In a CoB deployment, the blockchain plays a supplementary role in a cloud system, such that cloud datacenters are key to the system architecture. Tamper-resistant performance is applied only in the subsystem or subsystems in which the blockchain is used, as are other blockchain features (e.g., traceable data usage). The main benefits of this deployment mode include ease of establishment, a checkable specific function, and expansion-friendliness for existing cloud solutions. However, limitations include

- hard-to-use data over an entire consensus in the system and
- the challenge of connecting varying cloud subsystems when each subsystem runs its own blockchain or blockchains.

A BoC deployment emphasizes the role of the blockchain and the functionality of the clouds. In fact, utilizing this mode to reengineer cloud datacenters will further highlight the centralized-computing role of cloud computing, because in this "X as a service" service model the smart contract can take on complex or computationally heavy workloads. A primary advantage of this deployment is that blockchain-enabled applications are strengthened due to the power-up computation capability. Designing a smart contract on a blockchain will become a relatively simple job if the offloading option is available. A disadvantage is that the deployment will have limited influence on reengineering cloud datacenters, because the cloud plays a supporting role. In other words, this is a blockchain system.

Unlike the above two deployment modes, both blockchains and cloud computing play an equally important role in MBC-enabled applications. Blockchains' technical features are fully utilized in reengineering cloud datacenters. The virtue of this deployment is that many drawbacks of cloud datacenters can be addressed, to achieve both privacy-preserving identity validation in wireless communications and tamper resilience. However, embedding blockchain systems in validation operations and service deliveries with records requires a large reengineering effort.

Our future research will include implementing and evaluating a prototype of the proposed model, as well as investigating potential extensions to this work. For example, we will investigate the potential of having multiple blockchain systems interconnected with consented agreement over all participants.

REFERENCES

1. Y. Zhang et al., "Home M2M networks: architectures, standards, and QoS improvement," *IEEE Communications Magazine*, vol. 49, no. 4, 2011.
2. Z. Zhang et al., "When privacy meets economics: Enabling differentially-private battery supported meter reporting in smart grid," *Quality of Service (IWQoS), IEEE/ACM 25th International Symposium on*, 2017.
3. P. Kairouz, S. Oh, and P. Viswanath, "The composition theorem for differential privacy," *IEEE Transactions on Information Theory*, vol. 63, no. 6, 2017.
4. Q. Yan et al., "Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, 2016.
5. D. Puthal et al., "Threats to networking cloud and edge datacenters in the Internet of Things," *IEEE Cloud Computing*, vol. 3, no. 3, 2016.
6. K. Gai et al., "Privacy-preserving content-oriented wireless communication in internet-of-things," *IEEE Internet of Things Journal*, vol. 5, no. 4, 2018.
7. K. Gai and M. Qiu, "Blend arithmetic operations on tensor-based fully homomorphic encryption over real numbers," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, 2018.
8. W. Wang et al., "Exploring the feasibility of fully homomorphic encryption," *IEEE Transactions on Computers*, vol. 64, no. 3, 2015.

ABOUT THE AUTHORS

Keke Gai is an associate professor in the Beijing Institute of Technology's School of Computer Science and Technology. His research interests include cybersecurity, cloud computing, blockchains, combinatorial optimization, and edge computing. Gai received a PhD in computer science from Pace University. Contact him at gaikeke@bit.edu.cn.

Kim-Kwang Raymond Choo holds the Cloud Technology Endowed Professorship in the Department of Information Systems and Cyber Security at the University of Texas at San Antonio. His research interests include cyber and information security and digital forensics. Choo received a PhD in information security from Queensland University of Technology. He's a senior member of IEEE and a Fellow of the Australian Computer Society. Contact him at raymond.choo@fulbrightmail.org.

Liehuang Zhu is a professor at the Beijing Institute of Technology's School of Computer Science and Technology. His research interests include security protocol analysis and design, wireless sensor networks, and cloud computing. Zhu received his PhD in computer science from the Beijing Institute of Technology. He's the corresponding author. Contact him at liehuangz@bit.edu.cn.