A Comprehensive Study on Blockchain-based Cloud-Native Storage for Data Confidence

Hannie Zang
School of Electrical Engineering and Computer Science (EECS)
Gwangju Institute of Science and Technology (GIST)
Gwangju, Republic of Korea
hanizang@gist.ac.kr

Jongwon Kim
Artificial Intelligence Graduate School
Gwangju Institute of Science and Technology (GIST)
Gwangju, Republic of Korea
jongwon@gist.ac.kr

***Abstract*—This paper explores the concept of blockchain-based cloud-native storage, which combines the decentralized and immutable nature of blockchain with the scalability and agility of cloud-native architecture. We discuss the significant contributions of leveraging blockchain and cloud-native concepts in storage systems for edge computing, specifically focusing on data security, reliability, and scalability. Furthermore, we address the challenges associated with implementing blockchain and cloud-native concepts in edge storage systems, including network connectivity and resource constraints. By examining these challenges, we aim to provide insights into the practical considerations and potential solutions for effectively integrating blockchain and cloud-native technologies in edge storage environments.**

***Index Terms*—edge computing, storage system, cloud-native concepts, blockchain**

I. Introduction

Edge computing refers to a decentralized computing model where data processing, storage, and computation are performed closer to the edge devices or data sources, rather than relying solely on centralized cloud infrastructure [1]. It aims to reduce latency, improve real-time responsiveness, optimize bandwidth utilization, and address challenges associated with data-intensive applications and limited network connectivity.

Edge computing is being studied extensively recently for several reasons [2]. Firstly, the exponential growth of data generated at the edge, driven by IoT devices, sensors, and other edge devices, has created a need for localized storage and processing capabilities. By moving computation closer to the edge, edge computing enables faster data analysis and decision-making, which is critical for real-time applications and time-sensitive operations. Secondly, edge computing addresses the limitations of relying solely on centralized cloud infrastructure. The volume of data generated by edge devices can be overwhelming to transmit and store in the cloud, leading to network congestion, high bandwidth costs, and latency issues. Edge computing reduces the need for data transfer to the cloud by enabling local storage and processing at the edge, optimizing network bandwidth, and improving overall system performance.

In terms of storage systems in edge computing, several requirements should be considered:

- Security: Edge storage systems should implement robust security measures to protect data stored at the edge. Encryption, access controls, and authentication mechanisms are necessary to ensure the confidentiality and integrity of stored data.
- Reliability: Edge storage systems should ensure data durability and availability even in resource-constrained edge environments. Redundancy mechanisms and fault tolerance measures should be in place to handle failures and prevent data loss.
- Scalability: Edge storage systems should be designed to scale horizontally to accommodate the growing volume of data generated by edge devices. They should be able to handle increased data storage requirements without compromising performance.
- Data synchronization and management: As edge devices operate autonomously, storage systems need to support efficient data synchronization and management across distributed edge nodes. Data consistency and coordination mechanisms should be in place to maintain a unified view of data across the edge environment.
- Low latency: Storage systems in edge computing should provide fast access to data, minimizing latency in retrieving and storing information. This is crucial for real-time applications and time-sensitive operations that require quick response times.

To address the challenges and optimize storage systems in edge computing, blockchain and cloud-native concepts can be leveraged as Fig.1:

- Blockchain [3]: Blockchain technology can enhance data security, integrity, and trust in edge storage systems. Its decentralized and immutable nature ensures data immutability, tamper resistance, and audibility. Blockchain can provide a transparent and trustworthy framework for data sharing, access control, and secure transactions in edge environments.

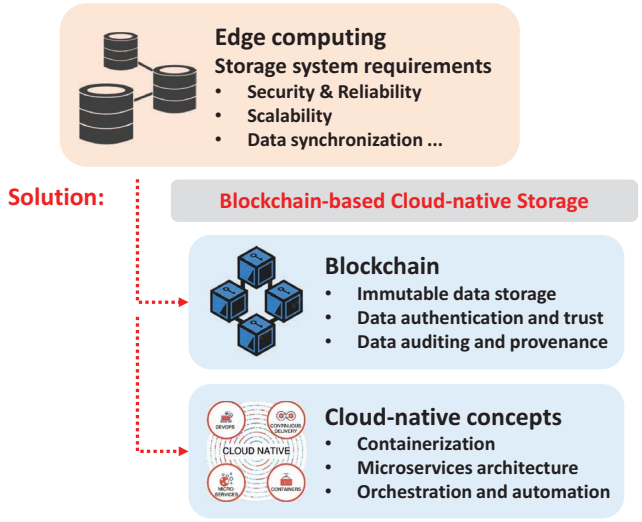

Fig. 1. Requirements of the storage system in edge computing and the blockchain-based cloud-native storage

- Cloud-native concepts [4]: Cloud-native concepts, such as containerization, microservices architecture, and orchestration tools like Kubernetes, can enable efficient resource utilization, scalability, and seamless deployment and management of storage services in edge computing. These concepts optimize storage performance, data availability, and resource management in distributed edge environments.

By integrating blockchain and cloud-native concepts, storage systems in edge computing can benefit from enhanced security, transparent data management, optimized resource utilization, and improved scalability. Blockchain provides a trust layer for secure data transactions, while cloud-native approaches enable efficient storage deployment and management in the distributed edge environment.

II. Blockchain-based Cloud-Native Storage

A. Definition

Blockchain-based cloud-native storage refers to a storage system that combines the principles of blockchain technology and cloud-native architecture. It integrates the decentralized and immutable nature of blockchain with the scalability and agility of cloud-native concepts to provide secure and efficient storage solutions. In this context, blockchain technology is used to ensure data integrity, transparency, and trust in the storage system. Data is stored in a distributed ledger, where each transaction is recorded and verified through cryptographic algorithms. This decentralized approach eliminates the need for a central authority and provides a tamper-proof and auditable record of data transactions. Cloud-native architecture, on the other hand, focuses on building applications and services that are designed to be scalable, portable, and resilient in cloud environments. It emphasizes containerization, microservices, and automation to enable rapid deployment, management, and scaling of storage resources.

By combining these two concepts, blockchain-based cloud-native storage offers several advantages. It provides a secure and transparent storage solution, ensuring data integrity and traceability. The decentralized nature of blockchain enhances data availability and fault tolerance, as data is distributed across multiple nodes. Cloud-native principles enable scalability and flexibility, allowing the storage system to adapt to changing demands and environments. Overall, blockchain-based cloud-native storage offers a robust and scalable approach to storing and managing data, leveraging the strengths of both blockchain and cloud-native concepts. It is particularly relevant in scenarios where data security, integrity, and transparency are paramount, such as in edge computing, IoT applications, supply chain management, and financial transactions.

B. Contributions

Blockchain and cloud-native concepts can significantly improve the storage systems in edge computing by addressing key challenges and providing enhanced capabilities. Here's how each of these concepts contributes to improving storage systems:

1) Blockchain:
   - Immutable data storage: Blockchain's decentralized and distributed ledger architecture ensures data immutability and tamper resistance. By storing data in a blockchain, it becomes virtually impossible to alter or modify it without consensus from the network participants, ensuring the integrity of stored data in edge environments.
   - Data authentication and trust: Blockchain enables secure and transparent data sharing, access control, and authentication mechanisms. It provides a trust layer where data transactions can be verified and validated, enhancing data security and trustworthiness in edge storage systems.
   - Data auditing and provenance: Blockchain's transparent and auditable nature enables a complete record of data transactions and modifications. This allows for better data governance and traceability, ensuring data provenance and accountability in edge computing environments.
2) Cloud-native concepts:
   - Containerization: By leveraging containerization technologies, such as Docker, storage systems can be packaged into lightweight and portable containers. This enables easy deployment, scaling, and management of storage services in the edge environment, ensuring efficient resource utilization.
   - Microservices architecture: Adopting a microservices architecture allows storage systems to be broken down into smaller, independent services. This modular approach enables flexible scaling, easier maintenance, and faster updates or replacements of specific storage components without disrupting the entire system.

- Orchestration and automation: Cloud-native tools like Kubernetes enable efficient orchestration and automation of storage resources in the edge environment. They facilitate dynamic allocation of storage capacity, load balancing, and fault tolerance, ensuring high availability and reliability of storage services.

Overall, by integrating blockchain and cloud-native concepts, storage systems in edge computing benefit from enhanced security, data integrity, scalability, efficient resource utilization, and improved performance. Blockchain ensures data security and trust, while cloud-native approaches enable the flexible deployment, management, and optimization of storage services in the dynamic and distributed edge computing environment.

C. Challenges

Leveraging blockchain and cloud-native concepts in storage systems for edge computing presents several challenges. Here are some of the key challenges to consider:

- Data privacy and security: Edge computing involves handling sensitive data, and maintaining data privacy and security is of utmost importance. Blockchain provides inherent data security through cryptography but ensuring secure data storage, access controls, and privacy-preserving mechanisms are still challenging, particularly in a distributed edge environment.
- Network connectivity: Edge computing environments often have intermittent or unreliable network connectivity. Blockchain networks require a consistent and reliable network connection for proper operation. Ensuring seamless data transfer, synchronization, and consensus in the presence of network disruptions or intermittent connectivity is a significant challenge.
- Scalability: Blockchain networks face scalability issues, especially when handling a large number of edge devices generating substantial amounts of data. Scaling a blockchain-based storage system to accommodate the growing data volume from edge devices while maintaining performance and efficiency is a complex challenge.
- Governance and regulation: Blockchain-based storage systems in edge computing raise governance and regulatory challenges. Determining the governance model, establishing consensus mechanisms, and complying with relevant regulations and policies become important considerations.
- Latency: Edge computing aims to process data in real time with minimal latency. However, blockchain networks often introduce additional latency due to consensus mechanisms and data validation processes. Addressing the latency challenge to meet the low-latency requirements of edge computing is crucial.
- Resource constraints: Edge devices typically have limited computing power, storage capacity, and network bandwidth. Implementing blockchain and cloud-native concepts can introduce additional overhead, potentially straining the limited resources of edge devices. Balancing resource usage and optimizing the system's performance become critical challenges.
- Interoperability and standards: Integrating multiple stakeholders, devices, and systems in a blockchain-based storage system requires interoperability and adherence to common standards. Achieving seamless integration and collaboration between various components and ensuring compatibility across different platforms and protocols is a complex challenge.

Overcoming these challenges requires careful design, optimization, and integration of blockchain and cloud-native concepts into the storage systems for edge computing. It involves developing efficient resource utilization strategies, scalability solutions, network resilience mechanisms, latency reduction techniques, robust data privacy and security measures, standardized interoperability frameworks, and appropriate governance models.

III. Conclusion

The integration of blockchain and cloud-native concepts in edge storage systems holds immense potential for enhancing data security, scalability, efficiency, and performance. By incorporating blockchain technology, the integrity, transparency, and trustworthiness of data transactions in edge environments can be significantly improved. Moreover, cloud-native concepts enable the deployment and management of storage services in the distributed edge environment with flexibility, scalability, and optimal resource utilization.

However, the successful implementation of blockchain and cloud-native concepts in edge storage systems requires addressing several key challenges. These challenges encompass resource constraints, scalability issues, intermittent network connectivity, latency concerns, data privacy and security, interoperability and standards, as well as governance and regulatory considerations. To overcome these challenges, it is crucial to focus on optimizing resource usage, standardizing protocols, integrating systems effectively, ensuring network resilience, safeguarding data privacy, and complying with regulatory frameworks.

Acknowledgment

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korean Government [Ministry of Science and ICT (MSIT)] under Grant 2023-2021-0-01835.

References

[1] Weisong Shi et al., "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp.637-646, Jun. 2016.

[2] Jianli Pan et al., "Future Edge Cloud and Edge Computing for Internet of Things Applications," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp.439-449, Oct. 2017.

[3] Ahmed Afif Monrat et al., "A Survey of Blockchain From the Perspectives of Applications, Challenges, and Opportunities," *IEEE Access*, vol. 7, pp.117134-117151, Aug. 2019.

[4] Dennis Gannon et al., "Cloud-Native Applications," *IEEE Cloud Computing*, vol. 4, no. 5, pp.16-21, Dec. 2017.