Blockchain applications to secure data transfer systems creation

Gleb Slepenkov[1][0009−0001−9978−396X]

Saint Petersburg State University, 7-9 Universitetskaya Embankment, St Petersburg, Russia, 199034

1 Introduction

In contemporary information societies, secure data transmission (SDT) is paramount for maintaining data confidentiality and integrity. SDT systems are critical infrastructure components across diverse sectors, including financial institutions, governmental agencies, and enterprise.

Implementing robust SDT mechanisms becomes particularly challenging when facilitating cross-organizational data exchange or connecting geographically dispersed departments within a single organization. Such scenarios necessitate data transfer across external, public networks, thereby significantly elevating the risk of unauthorized data disclosure or modification during transit. Consequently, rigorous data integrity verification and sender authentication protocols are indispensable. Furthermore, confirmation of data reception is often a critical requirement. The heterogeneity of data exchange technologies employed by different entities can further complicate the SDT process.

Blockchain represents an emerging technology for SDT solutions. This technology by design provides core features essential for SDT such as fault tolerance, ledger data immutability, zero trust between blockchain nodes. Moreover, many blockchains are designed to operate in unreliable public networks with nodes running in different heterogeneous clusters owned by separate stakeholders. By this reason blockchains are especially suitable for cross-organizational data transfer. Furthermore, blockchain platforms often provide smart contracts which allow to implement custom data transfer logic directly on blockchain.

This suitability has spurred significant research interest, as evidenced by comprehensive surveys exploring blockchain applications for data sharing and exchange [15] and for specific domains like smart transport [2]. These surveys highlight the potential of blockchain to address the challenges of SDT in diverse and demanding environments.

This study investigates existing paradigms of blockchain applications to SDT process and proposes a new approach, which allows to combine advantages of existing paradigms and create SDT blockchain-based system, which preserves key features of existing approaches, but has message broker like architecture and lax data size limitations. The study is organized in a following way. In section 2 existing data transfer paradigms are described in detail. In section 3 proposed system architecture is presented. In section 4 implementation details of created systems are described. In section 5 further research directions are provided.

2 Related Work

The integration of blockchain technology into data transfer systems presents two primary paradigms: blockchain as a data access interface and blockchain as a direct data transfer tool. This section examines the existing literature, classifying studies according to these two distinct roles of blockchain in facilitating data exchange.

2.1 Blockchain as a Data Access Interface

This approach leverages blockchain as a secure and transparent mechanism for managing and controlling access to data that is typically stored off-chain. The blockchain records metadata about the data, access permissions, and transaction histories, thereby ensuring data integrity, auditability, and secure access control. The data itself is transferred using conventional methods.

Wang et al. introduced BBS, a big data sharing system utilizing blockchain for access control and data integrity management [16]. Similarly, Wang et al. proposed a blockchain-based model for sharing big data in the oil and gas sector, employing blockchain to enable secure data access and provenance tracking [17]. Yang et al. developed a sharing platform for wild bird data based on blockchain and IPFS, where blockchain governs access rights to data residing in IPFS [19]. A cross-organizational data sharing framework that uses blockchain probes to orchestrate and audit data access across various entities was presented by Jia et al. [8]. Gupta et al. explored the application of blockchain for securing data access within e-healthcare applications [7].

2.2 Blockchain as a Data Transfer Tool

This approach employs the blockchain directly to transfer data between participants. While data may occasionally be stored on-chain, it is more common for the blockchain to facilitate the transfer of data stored off-chain, often in conjunction with technologies such as peer-to-peer networks or distributed file systems.

Lin et al. proposed a multi-level blockchain architecture for secure data transfer in the Internet of Vehicles (IoV), directly leveraging blockchain for data exchange [12]. Peng et al. demonstrated the feasibility of building a peer-to-peer file storage and sharing system on a consortium blockchain, where the blockchain assists in discovering and securely transferring file chunks [13]. Priyadarshini et al. introduced a system for secured data transfer between fog nodes utilizing blockchain to enhance the security of the data transfer process itself [14].

2.3 Enabling Technologies and Security Considerations

Regardless of the chosen paradigm, several enabling technologies and security considerations are universally relevant. Kim et al. presented a hybrid decentralized PBFT blockchain framework for OpenStack message queues, aimed at

improving fault tolerance and scalability, crucial aspects for reliable data transfer systems [10]. A publish-subscribe architecture to foster interoperability among different blockchain networks, thus facilitating data exchange across heterogeneous systems, was introduced by Ghaemi et al. [6]. Bagga et al. and Xu et al. explored blockchain-based authentication and key agreement protocols for IoV, enhancing the security of data transfer in vehicular contexts [3, 18]. Bogdanov et al. focused on the consensus mechanism, exploring a combination of PBFT and Raft [4], while Ai et al. proposed a Proof-of-Transactions consensus protocol [1], both striving for efficient and scalable agreement in distributed ledgers. A privacy-preserving authentication protocol specifically for VANETs was the focus of Lin et al. [11].

2.4 Comparison of Approaches

The two approaches, blockchain as a data access interface and blockchain as a data transfer tool, offer distinct advantages and disadvantages depending on the specific use case and requirements.

Table 1. Comparison of Blockchain Approaches for Data Transfer

Feature	Blockchain as Data Access Interface	Blockchain as Data Transfer tool
Data Storage	Primarily Off-Chain	Often Off-Chain, but metadata on-chain
Data Transfer	Traditional methods	Blockchain or Blockchain-assisted
Scalability	Higher, as data transfer is off-chain	Potentially Lower, dependent on blockchain throughput
Complexity	Lower	Higher
On-Chain Footprint	Smaller (Metadata only)	Larger (Metadata and potentially data chunks)
Suitability	Large datasets, access control focus	Smaller datasets, secure and direct transfer focus

As illustrated in Table 1, the choice between these two approaches hinges on factors such as data size, scalability requirements, security priorities, and the level of on-chain transparency desired. The data access interface approach is generally more suitable for scenarios involving large datasets and a primary focus on secure access control, while the data transfer tool approach is better suited for scenarios demanding secure and direct data transfer, even if it potentially introduces limitations in terms of scalability.

3 Proposed system architecture

3.1 Overview

This section describes the architecture of proposed system. The main goal of such an architecture is to make a combination of two main blockchain integration

into data transfer systems described in section 2 in order to create system with message broker like architecture, convenient for data transfer, and lax data size limitations comparing to existing solutions. Essentially architecture follows the trends described in related articles:

1. Blockchain usage as message broker: [6], [10]. The key drawback of theese solutions lies in usage of ledger for message transfer. It greatly limits the system scalability and transfered data size.
2. Blockchain as interface to off-chain data (e.g. [8], [16]). Such a solutions usually do not have any notification mechanisms. Hence, message receiving becomes a challenging task.

Proposed architecture assumes that system should contain the following components:

- Blockchain as main data transfer tool.
- Data storages internaly used by counterparties (organizations) to store the data to be sent;
- Connectors — some application used by SDT process counterparties in order to create connection between their data storages and blockchain nodes owned by organization (mainly inspired by blockchain probes described in [8])

A sample of such an architecture for SDT between two organizations is presented on figure 1.

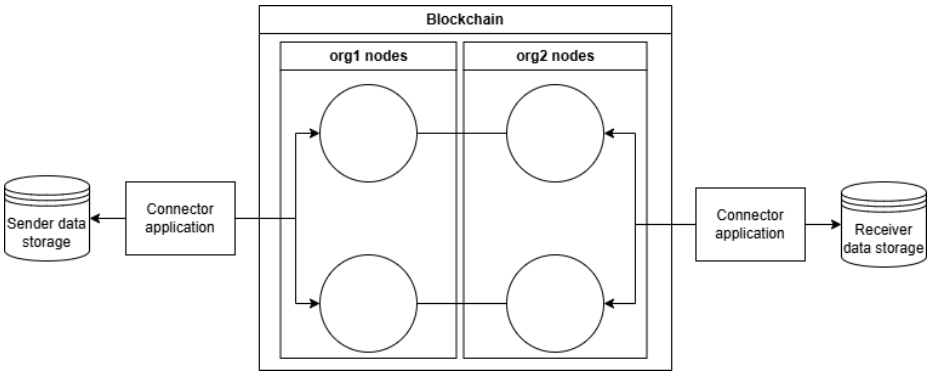

Fig. 1. System architecture for two organizations

3.2 SDT process scheme

The SDT process is presented on figure 2. It represents the following algorighm.

The data transfer process commences with the data sender (Org1) initiating a data transfer request to its connector, which serves as the interface to

the blockchain. The Org1 connector then persists the data in Org1's storage and gathers associated metadata. Subsequently, the Org1 connector queries the blockchain to initiate the execution of the Secure Data Transfer (SDT) smart contract. The smart contract validates the provided metadata and records it on the ledger. Upon successful metadata validation, the smart contract informs the Org2 connector about the availability of a new message. The smart contract also acknowledges the acceptance of the message delivery to Org1.

To retrieve the data, the Org2 connector queries the blockchain nodes and executes the SDT smart contract. The smart contract verifies Org2's authorization to access the requested data. If authorized, the smart contract requests the data from the Org1 connector. Upon receiving the request, the Org1 connector retrieves the data from Org1's storage and transmits it back to the smart contract.

The smart contract then prepares the data for transfer, which may involve encryption and digital signing depending on the specific smart contract implementation. The prepared data is subsequently returned to the Org2 connector. The Org2 connector then stores the received data in Org2's storage.

To finalize the process, the Org2 connector notifies the blockchain about the message reception via the smart contract. The smart contract updates the message registry, marking the message as delivered. The smart contract also notifies the Org1 connector about the successful completion of the message delivery process. Finally, the Org1 connector informs the data sender about the completion of the data transfer process.

Such an algorighm has following important features:

1. Loose coupling between sender and receiver due to blockchain usage as message broker.
2. Better fault tolerance comparing to regular message brokers.
3. More convenient message broker like SDT interface comparing to existing blockchain solutions.
4. Great SDT process customization capabilities due to smart contract usage.

4 Implementation related details

4.1 Blockchain selection

Blockchain is the core component of SDT system. Hence, the particular blockchain implementation has a great impact on SDT system parameters such as:

- Throughput — amount of data system is able to transfer during fixed time interval.
- Fault tolerance — fault types system is able to resist and maximum number of crashed nodes system is able to handle.
- Scalability — maximum number of nodes in the system.

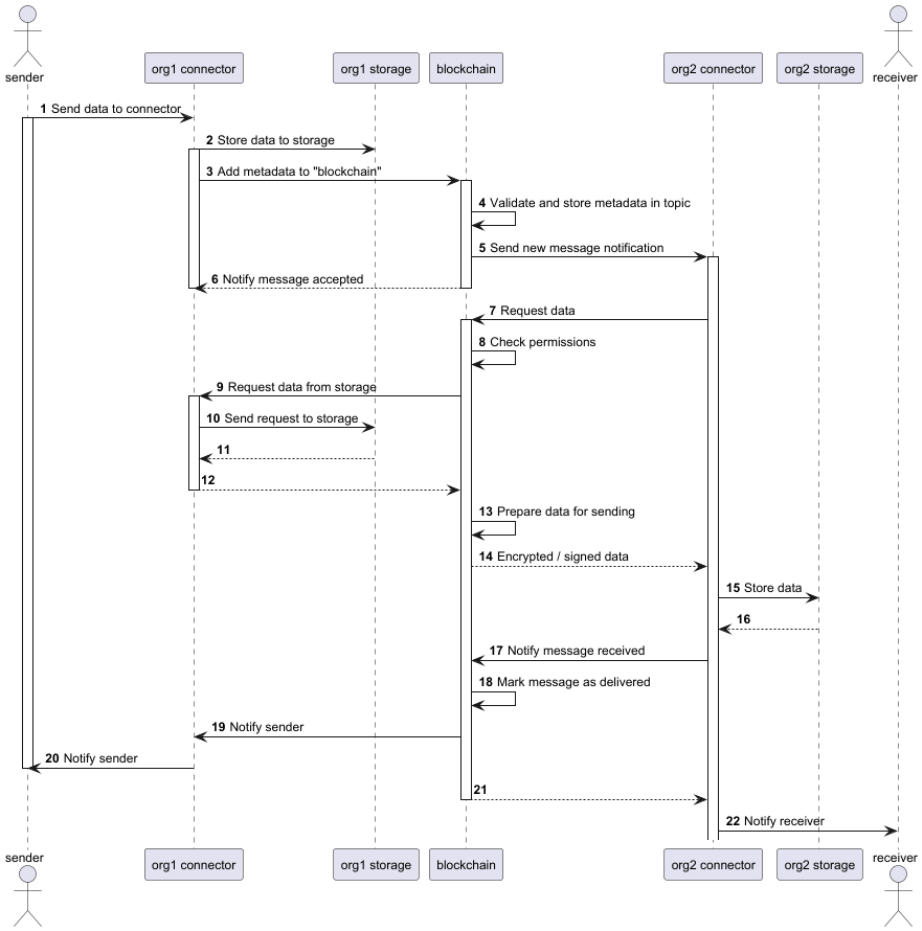

Fig. 2. SDT process scheme

All parameters presented above depends mainly on blockchain type and consensus algorighm used.

In literature reviewed a notable trend towards private blockchain solutions is observed. This preference stems from the need for controlled access, enhanced privacy, and potentially higher throughput in many SDT scenarios, particularly in applications like IoT and vehicular networks. Several papers mentioned the use of public blockchains but it wasn't their primary option. Specific examples include:

- [13] proposes a peer-to-peer file storage and sharing system based on a consortium blockchain.
- [8] presents a cross-organizational data sharing framework based on blockchain-probes, implying a permissioned setting.
- Several papers addressing IoT applications (e.g., [1, 7]) often implicitly or explicitly assume a permissioned blockchain context to address security and scalability constraints.

Thus, this paper consideres the private blockchain as the best option for SDT system implementation.

In addition to blockchain type selection it is important to use correct consensus algorighm. In order increase SDT system fault tolerance this paper proposes PBFT algorighm usage according to the trends observed in literature. However, PBFT has important limitations:

TODO describe PBFT limitations according to [4], [9], [5]

5 Further research directions

References

1. Ai, Z., Cui, W.: A proof-of-transactions blockchain consensus protocol for large-scale iot. IEEE Internet of Things Journal (2022). https://doi.org/10.1109/JIOT.2021.3108621
2. Bagga, P., Das, A.K.: Blockchain for Smart Transport Applications, pp. 125–154. Springer International Publishing, Cham (2022)
3. Bagga, P., et al.: Blockchain-based batch authentication protocol for internet of vehicles. Journal of Systems Architecture **113**, 101877 (2021)
4. Bogdanov, A., et al.: Combining pbft and raft for scalable and fault-tolerant distributed consensus. Physics of Particles and Nuclei **55**(3), 418–420 (2024)
5. Fan, C., et al.: Performance evaluation of blockchain systems: A systematic survey. IEEE Access **8**, 126927–126950 (2020)
6. Ghaemi, S., et al.: A pub-sub architecture to promote blockchain interoperability. arXiv preprint arXiv:2101.12331 (2021)
7. Gupta, S., Yadav, B., Gupta, B.: Security of IoT-based e-healthcare applications using blockchain, pp. 79–107. Springer International Publishing, Cham (2022)
8. Jia, X., et al.: Cross-organisational data sharing framework based on blockchain-probes. IET Networks **12**(2), 77–85 (2023)

9. Ke, Z., Park, N.: Performance modeling and analysis of hyperledger fabric. Cluster Computing **26**(5), 2681–2699 (2023)
10. Kim, Y., Park, J.: Hybrid decentralized pbft blockchain framework for openstack message queue. Human-centric Computing and Information Sciences **10**(1), 31 (2020)
11. Lin, C., et al.: Bcppa: A blockchain-based conditional privacy-preserving authentication protocol for vehicular ad hoc networks. IEEE Transactions on Intelligent Transportation Systems **22**(12), 7408–7420 (2020)
12. Lin, H.Y.: Secure data transfer based on a multi-level blockchain for internet of vehicles. Sensors **23**(5), 2664 (2023)
13. Peng, S., et al.: A peer-to-peer file storage and sharing system based on consortium blockchain. Future Generation Computer Systems **141**, 197–204 (2023)
14. Priyadarshini, R., Malarvizhi, N.: Secured data transfer between fog nodes using blockchain. In: Proceedings of the 2nd International Conference on Computational and Bio Engineering: CBE 2020. pp. 417–422. Springer Singapore (2021)
15. Song, R., et al.: A survey of blockchain-based schemes for data sharing and exchange. IEEE Transactions on Big Data (2023)
16. Wang, S., et al.: Bbs: A secure and autonomous blockchain-based big-data sharing system. Journal of Systems Architecture **150**, 103133 (2024)
17. Wang, Y.Y., Huang, S., Yu, X.: An oil and gas big data sharing model based on blockchain technology. IOP Conference Series: Earth and Environmental Science **651**(3), 032105 (2021)
18. Xu, Z., et al.: A blockchain-based roadside unit-assisted authentication and key agreement protocol for internet of vehicles. Journal of Parallel and Distributed Computing **149**, 29–39 (2021)
19. Yang, H., et al.: A research on the sharing platform of wild bird data in yunnan province based on blockchain and interstellar file system. Sensors **22**(18), 6961 (2022)