A NEW APPROACH TO THE DEVELOPMENT OF PROVENANCE METADATA MANAGEMENT SYSTEMS FOR LARGE SCIENTIFIC EXPERIMENTS

A.P. Demichev [1,a], J.Yu. Dubenskaya [1], A.P. Kryukov [1], S.P. Polyakov [1], N.V. Prikhod'ko [2]

[1] *Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia*

[2] *Yaroslav-the-Wise Novgorod State University, Velikiy Novgorod, Russia*

E-mail: [a] demichev@theory.sinp.msu.ru

Provenance metadata (PMD) contain key information that is necessary to determine the origin, authorship and quality of relevant data, their storage and usage consistency, and for interpretation and confirmation of relevant scientific results. The need for PMD is especially important when Big Data are jointly processed by several research teams, which is a very common practice in many scientific areas of late. Although a number of projects have been implemented in recent years to create management systems for such metadata, the vast majority of the implemented solutions are centralized, which is poorly suited to current trends of working in distributed environments and using metadata by organizationally unrelated or loosely coupled communities of researchers. We propose to solve this problem by employing a new approach to creating a distributed registry of provenance metadata based on blockchain technology and smart contracts. We have investigated the problem of the optimal choice of the type of blockchain for such a system, as well as the optimal choice of the blockchain platform. The architecture and algorithms of the system operation, as well as its interaction with the distributed storage resources management systems, are proposed.

Keywords: distributed storage, provenance metadata, blockchain, access rights, Hyperledger

1. Introduction

Currently, the implementation of large scientific, engineering and business projects is associated, as a rule, with the need to store and process large amounts of data. This leads to development of new, more economical and reliable, architecture and operating principles of information systems, including storage systems. Extreme options for architectural solutions for the latter are fully centralized and fully decentralized (peer-to-peer, P2P) storages. However, often such solutions are unacceptable, in particular because of their high cost or low reliability. In many use cases, the intermediate solution between the fully centralized and fully decentralized ones may prove to be optimal [1]. To implement such a solution, organizations participating in a large project pool their local resources into a unified distributed storage and, if necessary, additionally rent cloud storage resources, possibly from several providers. Especially profitable from the economic and technical points of view, this solution can be in the case when there is a need to store large amounts of data during a limited period of the project implementation and in a situation where the project brings together many organizationally unrelated participants. In general, such a distributed storage pool forms a dynamically changing environment (the new storage can enter the pool and the other can leave it), and the local storages entering the pool can have different data management systems. The challenge is to combine all these storages and data in them into a single system in a dynamically changing environment, as well as ensure the implementation of reciprocal access policies to the data of the parties involved. For example, the owner of data (the user who created these data or the organization to which they belong) should be able to manage access rights to them for other users. Another example is the ability of a cloud storage to grant access to data stored on it only to users from organizations that have paid for the provision of the storage services. This implies the availability of decentralized methods both for data access management in such a dynamically changing environment and for ensuring a reliable, immutable record of the history of committed transactions, that is, provenance metadata (PMD), for examination and resolving possible conflicts of project participants between themselves and with storage owners. Conflicts can be related to priority issues when obtaining the results of data processing, use of results, violation of access rights, etc.

In other words, it is necessary to provide tools to support the implementation of business processes of storage and exchange of scientific data in a distributed environment with administratively unrelated or loosely related organizations participating in joint projects or simply exchanging data on certain conditions. First of all, this requires a provenance metadata registry that is resistant to malicious changes as well as a method of ensuring consensus among participants in the business process about the content and order of transactions with data.

It should be noted that although a number of projects have been implemented in recent years to create systems for metadata storage and management, including the provenance of data, the vast majority of the implemented solutions are centralized [2, 3], which is poorly suited to distributed dynamically changing environment, and the possibility of using metadata by organizationally unrelated research communities. On the other hand, in recent years, distributed registries based on blockchain technology have become very popular in various applied areas due to a number of important advantages [4, 5]. Most recently, on the basis of the blockchains, developments have also been appeared for the PMD management systems [6, 7]. However, they are designed to work with one storage, do not solve the problem of providing business process for data exchange between administratively different organizations and data access managment.

2. Distributed storage with PMD driven data management

The basic scenario of using the proposed system assumes that a virtual organization (VO) is formed for the joint implementation of a certain project. VO includes several real organizations which, in turn, include data providers, data handlers and users affiliated with them. It is assumed that the implementation of such a project requires the use of a distributed data storage. This distributed storage can be formed by renting multiple cloud storage, as well as integrating the own storage resources of the organizations that form the VO. Thus, the hardware and software basis of the business environment in this case is formed by a set of storages (possibly of different types, e.g., cloud storages,

file servers, tape storages, etc.), each of which can be managed by its own data management system (DMS). Generally speaking, several VOs can coexist; the storages with which they interact can form partially overlapping sets. In addition to the task of recording the immutable history of working with data in a distributed storage environment, the task of providing distributed management of access rights to data is set. A natural solution for the establishment of a distributed immutable registry for the PMD records is the use of the blockchain technology. The latter guarantees that no records were inserted into the registry in hindsight, no entries were changed in the registry and the registry has never been branched or bifurcated. An important question is how to provide validation of the chain of blocks with transaction records in the case of PMD registry. The use of the most popular proof-of-work (PoW) method [5] on the basis of mining is very resource-intensive, and is poorly suited for management systems for provenance metadata for the processing of scientific data. Indeed, the calculations that are performed within the framework of PoW themselves do not serve any useful purpose, and this is a principle feature. It is very difficult to come up with a proof of work that would serve a socially useful role. Therefore, if possible, it is better to abandon it. Trying to solve these problems, a community of researchers in this field offers a variety of consensus algorithms that do not require "work". The choice of the algorithm heavily depends on the way of access to transaction processing. From this point of view, blockchains are classified as follows:

- permissionless (public) blockchains, in which there are no restrictions on the transaction handlers;
- permissioned blockchains, in which transaction processing is performed by specified entities.

Public blockchains are more known because cryptocurrency networks are based on them. In contrast to the permissionless blockchains, in the systems based on permissioned blockchains, the built-in coins are usually not used. Built-in coins are required in permissionless blockchains to provide a reward for processing transactions. Permissioned blockchains can form a more controlled and predictable environment than public blockchains and does not require calculations related to the PoW algorithms. In the distributed storage environment, the local data management systems, data owners, representatives of real organizations participating in the project, etc., can act as the authorized parties that create and sign the blocks. In order to maliciously change a transaction confirmed by all the authorized parties in the distributed storage environment, the attacker must gain access to all the secret keys of the block handlers. This is very unlikely, and thus this approach provides a high degree of protection for the distributed registry. It is this approach to the construction of the metadata registry that was implemented in our PMD management system.

To put this solution into practice, it is convenient to use existing blockchain platforms. Analysis of existing platforms shows that the required solution for the PMD management system most naturally can be implemented on the basis of the Hyperledger Fabric permissioned blockchain platform (HLF; www.hyperledger.org) [8] together with Hyperledger Composer (hyperledger.github.io/composer). The latter is a set of tools for simplified use of the blockchain. Hereafter we shall refer to these two components as HLF&C-platform. To describe the business process within the framework of HLF&C-platform, a number of concepts are used, the main ones are assets, participants, transactions and events. Assets are tangible or intellectual resources, services or property, records of which are kept in the blockchain. Assets must have a unique identifier, but they can also contain any properties defined for them. Participants are members of the business network which can own assets and make transaction requests. They also can have any properties if necessary. Transaction is the mechanism of interaction of participants with assets. Messages about the events can be sent by transaction processors to inform external components of changes in the blockchain. Very important that HLF&C-platform provides the operation of smart contracts (called chaincode), which allows us to organize the business process of sharing storage resources by project participants located in different administrative domains. The suggested system for managing provenance metadata, entitled ProvHL (Provenance HyperLedger), is a sophisticated adaptation of the HLF&C-platform for the business process of sharing storage resources.

From the general point of view, two approaches are possible. In the first approach, data management systems (DMS) manage data and use a blockchain simply as a distributed log (data driven data management). In the second approach, the metadata is written to the blockchain beforehand, and DMSs refer to the blockchain and performs the transactions recorded there (metadata

driven data management). In the first case, the functionality of the blockchain system is very limited, it only provides a distributed ledger which is resistant to occasional or malicious attempts to modify the history of data in distributed storage. HLF&C-platform enables one to implement the second approach, which in addition to simply maintaining the ledger allows us to solve the problem of distributed data access management.

In our case, participants (in the sense of the HLF&C-platform) include persons (users and administrators of different levels) and storage providers. The main assets are data files. Their properties (attributes) are provenance metadata, including local file name in a storage, storage ID, creator ID, file owner ID, type of the file (primary, secondary or replica), etc. Another important type of the assets are (local) storages constituting the distributed environment. We also defined user groups as assets, because we found it useful for managing data access rights. Finally, operations with files are treated as assets too because each operation actually comprises of a several atomic transactions. The basic operations can be of the following types: new file upload; file download; file copy within a storage; file deletion; file copy to another storage; file transfer to another storage.

The algorithm which we propose for recording transactions with provenance metadata and managing data access rights in the framework of ProvHL in a very simplified form reads as follows:

- the owner accesses the chaincode function, which, according to the acl-file ("acl" stands for access control language), allows the owner of the data to grant access rights to these data to another user or group of users;
- a user who is granted access rights by the owner accesses the chaincode with a request to make an operation (ClientRequest transaction) with data (for example, file download, upload or copy);
- chaincode verifies that such a transaction complies with the rules defined in the acl-file and, if it does, sends a request to the HLF environment to complete the transaction;
- HLF performs transaction processing (transaction workflow: simulation/endorsements — ordering — validation — state updating);
- HLF sends a message (event) to the user about the successful transaction and its recording in the blockchain; the message also contains the transaction identification number;
- the user accesses the data management system (DMS) with a request to perform a data operation that contains the number of the corresponding transaction;
- DMS checks for a record of this transaction in the blockchain;
- if there is a record of the valid transaction, the DMS performs the required operation and, in turn, initiates a transaction record confirming that a data operation was performed (ServerResponse transaction).

As it can be seen, for each data operation, at least two transaction records are made in the blockchain: one corresponds to the client request (ClientRequest), and the second corresponds to the server response (ServerResponse). In general case, an operation comprises of even more transactions. In the simplified description of the algorithm, some details specific to certain types of transactions are omitted for brevity. In particular, when the "new file upload" operation is performed, the creation of the new asset, that is the data file, is performed only after the actual upload of the file in the storage when DMS makes a ServerResponse transaction and turns the uploaded file into a fully valid asset.

Together with the above-mentioned splitting of transactions into the client and server parts, this makes the level of correspondence between the history recorded in the blockchain and the real history of the data in the distributed storage practically acceptable.

3. Conclusion

In this paper, through the use of a new approach based on the integration of blockchain technology, smart contracts and metadata driven data management, the principles and algorithms of the system, entitled ProvHL (Provenance HyperLedger), are developed that are fault-tolerant, safe and secure management system of provenance metadata, as well as access rights to data in distributed storages. The problems of optimal choice of the blockchain type for such a system, as well as the choice of the blockchain platform are studied. Namely, it is proposed to use a permissioned type of

blockchain and the Hyperledger blockchain platform, on the basis of which the ProvHL system is implemented.

At present, a testbed has been created on the basis of SINP MSU, where a preliminary version of the ProvHL prototype is deployed to implement the developed principles and refine the algorithms of the system. The creation of ProvHL production level system will significantly improve the quality and reliability of the results obtained on the basis of processing and analysis of Big Data in a distributed computer environment.

Acknowledgement

This work was funded by the Russian Science Foundation (grant No 18-11-00075).

References

[1] Kryukov A.P. and DemichevA.P. Architecture of Distributed Data Storage for Astroparticle Physics // Lobachevskii Journal of Mathematics. 2018. V. 39. No. 9. pp. 1199–1206

[2] Zafar F., Khan A., Suhail S., Ahmed I., Hameed K., Khan H.M., Jabeen F. and Anjum A. Trustworthy Data: A Survey, Taxonomy and Future Trends of Secure Provenance Schemes // Journal of Network and Computer Applications. 2017. V. 94. PP. 50-68

[3] da Cruz S. M. S., Campos M. L. M. and Mattoso M. Towards a Taxonomy of Provenance in Scientific Workflow Management Systems. // In: World Conference on Services-I. July 2009. pp. 259-266. - DOI: 10.1109/SERVICES-I.2009.18

[4] Staff E. Blockchains: The Great Chain of Being Sure about Things // The Economist. 2016. Available at: https://www.economist.com/news/briefing/21677228-technology-behind-bitcoin-lets-people-who-do-not-know-or-trust-each-other-build-dependable, (accessed 28.09.2018)

[5] Baliga A. Understanding Blockchain Consensus Models // Tech. rep., Persistent Systems Ltd. 2017, pp. 1-14. Available at: https://pdfs.semanticscholar.org/da8a/37b10bc1521a4d3de925d7ebc44bb606d740.pdf. (accessed 28.09.2018)

[6] Ramachandran A. and Kantarcioglu M. SmartProvenance: A Distributed, Blockchain Based Data Provenance System. // In: Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy. March 2018. pp. 35-42. - DOI: 10.1145/3176258.3176333

[7] Liang X., Shetty S., Tosh D., Kamhoua C., Kwiat K., and Njilla L. Provchain: A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability // In: Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2017. pp. 468-477. - DOI: 10.1109/CCGRID.2017.8

[8] Androulaki E., Cachin C., Ferris C., Muralidharan S., Murthy C., Nguyen B., Sethi M., and Stathakopoulou C. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains // In: Proceedings of the Thirteenth EuroSys Conference, April 2018. pp. 30 – 45. - DOI: 10.1145/3190508.3190538