The Blockchain Technology in State Institutions

Yakov M. Dalinger[1], Evgeny A. Saksonov[2]
St. Petersburg State University of Civil Aviation
St. Petersburg, Russia
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]iakovdalinger@gmail.com, [2]saksmiem@mail.ru

Abstract. **The problems of building a distributed registry (blockchain) for state institutions are considered. The main tasks, the solution of which is necessary when creating the registry, are highlighted. The model of formation of registry blocks is given.**

Results can be useful to designers and administrators of the state institutions distributed registry .

Key words: **blockchain, distributed registry, mathematical model, state institution.**

I. INTRODUCTION

The blockchain technology is becoming increasingly popular in creating information systems of state institutions at the federal, regional and municipal levels due to its features [1]:

- the impossibility of making changes to the entries (copies) created in the registry;
- fixation of all changes in previously made records in the form of new records;
- high functional reliability, associated with the presence of copies;
- a full match of all available copies of the registry;
- the possibility of unlimited expansion of the amount of the registry.

These features allow the use of blockchain technology in building archives and specialized information systems of various levels and purpose (for example, systems of personal data processing, systems of financial organizations, ministries and departments) where it is required to ensure high reliability of information storage, control of access to data and be able to increase the amounts of stored information.

Below we will use the synonyms: blockchain, the distributed registry or the registry [5].

II. DESCRIPTION OF THE PROBLEM

The processes of applying of the systems based on blockchain technology are significantly different from the known methods of creating and operating cryptocurrency related systems, and those differents must be considered when developing.

The main differences are as follows:

- the lack of the need to compete for the right to enter data into the register and receive remuneration;
- territorial localization of copies of register in the locations of representative offices of state institutions (for example, within the Russian Federation, region, municipal district, etc.);
- compliance of standards governing the processing and access to information in each case;
- ensuring the protection of information in the registry from specific threats related to the functioning of the registry, the properties of stored information and the requirements of state institutions;
- ensuring access to the registry of various groups (types) of users (employees of state institutions, representatives of external organizations, citizens of the Russian Federation) in compliance with the established access rights;
- the availability of special means of checking data recorded in the register, depending on the purpose of the data, the specific tasks of the state institution;
- compliance with the standards for the response time to requests, recovery from failures.

In addition, copies of the registry contain a large amount of various information, structured as blocks often presented in an encrypted form, which can complicate the search for data in the registry and requires the creation of means of its presentation in a convenient form for users of the registry.

III. REGISTRY CREATION TASKS

When creating a systems on base blockchain technology for a particular state institution, are require solution the following tasks:

- Development of methods for creating and provided operating a distributed registry. This includes determining the structure of the registry, the composition of the server groups for the registry (it is possible to store each copy on several servers), the procedure for forming data blocks, and writing blocks to the registry.
- Development of the structure of records and blocks of the distributed registry. This will allow you to more easily obtain a representation of registry entries in the required form for processing.
- The selection of different types of registry users and the assignment for users each type of access rights. Users can be employees of state institutions and have rights write and read or it can to be particular persons with only read rights.
- Development of methods for ensuring the coherence of data in accessible copies of the registry. After the blocks

are entered in the register, obtaining coherent copies is possible using various known methods, depending on the requirements of the state institution.

- Creation of mathematical models for evaluating the characteristics of a distributed registry depending on its organization, parameters and purpose (reliability, length of the procedure for forming and inserting blocks and ensuring the coherence of data in copies, response time to user requests, recovery time after of failurs, protection from threats attacks of various kinds).
- Simulation and development of recommendations for the registry operation depending on its parameters (the number of copies and their location, the number of users, the number of users of various types, etc.).
- Creation of a prototype for carrying out field tests, matching models with actual conditions.
- Creation of basic software (development of requirements and composition of a software, development of a prototype of basic software modules, etc.).
- Development of guides for the creation and operation of a distributed registry, taking into account the purpose and conditions of operation.

There are possible solutions to the above registry creation tasks. A possible solution for creating a registry and organizing its work is shown in Figure 1.

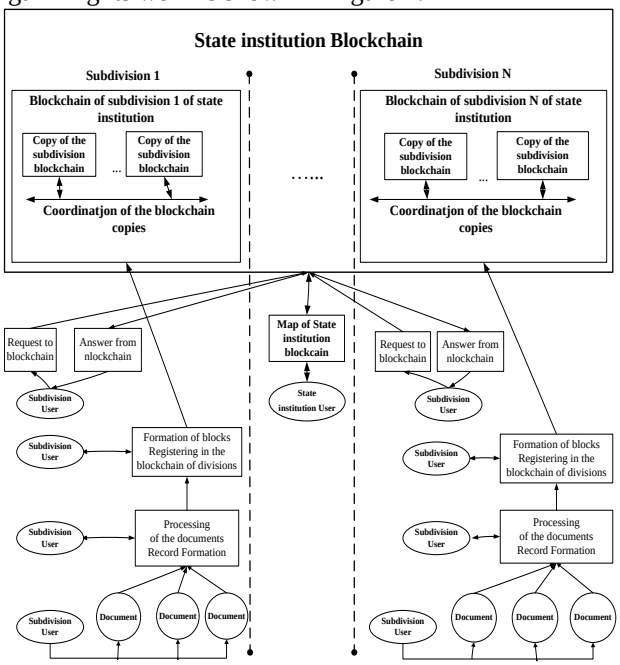

Fig. 1. Schema of the system based on the blockchain technology

Here, for writing of blocks, it is not necessary to solve the problem of finding a number with a given hash value. However, the hash of the block is also calculated and the hash of the created chain is generated. Since all users of the same subdivision (server) form a common queue, there will be no conflicts between them. Due to the specifics of the work of most state institutions, groups of users of each subdivision work with information of this subdivision, so other subdivisions create their own registers.

Here it is proposed for each subdivision of a public institution to create its own distributed registry. This solution allows to split data from different subdivisions, which often meets security requirements, reduce the amount of data in each registry, reduce the number of users with the right to enter data into the registry. For the access the registry uses known methods of control.

Consolidation of private registers into a general register of state institution is possible using a data allocation card. It is also possible to create a structure of a general registry in accordance with the structure of a state institution, for example, a network or hierarchical structure.

The number of copies of the registers of each subdivision is determined in accordance with the requirements for functional safety.

The structure of records and blocks in the registry should contain the creation date, the author and subdivision identifier.

The users of the registry of state institution and registries of divisions can be employees of state institution or private individuals (not employees). Employees are divided into those who are given the right to enter data in the registry and those who can only read the registry data. The private individuals are access rights only for reading data with restrictions on the amount and composition of data.

IV. MATHEMATICAL MODELING OF THE REGISTRY

When blocks are formed for entry in the registry, each block contains several documents. Thus, the intensity of the flow of blocks is less than the intensity of the flow of documents. This is the effect of information absorption in the formation of blocks.

A mathematical models of the block formation process are given in [2, 4, 6].

Below we consider the individual elements of the model with absorption [2].

System description

In general, we consider that the node for created of blocks receives N types document streams ($\infty > N \geq 1$), from which M types of blocks are created. The absorption algorithm is given by a matrix $\mathbf{M} = \|m_{ij}\|$, where $m_{ij} \geq 0$ the number of documents of the flow j, which are part of a block of the type i ($i = 1,2,...,M; j = 1,2,...,N$), the i-th row of the matrix is represented as a vector $\mathbf{m}_i = (m_{i1}, m_{i2}, ..., m_{iN})$.

Processing of a document consists in its inclusion in a block. Processing blocks are maintenance in the processing node (server of register) as a single (unified) message. Two types of queues can be formed in a node: a queue of documents and a queue of blocks waiting for service.

We investigate the case when the number of places for waiting for documents is limited. The number of places to wait for documents of each thread that are part of a block of type i is set by the vector $\mathbf{h}_i = (h_{i1}, h_{i2}, ..., h_{iN})$, where $\infty > h_{in} \geq m_{in}$ is the number of places to wait for documents of the flow number n.

As a model of the node, investigate the queueing system of a single service device, the input of which receives

N independent Poisson streams of documents, is studied. The rate of the flow of documents *i* – λ_i ($\infty > \lambda_i \geq 0$, *i* = 1, 2,…, *N*). The service device produces the formation and maintenance of blocks according to predefined rules.

Duration of processing of block of the type *i* (*i* = 1, 2, ..., *M*) is a random variable β_i with a distribution function $B_i(t)$, with first and second moments: $0 < b_{1i} = \int_0^\infty t dB_i(t) < \infty$ and $0 < b_{2i} = \int_0^\infty t^2 dB_i(t) < \infty$.

Here, to simplify the formulas, it is assumed that the duration of processing a block does not depend on its composition.

To analyze the operation of such a system, it is necessary to determine the characteristics of flows of blocks and probabilities of the documents loss.

Analysis of the system

The state of the system at time *t* is given by a vector $\mathbf{g}_i(t) = (g_{i1}(t), g_{i2}(t), ..., g_{iN}(t))$, where $g_{in}(t)$ is the number of documents of the flow number *n* ($m_{in} \neq 0$), which are in the system, and not included in the block type *i* ($g_{in}(t) = 0$, if $m_{in} = 0$).

If we consider the system at the time of arrive of documents that make up a block of type *i*, the set of states forms a nested finite Markov chain with the number of states - K_i.

Denote $\overline{\lambda_i} = \sum_{\substack{n=1 \\ (m_{in} \neq 0)}}^{N} \lambda_n$, *i* = 1 , 2,..., *M*. The probability that an arrived document will be a message of flow *n*: $q_{in} = \lambda_n / \overline{\lambda_i}$, *i* = 1 , 2,..., *M*; *n* = 1, 2,..., *N*.

For this Markov chain, a matrix of transient probabilities is constructed: $\mathbf{P}_i = \left\| p_{\mathbf{g}_i, g_i^*} \right\|$, where $\mathbf{g}_i$ and $\mathbf{g}_i^*$ are vectors of the states of chain; $p_{\mathbf{g}_i, g_i^*}$ is the probability of transition from state $\mathbf{g}_i$ to state $\mathbf{g}_i^*$.

The following formulas are obtained to calculate the values of transition probabilities:

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2, ..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2, ..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2,..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = \sum_{r=1}^{N} q_{ir}$;

This finite Markov chain is ergodic, and there is a limiting matrix: $\mathbf{A}_i = \lim_{r \to \infty} (\mathbf{P}_i)^r = \left\| a_{imn} \right\|, (m, n = 1, 2, ..., K_i)$, where $a_{imn} = a_{in}$ for all *n* = 1, 2, ..., K_i [3]. Here a_{in} is the limiting probability of hitting of the system at the next step in the state number *n*.

Among the set of States of the system, we select a subset of those for which a block is formed upon arrive of some document. We denote this subset - $\mathbf{U}_i$. The vector of the state of this subset will be denoted - $\mathbf{g}_{i\mathbf{U}_i}$. Let the limiting probability of such a state be $a(\mathbf{g}_{i\mathbf{U}_i}) = a_{ik}$ (the state number is *k*, $k \in \mathbf{K}(\mathbf{U}_i)$, where $\mathbf{K}(\mathbf{U}_i)$ is the set of state numbers included in the subset) and for the formation of a block from this state, the receipt of an document of the flow *n* is required. Then, the probability of occurrence of a block from this state is equal $a_{ik} q_{in}$. In this case, the probability of formation of a block when the next document of the total flow of documents is arrived is calculated as: $z_i = \sum_{\substack{j=1 \\ (j \in \mathbf{K}(\mathbf{U}_i)}}^{N} a_{ij} q_{ij}$, *i* = 1, 2,..., *M*.

Among the set of states of the system, we distinguish a subset of those for which, when a document of the flow *j* arrives, it will be lost - $\mathbf{H}_{ij}$. This subset of is denoted - $\mathbf{K}(\mathbf{H}_{ij})$. The probability of losing of document of the flow *j* of the block type *i*: $P_{ij} = \sum_{\substack{k=1 \\ k \in \mathbf{K}(\mathbf{H}_{ij})}}^{K_i} a_{ik} q_{ij}$, *i* = 1, 2,..., *M*; *j* = 1, 2,..., *N*.

Here is the limiting probability of getting the system to the state number *k*.

The probability of loss of any document constituting a block of type *i*: $P_i = \sum_{\substack{j=1 \\ m_{ij} \neq 0}}^{N} P_{ij}$, *i* = 1, 2,..., *M*.

After carrying out calculations on the constructed mathematical model, we obtain a set of characteristics of the of document flows that make up blocks of different types: sets of matrices: { $\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_M$ } and { $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_M$ }; a set of parameters of blocks flows: { $\Lambda_1, \Lambda_2, ..., \Lambda_M$ }; a set of values of the probability of loss of documents: { P_{ij} }, { P_i }, where (*i* = 1, 2, ..., *M*; *j* = 1, 2,.., *N*).

On practice, there are quite often cases when a block includes one document or when a block consists of a group of documents of only one flow. Applying the developed model for these cases, we obtain, for example: *M* = 1, *N* = 2, *h* = (1,1). The matrix of transient probabilities has the form:

$$\mathbf{P}_1 = \begin{pmatrix} 0 & q_{12} & q_{11} \\ q_{11} & q_{12} & 0 \\ q_{12} & 0 & q_{11} \end{pmatrix}.$$ The row of limiting matrix:

$a_{11} = (1 - q_{11})(1 - q_{12}) / (1 - q_{11} q_{12})$;

$a_{12} = q_{12}(1 - q_{11}) / (1 - q_{11} q_{12})$; $a_{13} = q_{11}(1 - q_{12}) / (1 - q_{11} q_{12})$.

The rate of the flow of blocks:

$$\Lambda_1 = \frac{\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)}{\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2}.$$

The probability of losing documents of the first and second flows: $P_{11} = a_{13}$, $P_{12} = a_{12}$.

Blocks of different types form a common queue, and are served one at a time in the order of receipt. The rate of the total flow of blocks:

$$\Lambda^* = \sum_{i=1}^{M} \Lambda_i \cdot$$

This system with the total flow of blocks can be considered as a queueing system type **M**/**G**/**1**/∞ [6].

The obtained results allow to calculate the loss probabilities of the documents, the parameters the flows of blocks and to explore the work of processing node.

V. CONCLUSION

The results can be considered as a method of preliminary design of the distributed registri of state institutions.

Mathematical model can be used to optimize the parameters of the registry.

REFERENCES

[1] Беларев И.А., Обаева А.С. О ра пределенном реестре и возможности его применения // Финансы: теория и практика. 2017. № 2(21). С 94–99. [Belarev I.A., Obaeva A.S. Distributed Ledger and its potential application. // Finance // Theory and practice. 2017. №2(21). P. 94–99].

[2] Dalinger J. M. Analysis of data flows in the systems with absorption of messages // Informatics and control systems: Publishing house of the Amur state University. 2012. №3 (33). P. 25–34.

[3] Kemeny J., Snell J. A finite Markov chains. Princeton; New York, 1967.

[4] Миролюбов А.Л., Саксонов Е.А. Система с комплексированием сообщений // Современные информационные и компьютерные технологии: Сб. науч. ст.: В 2 ч. / Гродненныский гос. университет им. Я. Купалы. Гродно, 2009. Ч. 2. С. 128–131.. [Mirolubov A.L., Saksonov E. A. The System with integration of the messages // The Modern information and computer technology: collection of scientific articles in 2 parts / The Ministry of education of the Republic of Belarus, Grodno State University Yankee Kupala. Grodno, 2009. Part 2. P. 128–131.

[5] Tanenbaum E., van Steen M. Distributed systems. Principles and paradigms. Upper Saddle River: Pearson Prentice Hall, 2007.

[6] Вишневский В.М. Теоретические основы построения компьютерных систем. М.: Техносфера, 2003. [Vishnevsky V. M. Theoretical bases of computer networks design. Moscow: Technosphere, 2003].