# Построение модели функционирования распределенной информационной системы на основе блокчейн-технологии

## Building a model for functioning a distributed information system based on blockchain technology

**Саенко / Saenko I.**
Игорь Борисович
(ibsaen@mail.ru)
доктор технических наук, профессор.
ФГКВОУ ВО «Военная академия связи
имени Маршала Советского Союза С. М. Будённого»
МО РФ (ВАС им. С. М. Буденного),
профессор кафедры автоматизированных систем
специального назначения (АС СН).
г. Санкт-Петербург

**Фабияновский / Fabiyanovski I.**
Игорь Николаевич
(fabik-spb@yandex.ru)
ВАС им. С. М. Буденного,
адъюнкт кафедры АС СН.
г. Санкт-Петербург

**Николаев / Nikolaev V.**
Владимир Викторович
(fortune-srus@yandex.ru)
ВАС им. С. М. Буденного,
адъюнкт кафедры АС СИ.
г. Санкт-Петербург

**Ясинский / Yasinski S.**
Сергей Александрович
(yasinsky777@mail.ru)
доктор технических наук, доцент.
ЗАО «Институт телекоммуникаций»,
ведущий специалист.
г. Санкт-Петербург

Ключевые слова: распределённая информационная система – distributed information system; информационный ресурс – information resource; блокчейн-технология – blockchain technology; распределенный реестр – distributed registry.

Рассматривается модель функционирования распределенной информационной системы на основе блокчейн-технологии. Модель предназначена для оценки доступности, целостности и эффективности обмена информационными ресурсами в распределенной информационной системе с произвольной топологией. Проведенные эксперименты подтверждают адекватность разработанной модели и показывают, что распределенные информационные системы, основанные на блокчейн-технологии, обладают более высокой эффективностью информационного обмена.

A model of functioning of a distributed information system based on blockchain technology is considered. The model is designed to evaluate the availability, integrity and efficiency of exchange of information resources in a distributed information system with an arbitrary topology. The experiments carried out confirm the adequacy of the developed model and show that distributed information systems based on blockchain technology have a higher efficiency of information exchange.

### Введение

В статье рассматривается модель функционирования распределенной информационной системы (РИС), построенной на основе блокчейн-технологии, которая предполагает обработку информационных ресурсов (ИР) в информационно-телекоммуникационной сети (ИТКС) с произвольной топологией. Одним из важных показателей скорости работы такой системы является среднее время реакции на запрос пользователя, а также время реакции системы в целом. Поэтому ставится задача разработки математической модели, позволяющей определить указанные характеристики для таких систем.

### Математическая модель функционирования распределенной информационной системы на основе блокчейн-технологии

Поскольку реальная РИС может иметь сложную топологию (структуру связей между подсетями и узлами), а интенсивность формирования запросов разными пользователями в общем случае описывается различными законами распределения, то для построения обобщенной математической модели, позволяющей оценить основные характеристики системы, необходимо ввести следующие упрощающие допущения.

Рассмотрим модель системы массового обслуживания (СМО) без отказов, без приоритетов, с дисциплиной диспетчеризации FIFO. Допустим, что отсутствует сегментация сообщений, а столкновения в канале не учитываются, поскольку используемые для передачи данных протоколы сводят вероятность возникновения коллизий к минимуму. Пусть в систему поступает пуассоновский поток заявок, длительность



4 2020.indd 73 24.12.2020 16:49:03

интервала между которыми распределена по экспоненциальному закону с плотностью $f(t) = \lambda e^{-\lambda t}$, где $\lambda$ – интенсивность потока. Использование такого входящего потока заявок обусловлено тем, что он служит для моделирования различных реальных процессов, в частности, для формирования запросов пользователями в информационных системах и осуществления электронного документа оборота с достаточной для большинства процессов точностью. При этом упрощается решение задач СМО, связанных с расчетом их эффективности [1–3]. Поступившая заявка направляется на свободный прибор. Если прибор занят, то поступившая заявка направляется в очередь неограниченной длины. Введенные допущения значительно упрощают модель, однако несущественно изменяют принцип обработки данных в РИС и отвечают общему порядку обслуживания пользовательских запросов. Целью моделирования является определение времени реакции на запрос конкретного пользователя, а также времени реакции всей системы.

Пусть $n$ – суммарное количество всех узлов $U_i$, $i = \overline{1,n}$. Введем в рассмотрение следующие формальные обозначения:

$A_i$ – $i$-ый пользователь, $i = \overline{1,n}$;

$F_j$ – $j$-ый блок узла, $j = \overline{1,m}$, где $m$ – суммарное количество блоков;

$\lambda_i$ – интенсивность формирования запросов из узла $U_i$, $i = \overline{1,n}$;

$Q_l$ – $l$-ый запрос, $l = \overline{1,q}$, где $q$ – количество запросов;

$VF_j$ – объем ИР на блоке узла $F_j, j = \overline{1,m}$;

$VV_i$ – скорость чтения информации в узле $U_i, j = \overline{1,n}$;

$VD_i$ – скорость записи в оперативную память узла $U_i$, $i = \overline{1,n}$;

$O$ – скорость передачи данных по каналу связи;

$O_o$ – постоянная задержка при передаче данных по каналу связи;

$\alpha_o$ – постоянная задержка при обработке в узле (связана с периодом опроса входной очереди в буфере электронных вычислительных машин);

$PU_i$ – производительность процессора узла $U_i$, $i = \overline{1,n}$;

$f_{AiQl}$ – вероятность формирования пользователем $A_i$ запроса $Q_l$, $i = \overline{1,n}, l = \overline{1,q}$;

$b_{QlFj}$ – объем считываемой информации при выполнении запроса $Q_l$ к фрагменту запроса $F_j, j = \overline{1,n}, l = \overline{1,q}$;

$b_{rQlFj}$ – количество процессорных операций, необходимых для выполнения запроса $Q_l$ к блоку $F_j, j = \overline{1,n}, l = \overline{1,q}$.

Модель описанной информационной системы можно рассматривать в виде разомкнутой СМО. Для моделирования канала передачи данных, коммутаторов и узлов используются обслуживающие приборы (в терминах СМО). Кроме того, в состав СМО входит буферная память каждого сетевого оборудования, которая хранит поступающие сообщения, а также буферная память каждого узла. Необходимо отметить, что рассматриваемая концептуальная модель предполагает следующее. Узлы сети, подсистема управления доступом контента (ПУДК), подсистема метаданных (ПМ), входящие в состав сети распределенной структуры, в совокупности образуют множество узлов $U_i$, $i = \overline{1,n}$. ПУДК самостоятельно не формируют пользовательские запросы, что в свою очередь достигается заданием вероятности формирования собственных исходящих запросов равной нулю.

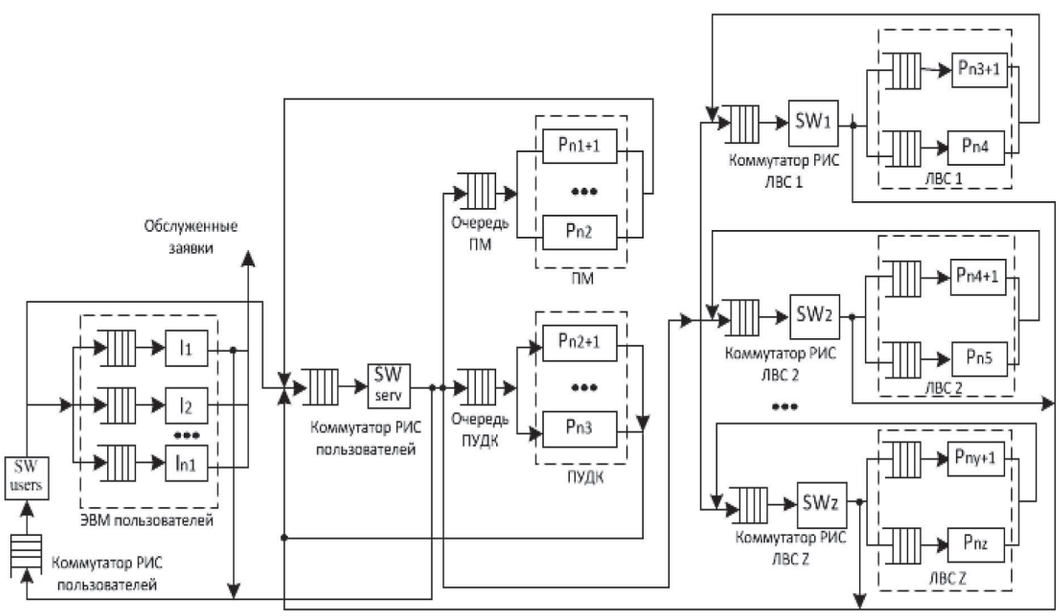


*Рис. 1. Концептуальная модель СМО*



4-2020.indd 74 24.12.2020 16:49:05

Описанную модель распределенной системы можно представить в виде концептуальной модели СМО, приведенной на рис. 1. На этом рисунке применяются следующие обозначения: $n_i$ – количество электронных вычислительных машин (ЭВМ) пользователей; $(n_2 - n_1)$ – количество СМ; $(n_3 - n_2)$ – количество систем управления доставкой (СУД); $(n_6 - n_3)$ – количество ЭВМ в подсетях облака; $I_i$, …., $I_{n1}$ – генераторы транзактов, имитирующих SQL-запросы; $P_{n1+1}$,…$P_{n2}$ – приборы, моделирующие СМ; $P_{n2+1}$,…$P_{n3}$ – приборы, моделирующие СУД; $P_{n3+1}$, ,,,, $P_{nz}$– приборы, моделирующие ЭВМ в подсетях облака; $SW_{user}$ – прибор, моделирующий коммутатор ЛВС пользователей; $SW_{serv}$ – прибор, моделирующий коммутатор облака; $SW_1$…, $SW_z$ – приборы, моделирующие коммутаторы подсетей.

Определение времени реакции на запрос конкретного $A_s$-го пользователя подразумевает применение декомпозиционного подхода, заключающегося в том, что для $A_s$-го пользователя определяются возможные схемы выполнения запросов, соответственно вычисляются составляющие временных задержек, после чего рассчитывается время реакции по всем возможным схемам. Для каждого обслуживающего прибора определяются временные задержки, при вычислении которых учитываются составляющие интегрированного потока запросов к этому прибору и процедур обслуживания.

Для $U_s$-го узла выделены две группы потока заявок, определяющие возможные схемы выполнения запросов:

– запросы, инициированные в узле $U_s$, выполнение которых подразумевает обращение к бокам данных находящимся в других узлах, а окончательная обработка полученной информации производится в $U_s$-м узле. Запросы данной группы формируются пользователями сети с регулярной структурой;

– запросы, инициированные пользователями других узлов, а в узле $U_s$ содержится блок данных, необходимых для их выполнения. Такие запросы обрабатываются узлами подсетей, входящими в состав РИС, а также СУД.

Режим функционирования системы и характеристики пользовательских запросов определяют процедуру обслуживания заявок. При выполнении запросов первой группы сначала в $U_s$-м узле осуществляется получение данных, которые были предварительно считаны в другом узле, а затем их обработка. Выполнение запросов второй группы, инициированных пользователями других узлов информационной системы, включает чтение и обработку данных, размещенных в узле $U_s$. Для передачи запросов от $A_s$-го пользователя к другим узлам и для передачи из $U_s$-го узла данных в другие узлы требуются ресурсы канала.

Обозначим вероятность формирования пользователем $A_s$ запроса к узлу $U_z$ как $PS_{AsUz}$, $s = \overline{1,n}$, $z = \overline{1,n}$. Тогда время $TA_s$ реакции на запросы $A_s$-го пользователя можно определить по следующей формуле:

$$TA_s = \sum_{\substack{z=1 \\ z \neq s}}^{n} PS_{A_s u_z} TA_{A_s u_z}, s = \overline{1,n}.$$

Согласно [4–8], время $TA$ реакции всей системы вычисляется по формуле

$$TA_s = \sum_{s=1}^{n} \lambda_s TA_s / \sum_{s=1}^{n} \lambda_s.$$

Время $t_0$ ожидания обслуживания заявок для разомкнутых СМО с γ входными потоками и безприоритетной дисциплиной диспетчеризации FIFO определяется следующим образом:

$$TA_s = \sum_{i=1}^{n} \lambda_i x_i^{-2} / 2\left(1 - \sum_{i=1}^{\gamma} \lambda_i \omega_i\right),$$

где $\lambda_i$ – интенсивность $i$-го потока заявок; $x_i^{-2}$ – второй момент обслуживания заявок; $\omega_i$ – среднее время обслуживания заявок.

Чтобы вычислить время реакции на запросы $A_s$-го пользователя, основываясь на имеющихся исходных данных, нужно определить:

– вероятность формирования запросов $A_s$-м пользователем к $F_j$-му блоку данных:

$$PI_{A_s F_j} = \sum_{i=1}^{q} f_{A_s Q_i} \delta(b_i F_j), s = \overline{1,n}; j = \overline{1,m},$$

где $\delta(b_i F_j)$ – символ Кронекера;

– вероятность формирования запросов $A_s$-м пользователем к $U_z$-му узлу:

$$PS_{A_s U_z} = \sum_{i=1}^{m} PI_{A_s F_j} x_{F_j U_z}, s = \overline{1,n}; z = \overline{1,n},$$

где $x_{F_j U_z}$ – элемент матрицы распределения блоков по узлам РИС, который равен 1, если $F_j$-й фрагмент размещен в $U_z$-м узле, и 0 – в противном случае;

– интенсивность входного потока для первой группы заявок:

$$\lambda_{1s} = \lambda_s \sum_{\substack{z=1 \\ z \neq s}}^{n} PS_{A_s U_z}, s = \overline{1,n};$$

– интенсивность входного потока для второй группы заявок:

$$\lambda_{2s} = \lambda_s \sum_{\substack{z=1 \\ z \neq s}}^{n} \lambda_z PS_{A_s U_z}, s = \overline{1,n};$$

Чтобы определить время обслуживания запросов первой группы, необходимо вычислить следующие показатели:

– средний объем информации, передаваемой в $U_s$-й узел:

$$VS_{1s} = \sum_{i=1}^{q} f_{A_s Q_i} \sum_{j=1}^{m} b_{Q_i F_j} \sum_{\substack{z=1 \\ z \neq s}}^{n} x_{F_j U_z}, s = \overline{1,n};$$



4-2020.indd 75 24.12.2020 16:49:06

– среднее время передачи информации в $U_s$-й узел:

$$WS_{1s}=VS_{1s}/O + O_o + VS_{1s}/ VD_s,\ s = \overline{1,n};$$

– среднее количество процессорных операций, связанных с выполнением процедур управления:

$$PV_{1s} = \sum_{i=1}^{q} f_{A_sQ_i} \sum_{j=1}^{m} br_{Q_iF_j} \sum_{\substack{z=1 \\ z\neq s}}^{n} x_{F_jU_z}, s=\overline{1,n};$$

– среднее время выполнения процессорных операций:

$$WP_{1s} = PV_{1s} / PU + \alpha_0,\ s=\overline{1,n};$$

На основе введенных характеристик можно вычислить среднее время обслуживания запросов первой группы в узле $U_s$:

$$W_{1s} = WS_{1s} + WP_{1s},\ s=\overline{1,n};$$

Аналогичные характеристики вводятся для запросов второй группы:

– средний объем считываемой информации:

$$VS_{2s} = \sum_{\substack{i=1 \\ i\neq s}}^{n} \lambda_i / \sum_{\substack{u=1 \\ u\neq s}}^{n} \lambda_u \sum_{i=1}^{q} f_{A_sQ_i} \sum_{j=1}^{m} \delta\left(b_{Q_iF_j}\right) \cdot V_{F_jU_z}, s=\overline{1,n};$$

– среднее время чтения и передачи информации:

$$WS_{2s} = \frac{VS_{2s}}{O} + O_o + \frac{VS_{2s}}{VV_s}, s=\overline{1,n};$$

– среднее количество процессорных операций, связанных с выполнением процедур управления:

$$PV_{2s} = \sum_{i=1}^{q} f_{A_sQ_i} \sum_{j=1}^{m} br_{Q_iF_j} x_{F_jU_z}, s=\overline{1,n};$$

– среднее время выполнения процессорных операций:

$$WP_{2s} = PV_{2s} / P_{us} + \alpha_0,\ s=\overline{1,n};$$

Тогда среднее время обслуживания запросов второй группы в $U_s$-м узле вычисляется как

$$W_{2s} = WS_{2s} + WP_{2s},\ s=\overline{1,n};$$

Время $W_{SUD}^{III}$ формирования подсистемой управления доставкой окончательного набора данных из результатов обработки запроса узлами определяется на основе следующих соотношений:

– среднего объема полученной информации:

$$VS_s = \sum_{\substack{i=1 \\ i\neq s}}^{n} \lambda_i / \sum_{\substack{u=1 \\ u\neq s}}^{n} \lambda_u \sum_{i=1}^{q} f_{A_sQ_i} \sum_{j=1}^{m} \delta\left(b_{Q_iF_j}\right) \cdot V_{F_j} \sum_{\substack{z=1 \\ z\neq s}}^{n} x_{F_jU_z}, s=\overline{1,n};$$

– среднего времени чтения информации:

$$WSs = VSs / O + O_o + VS_s / VV_s,\ s=\overline{1,n};$$

– среднего количества процессорных операций, связанных с выполнением процедур управления:

$$PV_s = \sum_{i=1}^{q} f_{A_sQ_i} \sum_{j=1}^{m} br_{Q_iF_j} x_{F_jU_z}, s=\overline{1,n};$$

– среднего времени выполнения процессорных операций:

$$WP_s = PV_s / P_{us} + \alpha_0,\ s=\overline{1,n};$$

где $s$ – номер узла, соответствующего ПУДК.

Тогда среднее время формирования подсистемой управления доставкой окончательного набора данных вычисляется как $W_{SUD}^{III} = WS_s + WP_s$.

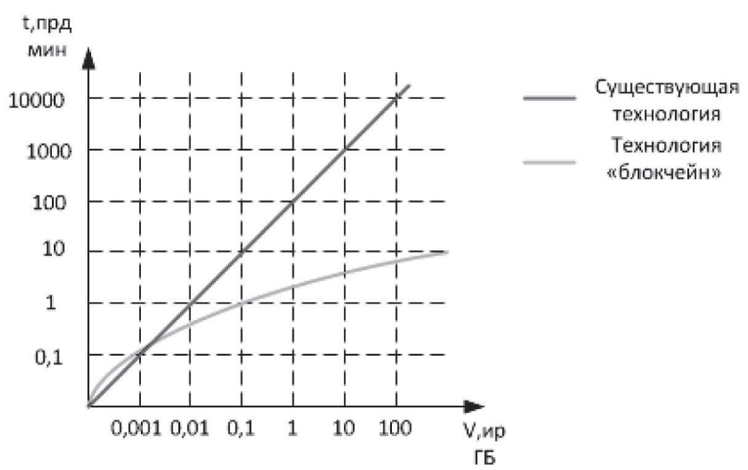


*Рис. 2. Графики зависимостей объема ИР от времени передачи*



4-2020.indd 76 24.12.2020 16:49:08

Время ожидания обслуживания $U_s$-м узлом запросов $k$-й группы определяется по формуле:

$$TO^{I}_{SUD} = TO^{II}_{SUD} = TO^{III}_{SUD} = TO_{SM} = TO_{1S} = TO_{2S} = TO_S =$$

$$= \frac{\lambda_{1s} x_{1s}^{-2} + \lambda_{2s} x_{2s}^{-2}}{2\left(1-\left(\rho_{1s}+\rho_{2s}\right)\right)} = \left(\sum_{k=1}^{2} \lambda_{ks} x_{ks}^{-2}\right) \Big/ \left[2\left(1-\sum_{k=1}^{2}\rho_{ks}\right)\right], s = \overline{1,n}.$$

Ввиду того, что размер запроса, как правило, не превышает размер кадра Ethernet, среднее время передачи запроса не зависит от пользователя, т.е. $W^{sv}_{users} = W^{sv}_{cloud} = W^{sv}_{lan} = WK_1$, где $WK_1$ задается в качестве исходных данных и может быть определено на основе анализа статистических данных.

Время обслуживания каналом запросов второй группы вычисляется следующим образом:

$$WS2s = VS2s / O + O_o,\ s = \overline{1,n}.$$

Время обработки служебных сообщений не зависит от номера пользователя и принимается равным $W_{serv}$, т.е. $W^{I}_{SUD} = W_{SM} = W^{II}_{SUD} = W_{serv}$, т.е. $W_{serv}$ где $W_{serv}$ задается в качестве исходных данных и определяется характеристиками используемого оборудования.

Загрузка канала вычисляется исходя из интенсивности запросов $\lambda_{ks}$ по следующей формуле:

$$\rho_{ks} = \lambda_{ks} W_{ks},\ k = 1,2,\quad s = \overline{1,n}.$$

Загрузка прибора может быть вычислена исходя из интенсивности формирования запросов $\lambda_{ks}$ и времени их обслуживания $W_{ks}$ следующим образом: $\rho_{ks} = \lambda_{ks} W_{ks}$, $k = 1,2;\ s = \overline{1,n}$. При экспоненциальном законе распределения вторые моменты времени обслуживания равны $x_{ks}^{-2} = 2W_{ks}^{2}$, а при постоянном времени обслуживания $x_{ks}^{-2} = 2W_{ks}^{2},\ k = 1,2;\ s = \overline{1,n}$.

Выше было отмечено, что время передачи запросов из одного узла в другой принимается постоянным. Длительность передачи данных, возвращаемых запросом, распределено по экспоненциальному закону. Поэтому вторые моменты времени обслуживания заявок первой и второй групп, соответственно, вычисляются по следующим формулам:

$$\overline{XK}_{ts}^{2} = WK_{1s}^{2} \cdot \overline{XK}_{2s}^{2} = 2WK_{2s}^{2},\ s = \overline{1,n}.$$

В модели не учитываются столкновения в канале, так как используемые протоколы сводят вероятность возникновения коллизий к минимуму, а ресурсы канала выделяются по правилу FIFO без приоритетов. Тогда среднее время ожидания обслуживания заявок $k$-й группы не зависит от номера s пользователя и определяется следующим выражением:

$$TOK_{1s} = TOK_{2s} = TOK^{SW}_{ris} = TOK =$$

$$= \left[\sum_{s=1}^{2}\left(\lambda_{1s}\overline{XK}_{1s}^{2} + \lambda_{2s}\overline{XK}_{2s}^{2}\right)\right] \Big/ \left[2\left(1-\sum_{k=1}^{2}\left(\rho_{1s}+\rho_{2s}\right)\right)\right], s = \overline{1,n}.$$

## Экспериментальные результаты применения блокчейн- технологии в распределенной информационной системе

Результаты моделирования РИС, в которых определены все составляющие для расчета времени $TA_{A_sU_z}$ выполнения запроса $A_s$-го пользователя к $U_z$-му узлу ($s = \overline{1,n},\ z = \overline{1,n},\ s \neq z$) и найдены интегральные характеристики $TA_s$ (время реакции на запросы пользователя $A_s$) и $TA$ (время реакции системы), представлены

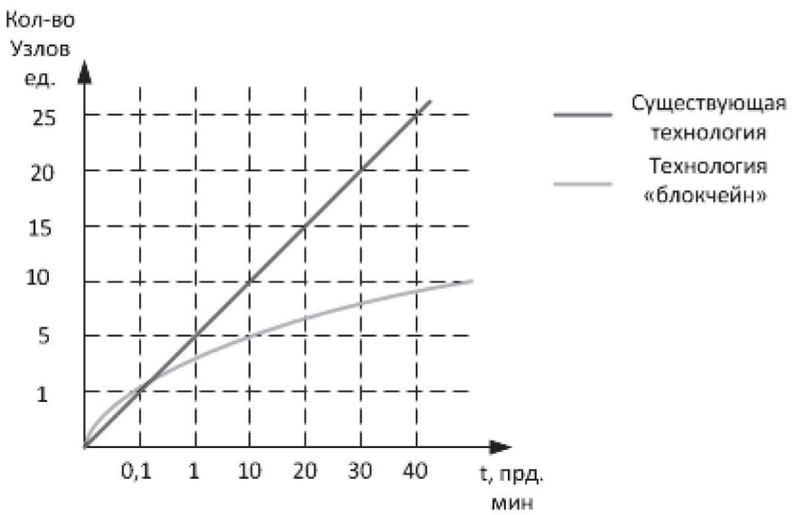


*Рис. 3. Графики зависимостей количества узлов от времени на передачу*



4-2020.indd 77 24.12.2020 16:49:10

на рис. 2 и рис. 3 в виде графиков, соответственно для зависимостей объема ИР от времени передачи и зависимостей количества узлов от времени на передачу.

В экспериментах, результаты которых представлены на рис. 2, предполагалось, что пропускная способность каналов, обеспечивающих передачу данных между узлами, ограничивается величиной 100 Мбит/сек. В экспериментах, результаты которых представлены на рис. 3, предполагалось, что объем ИР был ограничен величиной 10 Мбайт, а пропускная способность каналов ограничивается величиной 100 Мбит/сек.

Представленная модель отличается от уже известных тем, что учитывает специфику распределенной архитектуры. На ее основе можно производить моделирование распределенных систем различных областей приложений (системы электронного документооборота, системы управления технологическими процессами и т.д.). Данное утверждение справедливо, во-первых, потому что перечисленные информационные системы относятся к классу транзакционных систем, так как работают с малыми по размеру транзакциями, идущими большим потоком. Все осуществляемые операции в модели выполняются в одной транзакции, а их результат записывается в распределенный реестр. Во-вторых, в перечисленных транзакционных системах источниками заявок являются пользователи. Заявки обрабатываются в порядке поступления (дисциплина обслуживания FIFO), а результаты отправляются инициаторам. Интенсивность потока определяется количеством заявок, генерируемых за единицу времени. В таком виде она не зависит от конкретной предметной области.

Работа выполнена при поддержке РФФИ (проект 18-07-01369).

## Заключение

В статье рассмотрена модель функционирования РИС на основе блок-чейн технологии, предполагающая размещение блоков по узлам РИС с произвольной топологией. Математическая модель, описывающая такую систему как разомкнутую СМО, учитывает в качестве показателей своего функционирования среднее время реакции на запрос пользователя, а также время реакции системы в целом.

Дальнейшее направление работ связывается с разработкой методики управления информационным обменом в РИС на основе блокчейн-технологии в целях достижения требуемых значений оперативности и достоверности функционирования системы.

## Литература

1. Горобец, В. В. Модель архитектуры информационных систем для развертывания в рамках облачной среды / В.В. Горобец // Вестник Тверского государственного университета. Серия Прикладная математика. – 2012. – № 31. – С. 115–128.

2. Горобец, В. В. Облачная модель транзакционной системы / В.В. Горобец // Вестник компьютерных и информационных технологий. – 2013. – № 4. – С. 19–24.

3. Gorobets, V. V. The model of on-line transaction processing system in the cloud environment / V.V. Gorobets // Академические фундаментальные исследования молодых ученых России и Германии в условиях глобального мира и новой культуры научных публикаций: материалы Международной молодежной конференции, Новочеркасск, 4–5 октября 2012. – С. 157–159.

4. Горобец, В. В. Альтернативные варианты построения архитектуры транзакционных систем / В.В. Горобец // Методы и алгоритмы прикладной математики в технике, медицине и экономике: материалы XI Международной научно-практической конференции, Новочеркасск, 28 февраля 2011. – С. 52–57.

5. Воробьев, С. П. Обзор математических моделей для построения транзакционных систем на базе различных архитектурных решений / С.П. Воробьев, В.В. Горобец // Моделирование. Теория, методы и средства: материалы XI Международной научно-практической конференции, Новочеркасск, 31 марта 2011. – С. 146–153.

6. Стоунбрейкер, М. Один размер пригоден для всех: идея, время которой пришло и ушло [Электронный ресурс] / М. Стоунбрейкер, У. Кетинтемел // Центр информационных технологий. – Режим доступа: http://citforum.ru/database/articles/one_size_fits_all/ [дата обращения: 24.11.2012], свободный. – Загл. с экрана.

7. Брювер, Л. Способы повышения производительности, или когда скорость решает все [Электронный ресурс] / Л. Брювер; пер. с англ. И. Луканин // PHP Magazin. – Режим доступа: http://php-mag.net/itr/online_artikel/psecom.id.502.node.id. 114.html, свободный. – Загл. с экрана (дата обращения: 28.08.2020).

8. Сиротинский, В. В. Проблема производительности корпоративных хранилищ данных / В.В. Сиротинский, Д.В. Семенов // Вестник Московского авиационного института. – 2010. – Т. 17, № 1. – С. 154–159.



4 2020.indd 78 24.12.2020 16:49:10