

И.Д. Котилевец, Я.А. Миноцкий

СРАВНЕНИЕ СКОРОСТИ ОБРАБОТКИ ТРАНЗАКЦИЙ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ НА ОСНОВЕ РАСПРЕДЕЛЕННОГО РЕЕСТРА И НАПРАВЛЕННОГО АЦИКЛИЧЕСКОГО ГРАФА

Аннотация. Приводится сравнительный анализ скорости обработки транзакций в технологиях распределенного реестра и направленного ациклического графа. В расчетах величины среднего времени обработки транзакций применяется теория массового обслуживания.

Ключевые слова: ациклический направленный граф, распределенный реестр, транзакции, система массового обслуживания, среднее время обработки транзакции, пропускная способность.

I.D. Kotilevets, Ya.A. Minotsky

COMPARISON OF TRANSACTION PROCESSING SPEED IN DISTRIBUTED LEDGER AND DIRECTED ACYCLIC GRAPH

Abstract. This paper provides a comparative analysis of the transaction processing speed in distributed registry technologies and directed acyclic graph. The queuing theory is used to calculate the value of the average transaction processing time.

Keywords: directed acyclic graph, distributed registry technologies, transactions, queuing system, average transaction processing time, throughput.

Введение

В настоящее время наблюдается повышенный интерес к технологии распределенных реестров. Предполагается, что данная технология позволяет устранить ряд ограничений, присущих применяемым сегодня методам хранения данных, например, проблемы безопасности хранения и передачи данных и масштабируемости систем.

Технология распределенных реестров – это подход к обмену и хранению информации, при котором [1; 2]:

- каждый участник может обладать полноценной копией хранимых данных;
- синхронизация копий реестра происходит на основе протокола достижения распределенного консенсуса, то есть соглашения среди участников на добавление новой информации;
- каждый участник взаимодействия может иметь доступ к истории транзакций.

В классической технологии построения распределенного реестра у пользователей нет прямого доступа к самому реестру. Когда пользователь хочет добавить транзакцию в реестр, ему приходится отдавать поручение производителю верифицированных записей реестра (записи назовем блоками, а производителей – блок-продюсерами или майнерами). Майнеры решают, какую транзакцию добавить в следующий блок. У них есть эксклюзивный доступ к блокам и право решать, чью транзакцию принять для добавления в реестр с учетом применяемого алгоритма консенсуса. Таким образом, майнеры являются посредниками между пользователями и распределенным реестром. Но также майнеры являются гарантом роста сети и поддержания безопасности за счет внутренних механизмов взаимодействия узлов распределенных реестров.

Котилевец Игорь Денисович

старший преподаватель кафедры цифровых технологий обработки данных, МИРЭА – Российский технологический университет, Москва. Сфера научных интересов: базы данных, теория графов, системы массового обслуживания, технология распределенных реестров, интернет вещей. Автор более 50 опубликованных научных работ.

Электронный адрес: ikotilevets@gmail.com

Миноцкий Ян Анатольевич

студент, МИРЭА – Российский технологический университет, Москва. Сфера научных интересов: кибербезопасность, технологии распределенных реестров, интеллектуальный анализ данных.

Электронный адрес: yan.minotskiy@mail.ru

Однако данная технология имеет ограничение, которое замедляет ее внедрение в различные сферы жизни, например промышленность. Классический распределенный реестр не решает проблему масштабируемости, так как происходит снижение пропускной способности и скорости транзакций при существенном росте объемов транзакций из-за необходимости добавления транзакций в единую цепочку блоков и невозможности разделить процесс обработки в несколько потоков [3–5].

Для решения данной проблемы необходимо изменить архитектуру классического реестра и отказаться от логики блоков и блок-продюсеров. Для выстраивания цепочки блоков необходимо соединение самих транзакций путем включения в каждую транзакцию хешей нескольких предыдущих транзакций. В результате получается структура, известная в математике как направленный ациклический граф (Directed Acyclic Graph – DAG).

В данной работе приводится сравнительный анализ параметров математических моделей классического распределенного реестра и системы на основе направленного ациклического графа.

Модель распределенного реестра

В качестве параметров сравнения классического распределенного реестра и системы на основе направленного ациклического графа выбраны: время обработки одной транзакции, пропускная способность, количество устройств в сети, загруженность системы. Данные параметры необходимо рассчитать на основе входных параметров: количество устройств в сети, интенсивность поступления транзакций, время обработки транзакции одним устройством. Для расчета распределенной сети, построенной на основе классического распределенного реестра, можно представить систему обработки транзакции и создания нового блока как систему массового обслуживания. Системы массового обслуживания (далее – СМО) – это системы, в которые в случайные моменты времени поступают заявки на обслуживание, при этом поступившие заявки обслуживаются с помощью имеющихся в распоряжении системы каналов обслуживания. Данную математическую модель можно использовать в моделировании системы распределенных реестров, так как в качестве заявок могут выступать транзакции, а в качестве системы каналов обслуживания – устройства обрабатывающие транзакции. Рассмотрим классический распределенный реестр как одноканальную СМО с неограниченной очередью. Пусть все устройства в сети имеют одинаковую скорость обработки транзакции μ (пропускную способность):

Сравнение скорости обработки транзакций в распределенных системах ...

$$\mu = \frac{1}{t_{об}}, \tag{1}$$

где $t_{об}$ – время обработки транзакции.

Данный параметр зависит от множества факторов: мощность устройства, загруженность устройства другими процессами, пропускная способность канала передачи, скорость передачи по каналу связи и др.

Новые транзакции поступают в систему с интервалом по заданному закону распределения с математическим ожиданием $M(x)$. Тогда интенсивность поступления заявок $\lambda = M(x)$.

СМО с неограниченной очередью может находиться в бесконечном числе состояний. На Рисунке 1 представлен граф состояний системы обработки транзакций (S_0 – система свободна; S_1 – система занята, обрабатывается транзакция, транзакций в очереди нет; S_2 – система занята обработкой транзакции, в очереди на обработку 1-я транзакция; S_3 – система занята обработкой транзакции, в очереди на обработку 2-я транзакция, и др.) [6].

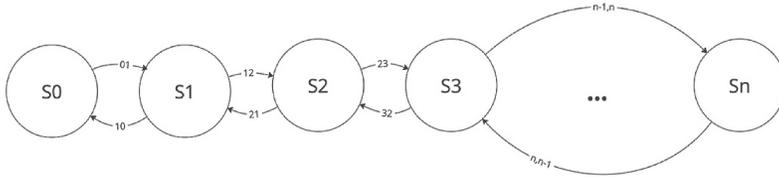


Рисунок 1. Граф состояний системы обработки транзакций

Загруженность системы обработки входящих транзакций можно определить как вероятность ненахождения системы в состоянии S_0 . Обозначим величину вероятности нахождения системы в состоянии S_0 как p_0 . Тогда будет справедливо равенство

$$\lambda_{01}p_0 = \lambda_{10}p_1, \tag{2}$$

где λ_{10} – интенсивность перехода системы из состояния S_0 ; λ_{01} – интенсивность перехода из состояния S_0 в состоянии S_1 .

Для p_1 справедливо выражение

$$\lambda_{12}p_1 = \lambda_{21}p_2. \tag{3}$$

Аналогично для p_k . В общем случае вероятность пребывания системы в состоянии определяется следующим соотношением:

$$\lambda_{k,k+1}p_k = \lambda_{k+1,k}p_{k+1}. \tag{4}$$

Тогда система уравнений примет вид

$$\begin{cases} \lambda_{01}p_0 = \lambda_{10}p_1, \\ \lambda_{12}p_1 = \lambda_{21}p_2, \\ \dots \\ \lambda_{k,k+1}p_k = \lambda_{k+1,k}p_{k+1}. \end{cases} \tag{5}$$

Для решения данной системы уравнений выразим все значения через p_0 . Например, для p_1 справедливо

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} p_0; \tag{6}$$

Для p_2

$$p_2 = \frac{\lambda_{12}}{\lambda_{21}} p_1 = \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} p_0. \quad (7)$$

Следует обратить внимание, что в числителе выражения (7) стоят интенсивности, идущие слева направо (см. Рисунок 1). Эти интенсивности одинаковы для всех переходов направо и равны λ , а переходы налево, которые представлены интенсивностями в знаменателе, равны μ [7]. Тогда справедливо выражение:

$$p_k = \left(\frac{\lambda}{\mu}\right)^k p_0. \quad (8)$$

Вероятность всех возможных n состояний

$$\sum_{i=0}^n p_i = 1, \quad (9)$$

то есть

$$p_0 + \frac{\lambda}{\mu} p_0 + \left(\frac{\lambda}{\mu}\right)^2 p_0 + \left(\frac{\lambda}{\mu}\right)^3 p_0 + \dots + \left(\frac{\lambda}{\mu}\right)^n p_0 = 1. \quad (10)$$

Отсюда p_0 можно найти по формуле

$$p_0 = \frac{1}{\sum_{i=0}^{n \rightarrow \infty} \left(\frac{\lambda}{\mu}\right)^i}. \quad (11)$$

В знаменателе получен ряд геометрической прогрессии

$$S_n = \frac{a(1-q^n)}{1-q} \quad (12)$$

Числовое значение выражения (12) (то есть сумма всех членов геометрической прогрессии) находится по формуле

$$S_n = \frac{a(1-q^n)}{1-q}, \quad (13)$$

при $n \rightarrow \infty$ и при $a = 1$.

Тогда выражение суммы ряда геометрической прогрессии

$$\lim_{n \rightarrow \infty} \frac{1-q^n}{1-q}. \quad (14)$$

Решение данного предела

$$\begin{cases} \frac{1}{1-q}, & |q| < 1, \\ \infty, & |q| \geq 1. \end{cases} \quad (15)$$

Так как интенсивность и пропускная способность не могут быть отрицательными числами, то знак модуля можно убрать. Таким образом, ряд сходится, когда выполняется условие

$$\frac{\lambda}{\mu} < 1. \quad (16)$$

При выполнении условия (16) система обработки транзакций, представляющая из себя СМО, находится в стационарном состоянии, при этом очередь не накапливается до бесконечности, и можно рассчитать среднее время обработки транзакций.

Сравнение скорости обработки транзакций в распределенных системах ...

Для удобства обозначим величину λ/μ как ρ . Чтобы найти среднее время обработки транзакций, сначала найдем среднее количество устройств, находящееся в системе обработки транзакций в случайный момент времени. Для одноканальной СМО с неограниченной очередью данное значение можно найти по формуле

$$L_{\text{сист}} = \frac{c}{1-c}. \quad (17)$$

По формуле Литтла найдем среднее время обработки транзакции:

$$T_{об} = \frac{L_{\text{сист}}}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{t_{об}}{1-\lambda t_{об}} = \left(\frac{1}{t_{об}} - \lambda \right)^{-1}. \quad (18)$$

Из выражения (18) можно сделать вывод, что при увеличении $t_{об}$ увеличивается $T_{об}$, и при увеличении λ также увеличивается $T_{об}$.

Стоит отметить, что данная закономерность прослеживается именно на интервале области допустимых значений параметров $t_{об}$ и λ :

$$T_{об} = \begin{cases} t_{об} \in [0; +\infty), \\ \lambda \in [0; +\infty), \\ t_{об}\lambda \in [0; 1). \end{cases} \quad (19)$$

На Рисунке 2 представлен график зависимости $T_{об}$ от $t_{об}$ и λ .

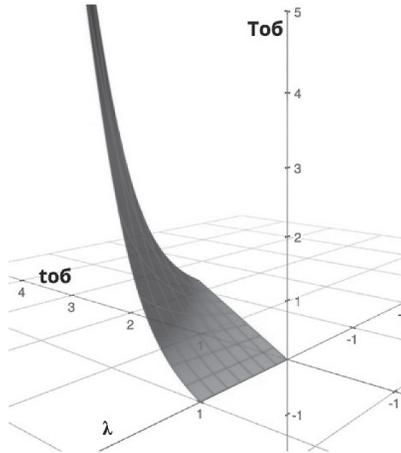


Рисунок 2. График зависимости времени обработки транзакции от $t_{об}$ и λ

Из (16) и Рисунка 2 можно сделать вывод, что для поддержания системы в стационарном состоянии необходимо ограничить число входящих транзакций, а значит, ограничить количество устройств в распределенном реестре (проявляется взаимная зависимость параметров n (количество устройств в сети) и λ). То есть $n \rightarrow \infty, \lambda \rightarrow \infty$, а значит, $T_{об} \rightarrow \infty$. Из этого следует, что горизонтальное масштабирование системы с помощью увеличения количества устройств приведет к увеличению времени обработки транзакций устройствами сети.

Проблему ограниченности потока входящих заявок в распределенную систему может решить DAG – ациклический направленный граф.

Транзакции, исходящие из устройств, составляют набор узлов DAG. Множество его ребер получается следующим образом: когда появляется новая транзакция, она должна подтвердить k других транзакций для общего $k \geq 2$ предыдущих транзакции; данные подтверждения представлены направленными ребрами. Чтобы создать транзакцию, пользователи-устройства должны находиться в рабочем состоянии (например, передавать данные другим устройствам) и подтверждать другие транзакции.

На Рисунке 3 представлен пример направленного ациклического графа. Вершины графа – это транзакции, а направленные дуги показывают, какие транзакции подтверждаются другими транзакциями. А, В, С, D, E, F, G, H – это обозначения набора устройств распределенного реестра. Линии перпендикулярные оси абсцисс показывают принадлежность транзакции к определенному устройству распределенной сети. По оси ординат обозначены номера транзакций в порядке добавления в DAG. Транзакции под номерами 16–18 пока не являются подтвержденными.

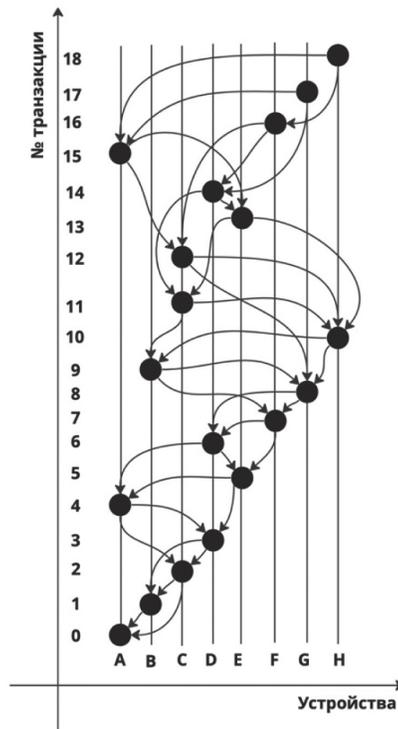


Рисунок 3. Пример ациклического направленного графа

Главное преимущество DAG по сравнению с классическим распределенным реестром в том, что система может обрабатывать транзакции в многопоточном асинхронном режиме (то есть несколько транзакций могут обрабатываться различными устройствами в несколько потоков, при этом время начала и конца обработки транзакции у различных обрабатывающих транзакцию устройств не совпадает). На Рисунке 4 серым цветом в качестве примера отмечены новые транзакции, которые поступили в систему одновременно. Данные транзакции будут обрабатываться параллельно в несколько потоков, при этом время конца обработки может не совпадать.

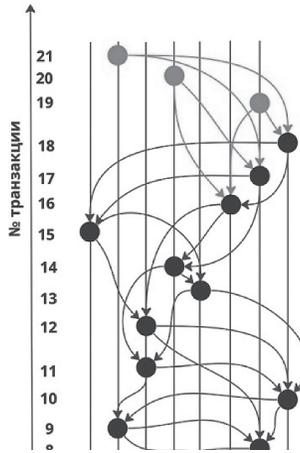


Рисунок 4. Одновременное формирование новых узлов

В соответствии с теорией массового обслуживания данная система принадлежит к классу многоканальных СМО с неограниченной очередью. Количество каналов n в данной системе будет равно количеству устройств, отправляющих транзакции.

Пусть, как и в случае с распределенным реестром, рассмотренном выше, время обработки транзакций для всех устройств в любой момент времени равно $t_{об}$, тогда в данном случае условием стационарности системы будет являться выражение (условие стационарности для многоканальной СМО)

$$\frac{\lambda t_{об}}{n} < 1. \tag{20}$$

Из (20) следует, что у сети DAG есть преимущество по стационарности в сравнении с классическим реестром в n раз, если сравнивать с выражением стационарности системы классического распределенного реестра (16).

Количество транзакций в системе обработки транзакций можно определить следующим образом:

$$L_{сист} = L_{оч} + \rho = \frac{\rho^{n+1} p_0}{n \cdot n! \left(1 - \frac{\rho}{n}\right)^2} + \rho, \tag{21}$$

где $L_{оч}$ – количество заявок, находящееся в очереди на обработку.

Среднее время обработки транзакции определяется по формуле (18), как и в случае с реестром. Подставив в выражение (18) результаты из формулы (21), получим выражение

$$T_{об} = \frac{(\lambda t_{об})^{n+1} p_0}{n \cdot n! \left(1 - \frac{\lambda t_{об}}{n}\right)^2 \lambda} + t_{об}. \tag{22}$$

Значение p_0 можно найти с помощью вычисления суммы модифицированного ряда геометрической прогрессии:

$$\sum_{i=0}^{n \rightarrow \infty} \frac{\rho^i}{i!}. \tag{23}$$

Данный ряд является степенным рядом, который сходится в значении e^ρ . Таким образом, p_0 определяется соотношением

$$e^{\rho} p_0 = \left(\sum_{i=0}^{n \rightarrow \infty} \frac{\rho^i}{i!} \right)^{-1} = e^{-\rho}. \quad (24)$$

В итоге получим соотношение для расчета времени обработки транзакции в распределенной сети ациклического направленного графа

$$T_{об} = \frac{(\lambda t_{об})^{n+1} e^{-\lambda t_{об}}}{n \cdot n! \left(1 - \frac{\lambda t_{об}}{n}\right)^2 \lambda} + t_{об}. \quad (25)$$

При $n \rightarrow \infty$ первое слагаемое выражения (21) будет стремиться к нулю, и $T_{об} \rightarrow t_{об}$ (см. Рисунок 5). По оси ординат отмечено время обработки системой относительно времени обработки транзакции отдельным устройством, по оси абсцисс – количество устройств в сети.

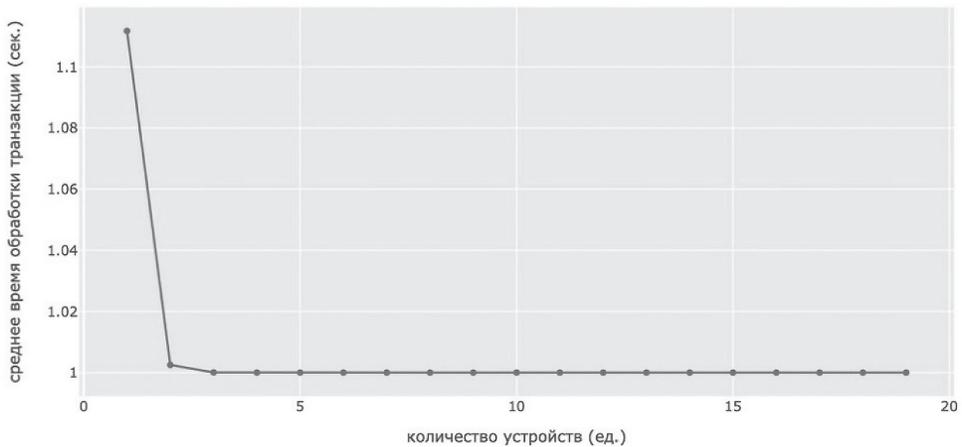


Рисунок 5. График зависимости времени обработки транзакции от количества устройств в сети DAG

В сети DAG время практически не зависит от количества устройств.

Стоит отметить, что при одной и той же величине $t_{об}$ пропускная способность с ростом количества устройств в распределенном реестре остается неизменной, а в DAG увеличивается в n раз. Таким образом, в DAG становится выгодно увеличивать объем сети (см. Рисунок 6).

Если допустить, что увеличение количества устройств приведет к повышению интенсивности (что может быть обусловлено отправкой транзакций подтверждения), то получится зависимость, представленная на Рисунке 7.

Можно заметить, что для DAG большее число устройств в сети повышает ее устойчивость.

В работе были учтены такие факторы, как время обработки транзакции устройствами, интенсивность поступления транзакций в систему, количество устройств. Пропускной способностью каналов связи можно пренебречь в связи с тем, что устройства передают малые по объему пакеты данных. Тем не менее есть неучтенный незначительный фактор – время отклика устройств. Для этого проведен сравнительный анализ показателей скорости, результаты которого представлены в Таблице.

Сравнение скорости обработки транзакций в распределенных системах ...

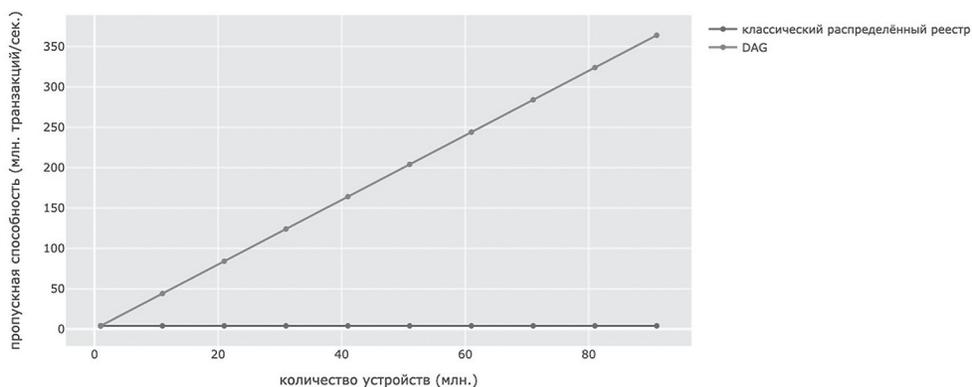


Рисунок 6. График зависимости пропускной способности от количества устройств в сети

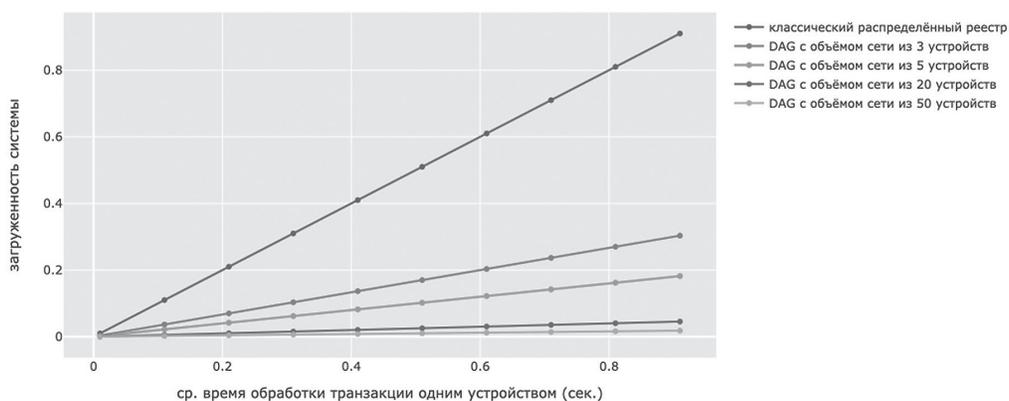


Рисунок 7. Зависимость загрузки системы от среднего времени обработки транзакций для различных распределенных реестров

Таблица

Показатели скорости работы распределенного реестра и ациклического направленного графа при различных величинах интенсивности

Интенсивность нагрузки	Время обработки транзакций распределенным реестром, сек.	Время обработки транзакций ациклическим направленным графом, сек.
0,1	0,006	0,005
0,25	0,007	0,005
0,5	0,011	0,006
0,75	0,018	0,006
0,9	0,081	0,007
0,95	0,126	0,007
0,99	0,708	0,007

Заключение

В работе с помощью законов теории массового обслуживания для классического распределенного реестра и ациклического направленного графа выведены формулы, по которым можно рассчитать время обработки транзакции в системе. Математически доказаны преимущества направленного ациклического графа по скорости обработки транзакций и по пропускной способности по сравнению с классическим распределенным реестром. Таким образом, переход от классического распределенного реестра к DAG может решить проблему масштабирования, так как при росте объемов транзакций данная структура только повышает свою пропускную способность.

Литература

1. Доклад для общественных организаций «Развитие технологий распределенного реестра». Банк России. М., 2017 [Электронный ресурс]. URL: [https://cbr.ru/Content/Document/File/50678/Consultation_Paper_171229\(2\).pdf](https://cbr.ru/Content/Document/File/50678/Consultation_Paper_171229(2).pdf) (дата обращения: 19.04.2023).
2. Котилевец И.Д. Сочетание технологии блокчейн и управления мультиверсионным параллелизмом для безопасной и масштабируемой системы индустриального интернета вещей // Промышленные АСУ и контроллеры. 2021. № 2. С. 22–28. DOI: 10.25791/asu.2.2021.1257
3. Сравнение блокчейн-платформ Hedera Hashgraph, Blockchain и Tangle [Электронный ресурс]. URL: <https://merehead.com/ru/blog/difference-between-hedera-hashgraph-vs-blockchain-vs-tangle/> (дата обращения: 20.04.2023).
4. Котилевец И.Д., Иванова И.А. Применение направленного ациклического графа в блокчейн-сети для повышения безопасности и скорости транзакций в промышленности // Промышленные АСУ и контроллеры. 2019. № 1. С. 53–59. EDN YTOPVJ.
5. Kotilevets I.D., Ivanova I.A., Romanov I O. (2018) Implementation of directed acyclic graph in blockchain network to improve security and speed of transactions. IFAC-Papers OnLine, Baku, 13–15 сентября 2018 года, Vol. 51, Iss. 30, Pp. 693–696. DOI: 10.1016/j.ifacol.2018.11.213
6. Гельдиев Б.А., Хыдыров Р.Б., Оразбердиев М.Ч. Применение теории системы массового обслуживания в эффективном использовании международных транспортных коридоров региона // Вестник науки и образования. 2022. № 10-1(130). URL: <https://cyberleninka.ru/article/n/primenenie-teorii-sistemy-massovogo-obsluzhivaniya-v-effektivnom-ispolzovanii-mezhdunarodnyh-transportnyh-koridorov-regiona> (дата обращения: 07.05.2023).
7. Венцель Е. С. Исследование операций: задачи, принципы, методология. 2-е изд., стер. М.: Наука, 1988. 208 с.

Literature

1. Report for public organizations «Development of distributed registry technologies». Bank of Russia. Moscow, 2017. Available at: [https://cbr.ru/Content/Document/File/50678/Consultation_Paper_171229\(2\).pdf](https://cbr.ru/Content/Document/File/50678/Consultation_Paper_171229(2).pdf) (accessed: 19.04.2023) (in Russian).
2. Kotilevets, I.D. (2021) [The Combination of Blockchain Technology and Multiversion Concurrency Management for a Secure and Scalable Industrial Internet of Things]. *Industrial ACS and Controllers*, 2021, No. 2, Pp. 22–28. DOI: 10.25791/asu.2.2021.1257 (in Russian).
3. Comparison of blockchain platforms: Hedera Hashgraph, Blockchain and Tangle. Available at: <https://merehead.com/ru/blog/difference-between-hedera-hashgraph-vs-blockchain-vs-tangle/> (accessed: 04.20.2023) (in Russian).

4. Kotilevets I.D., Ivanova I.A. (2019) [Application of a directed acyclic graph in a blockchain network to improve the security and speed of transactions in the industry]. *Industrial ACS and controllers*, 2019, No. 1, Pp. 53–59. EDN YTOPVJ (in Russian).
5. Kotilevets I.D., Ivanova I.A., Romanov I O. (2018) Implementation of directed acyclic graph in blockchain network to improve security and speed of transactions. IFAC-Papers OnLine, Baku, 13–15 сентября 2018 года. Vol. 51, Iss. 30, Pp. 693–696. DOI: 10.1016/j.ifacol.2018.11.213
6. Geldiev B.A., Khydyrov R.B., Orazberdiev M.Ch., (2022) [Application of the theory of the queuing system in the effective use of international transport corridors of the region]. *Vestnik nauki i obrazovaniya*, 2022, No. 10-1(130). Available at: <https://cyberleninka.ru/article/n/primeneniye-teorii-sistemy-massovogo-obsluzhivaniya-v-effektivnom-ispolzovanii-mezhdunarodnyh-transportnyh-koridorov-regiona> (accessed: 05.07.2023) (in Russian).
7. Venttsel E. S. (1988) *Issledovanie operacij: zadachi, principy, metodologija* [Research operations: tasks, principles, methodology]. Moscow, Nauka Publishing, 1988, 208 p. (in Russian).