

Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0824U001862

Особливі позначки: відкрита

Дата реєстрації: 09-05-2024

Статус: Наказ про видачу диплома

Реквізити наказу МОН / наказу закладу: № НСВС_62_24 від 23.07.2024



II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Нікітін Валерій Андрійович

2. Valerii Nikitin

Кваліфікація:

Ідентифікатор ORCID ID: 0000-0002-4509-1204

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 126

Назва наукової спеціальності: Інформаційні системи та технології

Галузь / галузі знань:

Освітньо-наукова програма зі спеціальності: Інформаційні системи та технології

Дата захисту: 27-06-2024

Спеціальність за освітою: Інформаційні системи та технології

Місце роботи здобувача: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): ДФ 26.002.160; ID 5546

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 50.41.21

Тема дисертації:

1. Методи підвищення ефективності узгодженості даних в інформаційних системах
2. Methods of increasing the efficiency of data consistency in information systems

Реферат:

1. Дисертаційна робота присвячена розробці методів для розподілених документо-орієнтованих баз даних, які дозволяють пришвидшити узгодження даних та покращити колізійну стійкість в процесі пошуку неузгоджених даних в різноманітних мережевих інформаційних системах, таких як, IoT, гетерогенні мультимедійні системи, аналітичні системи органів адміністративного управління, фінансові системи, дослідження екологічної безпеки та природокористування та інші. Також, було реалізовано спеціальне програмне забезпечення для проведення дослідів. В результаті виконання дисертаційної роботи було отримано такі результати. Розроблено метод забезпечення узгодженості даних у розподілених нереляційних документо-орієнтованих базах даних з використанням транзакційного годинника. Він отримує транзакції від клієнтських додатків та зберігає їх у відповідних чергах. Черги опрацьовуються в залежності від пріоритету

транзакції. При найвищому пріоритеті черга опрацьовується раніше, в порівнянні з тими, які мають нижчий пріоритет. Це дозволяє виділяти критично важливі дані (такі, як фінансові транзакції), які мають бути опрацьовані першочергово. При опрацюванні черги, транзакційний годинник об'єднує транзакції в результуючу транзакцію. Для цього він використовує час створення транзакції, що дозволяє глобально впорядковувати та об'єднувати їх в тому порядку, в якому вони були створені, а не отримані годинником. Після отримання результуючої транзакції, вона передається іншим реплікам, на яких відбувається запис. Слід зазначити, що при використанні транзакційного годинника, операції читання відбуваються напряму з реплік, що дозволяє зменшити навантаження на хостову машину. Удосконалено метод активної антиентропії з використанням спектрального фільтру Блума та алгоритму хешування замість дерева Меркла. Його створення обумовлено тим, що класичний механізм активної антиентропії використовує багато затратних операцій хешування. Також, при хешуванні великої кількості даних збільшується вірогідність колізій, яка може вплинути на вчасну ідентифікацію неузгодженості. Якщо таке буде відбуватися, то система може знаходитись у неузгодженому стані через те, що запуск процедури узгодження розпочнеться пізніше. Тому, для розподілених нереляційних документо-орієнтованих баз даних було вирішено використати протокол пліток, який полягає у децентралізованому способі взаємодії вузлів. Це забезпечує надійність, в порівнянні з централізованим підходом, оскільки виведення з ладу одного вузла не вплине на доступність системи. Проблема використання централізованого підходу полягає в тому, що виведення головного вузла з ладу вимагає певної затримки через використання консенсусного протоколу, який використовується для обрання нового головного вузла з існуючих реплік. Децентралізований підхід дозволяє інформаційній системі бути доступною для операцій запису, хоча це ускладнює підтримку узгодженості. Для пошуку неузгодженостей використовується певний знімок, який складається зі спектрального фільтру Блума та хеш-значення. Спеціально для розробленого методу активної антиентропії було модифіковано алгоритм формування спектрального фільтру Блума, який дозволяє пришвидшити ідентифікування неузгодженості даних в методі активної антиентропії. Швидкість формування спектрального фільтру Блума було збільшено через використання алгоритму на основі простих чисел замість використання хеш-функцій. Результати дослідів показують, що розроблений алгоритм має вищу колізійну стійкість, ніж при використанні однієї хеш-функції та вищу швидкодію у порівнянні з використанням декількох функцій хешування. Також, спеціально для розробленого методу активної антиентропії було розроблено спосіб хешування, який стійкий до колізій. Його мета полягає у зменшенні кількості колізій при хешуванні даних, які відрізняються за розміром. Це є дуже важливим для активної антиентропії, оскільки дозволяє вчасно знайти неузгодженість. Реалізовано спеціальний сервіс транзакційного годинника із власним програмним інтерфейсом для можливості застосування в інформаційних системах розробленого методу забезпечення узгодженості з використанням транзакційного годинника з розподіленою базою даних MongoDB. Реалізовано спеціальний сервіс активної антиентропії із власним програмним інтерфейсом для використання в інформаційних системах розробленого методу активної антиентропії з розподіленою базою даних MongoDB. Реалізовано прототип фінансової інформаційної системи, в якій розподілена база даних складається з восьми вузлів. З її використанням були отримані результати дослідження розроблених методів забезпечення узгодженості даних. Програмний інтерфейс було реалізовано з використанням мови програмування Python3. Для створення тестового середовища використовувався Docker та docker-compose для оркестрації необхідними компонентами.

2. The dissertation work is devoted to the development of methods for distributed document-oriented databases that allow speeding up data reconciliation and improving collision resistance in the process of searching for inconsistent data in various network information systems, such as IoT, heterogeneous multicomputer systems, analytical systems of administrative management, financial systems, research on environmental safety and nature management, and others. Also, special software for conducting experiments was implemented. The following results were obtained as a result of the dissertation work. A method of ensuring data consistency in distributed NoSQL document-oriented databases using a transactional clock has been developed. It receives transactions from client applications and stores them in appropriate queues. Queues are processed depending on the priority

of the transaction. With the highest priority, the queue is processed earlier, compared to those with a lower priority. This allows you to figure out critical data (for example, transmitting financial transactions) that should be processed first. When the queue is processed, the transaction clock merges the transactions into a resulting transaction. To do this, it uses the creation time of the transaction, which allows them to be globally ordered and merged in the order in which they were created, rather than received by the transactional clock. After receiving the resulting transaction, it is transmitted to other replicas on which writing takes place. It should be noted that when using a transactional clock, read operations take place directly from replicas, which allows to reduce the load on the host machine. The active anti-entropy method has been improved using a Bloom spectral filter and a hashing algorithm instead of a Merkle tree. Its creation is due to the fact that the classical mechanism of active anti-entropy uses many expensive hashing operations. Also, when hashing a large amount of data, the probability of collisions increases, which can affect the timely identification of inconsistencies. If this happens, the system may be in an inconsistent state because the start of the reconciliation procedure will occur later. Therefore, for distributed NoSQL document-oriented databases, it was decided to use the gossip protocol, which consists in a decentralized method of node interaction. This provides reliability, compared to a centralized approach, because the failure of one node will not affect the availability of the system. The problem with using a centralized approach is that failing a master requires some latency due to the consensus protocol used to elect a new master from existing replicas. A decentralized approach allows the information system to be available for record operations, although this makes it difficult to maintain consistency. To search for inconsistencies, a certain snapshot is used, which consists of a spectral Bloom filter and a hash value. The algorithm to produce the Bloom spectral filter was modified specifically for the developed method of active anti-entropy, which makes it possible to speed up the identification of data inconsistencies in the method of active anti-entropy. The speed of forming the spectral Bloom filter has been increased by using an algorithm based on prime numbers instead of using hash functions. The results of the experiments show that the developed algorithm has higher collision resistance comparing to single hash function and higher speed comparing to usage of a several hashing functions. Also, a collision-resistant hashing method was developed specifically for the developed active anti-entropy method. Its purpose is to reduce the number of collisions when hashing data that differs in size. This is very important for active anti-entropy, as it allows early mismatch detection. A special transaction clock service with its own application programming interface has been implemented for the possibility of using the developed method of ensuring consistency using a transaction clock with a distributed MongoDB database in information systems. A special service of active anti-entropy with its own application programming interface has been implemented for the use in information systems of the developed method of active anti-entropy with the MongoDB distributed database. A prototype of the financial information system was implemented, in which the distributed database consists of eight nodes. With its use, the results of the study of the developed methods of ensuring data consistency were obtained. The application programming interface was implemented using the Python3 programming language. Docker and docker-compose were used to create test environment and orchestrate the necessary components.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- Nikitin V., Krylov E., Kornaga Y., Anikin V. Combined indexing method in NoSQL databases // Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection. №1(38), 2021. P. 3 – 9. DOI: <https://doi.org/10.20535/1560-8956.38.2021.232948>

- Nikitin V., Krylov E., Kornaga Y., Anikin V. Modification of hashing algorithm to increase rate of operations in NoSQL databases // Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection. № 2 (39), 2021. P. 39 – 43. DOI: <https://doi.org/10.20535/1560-8956.39.2021.247395>
- Mukhin V., Kornaga Y., Zavgorodnii V., Fartushnyi I., Pashov R., Nikitin V., Stepanov A. Method of determining the required number of database nodes in a distributed data processing system / 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT) // IEEE. DOI: <https://doi.org/10.1109/ATIT54053.2021.9678569>
- Nikitin V., Krylov E. Comparison of hashing methods for supporting consistency in distributed databases // Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection. № 1 (40), 2022. P. 48 – 53. DOI: <https://doi.org/10.20535/1560-8956.40.2022.261646>
- Nikitin V., Krylov E. A collision-resistant hashing algorithm for maintaining consistency in distributed nosql databases // Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection. № 2 (41), 2022. P. 45 – 57. DOI: <https://doi.org/10.20535/1560-8956.41.2022.271338>
- Крилов Є. В., Нікітін В. А. Використання транзакційного годинника для пришвидшення процесу узгодження даних в розподілених системах // Фахове видання категорії Б “Науковий вісник Ужгородського університету. Серія «Математика і інформатика»”. № 1 (42), 2023. С. 188 – 192. DOI: [https://doi.org/10.24144/2616-7700.2023.42\(1\).188-192](https://doi.org/10.24144/2616-7700.2023.42(1).188-192)
- Nikitin V., Krylov E. Primary-based Spectral Bloom filter for the ensuring consistency in distributed document-based NoSQL databases using active anti-entropy mechanism // Computer Systems and Information Technologies. №3, (2023). P. 75 – 80. DOI: <https://doi.org/10.31891/csit-2023-3-9>
- Nikitin V., Krylov E. Active anti-entropy mechanism based on Spectral Bloom filter and PH-2 hash algorithm for reconciliation of replicas of NoSQL distributed document oriented databases // Information Technology and Society. №3 (9), 2023. P. 63 – 67. DOI: <https://doi.org/10.32689/maup.it.2023.3.8>
- Nikitin V., Krylov E. Consistency optimization methods in distributed NoSQL databases // Інженерія програмного забезпечення і передові інформаційні технології (SoftTech-2022): матеріали тез доповідей III Всеукраїнської наук.-практ. конф. молодих вчених та студентів (м. Київ, 23–25 листопада 2022 р.). Секція кафедри інформатики та програмної інженерії. – К. : КПІ ім. Ігоря Сікорського, 2022. – С. 27 – 30. URL: https://drive.google.com/file/d/1CP9EaBTT_rJAXsINbanSVGnP2jkg9FJ0/view
- Белоус Р. В., Нікітін В. А. Варіанти забезпечення суворої узгодженості в NoSQL // Grail of Science, (24), с. 364 – 365. DOI: <https://doi.org/10.36074/grail-of-science.17.02.2023.065>

Наукова (науково-технічна) продукція: технології; методи, теорії, гіпотези

Соціально-економічна спрямованість: оптимізація роботи веб-орієнтованих систем з великим об'ємом даних

Охоронні документи на ОПВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами: 0117U004913

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Крилов Євген Володимирович

2. Yevhen O. Krilov

Кваліфікація: к.т.н., доц., 05.13.09

Ідентифікатор ORCID ID: 0000-0003-4313-938X

Додаткова інформація: <https://scholar.google.ru/citations?user=jvBZVWUAAAAJ>

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Триснюк Василь Миколайович

2. Vasily M. Trisnuk

Кваліфікація: д. т. н., професор, 21.06.01

Ідентифікатор ORCID ID: 0000-0001-9920-4879

Додаткова інформація:

Повне найменування юридичної особи: Інститут телекомунікацій і глобального інформаційного простору

Код за ЄДРПОУ: 26022051

Місцезнаходження: Чоколовський бульвар, буд. 13, Київ, 03186, Україна

Форма власності: Державна

Сфера управління: Національна академія наук України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Соколов Володимир Юрійович

2. Volodymyr Sokolov

Кваліфікація: к. т. н., доцент, 05.13.06

Ідентифікатор ORCID ID: 0000-0002-9349-7946

Додаткова інформація:

Повне найменування юридичної особи: Київський столичний університет імені Бориса Грінченка

Код за ЄДРПОУ: 45307965

Місцезнаходження: вул. Бульварно-Кудрявська, 18/2, Київ, 04053, Україна

Форма власності: Комунальна

Сфера управління: Держадміністрація

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Корнієнко Богдан Ярославович

2. Bogdan Y. Korniyenko

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCID ID: 0000-0002-2521-0878

Додаткова інформація:

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Олійник Юрій Олександрович

2. Yurii O. Oliinyk

Кваліфікація: к. т. н., доцент, 05.13.06

Ідентифікатор ORCID ID: 0000-0002-7408-4927

Додаткова інформація: ;<http://orcid.org/0000-0002-7408-4927>

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові

голови ради

Жураковський Богдан Юрійович

**Власне Прізвище Ім'я По-батькові
головуючого на засіданні**

Жураковський Богдан Юрійович

**Відповідальний за підготовку
облікових документів**

Нікітін Валерій Андрійович

Реєстратор

УкрІНТЕІ

**Керівник відділу УкрІНТЕІ, що є
відповідальним за реєстрацію наукової
діяльності**



Юрченко Тетяна Анатоліївна