

Інститут проблем машинобудування ім. А. М. Підгорного  
Національної академії наук України

Жолткевич Галина Григоріївна



УДК 004.942

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ПРОЦЕСІВ РЕПЛІКАЦІЇ  
ДАНИХ У РОЗПОДІЛЕНИХ СХОВИЩАХ

Спеціальність 01.05.02 – математичне моделювання та  
обчислювальні методи

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня  
кандидата технічних наук

Харків-2021

Робота виконана в Харківському національному університеті імені В.Н. Каразіна Міністерства освіти і науки України.

Науковий керівник

доктор технічних наук, доцент  
Руккас Кирило Маркович,  
Харківський національний  
університет імені В.Н.Каразіна,  
професор кафедри теоретичної та  
прикладної інформатики

Офіційні опоненти

доктор технічних наук, професор  
Гребеннік Ігор Валерійович,  
завідувач кафедри системотехніки  
Харківського національного  
університету радіоелектроніки,

доктор технічних наук, доцент  
Глибовець Андрій Миколайович,  
декан факультету інформатики  
Національного університету «Києво-  
Могилянська академія».

Захист відбудеться « 1 » квітня 2021 року о 14.00 годині на засіданні спеціалізованої вченої ради Д 64.180.01 при Інституті проблем машинобудування ім. А.М. Підгорного НАН України за адресою: 61046, м. Харків, вул. Пожарського, 2/10.

З дисертацією можна ознайомитися у науково-технічній бібліотеці Інституту проблем машинобудування ім. А.М. Підгорного НАН України за адресою: 61046, м. Харків, вул. Пожарського, 2/10.

Автореферат розісланий « 26 » лютого 2021 року.

Учений секретар спеціалізованої  
вченої ради  
д.т.н., проф.



О.О. Стрельнікова

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** В епоху повсюдного використання сучасних технологій, таких як Big Data, Internet of Things, хмарні обчислення тощо, дані стають все більш важливими і потребують більшого, надійного зберігання. Це призводить до збільшення розміру розподілених сховищ, що далі потребують величезної мережі по всіх розподілених вузлах. Але існує декілька невирішених проблем із використанням великих розподілених сховищ даних, і деякі з них тісно пов'язані з теоремою CAP (Consistency, Availability, Partition tolerance).

Враховуючи, що відома стратегія ACID (atomicity, consistency, isolation, durability), яка використовується скрізь у реляційних базах даних, не може задовольнити всі потреби систем класу розподілених сховищ даних, механізми передачі даних крізь розподілене сховище не забезпечують в достатній мірі узгодженість, надійність та толерантність до мережевих розділів (основні фактори надійного сховища даних).

Масштабуванню розподілених систем останніми роками і дослідженню їх продуктивності приділяється багато уваги. Проблема CAP-теорема (не можна одразу задовольнити три основні характеристики у розподіленому сховищі) і порівняння BASE (Basically Available, Soft state, Eventual consistency) та ACID моделі вивчається вже понад 15 років. За останні 10 років є понад 40 тисяч статей, що присвячені цій темі. Є статті з аналізом CAP-теорема відомих американських учених Гілберта та Лінч, де розглядаються варіанти досягнення компромісу між AP (доступність, стійкість до розділень) або CP (узгодженість, стійкість до розділень) і CA (узгодженість, доступність). Google нещодавно розробили розподілену базу даних, яка забезпечує CP і майже забезпечує всі три гарантії, але це рішення дуже дороге у його підтримці. Тобто стає зрозумілим, що проблема CAP-теорема досліджена точково з різних боків, і немає систематизованого алгоритму підтримки узгодженості, доступності та стійкості до розділень, який був би стверджений стандартом. У роботі переслідуються мета розробити універсальний метод прийняття рішень, на який можливо спиратись, будуючи мережу для розподіленого сховища даних, що дозволить мати конкретні критерії для формування сховища за умов різноманітних потреб конкретної розподіленої системи. Для цього необхідно дослідити розподілені сховища даних, і щоб не порушити і не втрутитись у його роботу та заради уникнення помилок у дослідженні, будуються математичні і імітаційні моделі для такого сховища. Таким чином, постановка задачі дослідження даної дисертаційної роботи є актуальною.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційна робота виконувалася на кафедрі теоретичної та прикладної інформатики факультету математики та інформатики Харківського національного університету імені В.Н. Каразіна відповідно до завдань прикладної НДР ДР № 0116U003141.

**Мета і задачі дослідження.** Метою дисертаційної роботи є забезпечення контрольованого рівня узгодженості реплік даних в РСД (розподіленого сховища даних) при необхідних рівнях доступності даних і стійкості сховища даних до розділення на етапі проектування розподіленої системи задля управління ризиками некоректного функціонування системи за рахунок побудови математичних

моделей і розробки обчислювальних методів їх дослідження. Для досягнення даної мети в роботі поставлені та вирішені такі задачі:

- побудувати математичні моделі для формалізації, обґрунтування і розробки методів аналізу процесів реплікації;
- запропонувати метрики оцінювання рівня задоволення CAP-вимог;
- запропонувати обчислювальні методи для підвищення рівня узгодженості реплік РСД на основі раціонального збалансування запропонованих метрик;
- реалізувати імітаційні моделі для експериментального дослідження розроблених методів;
- оцінити часову та просторову складність розроблених обчислювальних методів.

**Об'єктом дослідження** є процеси реплікації даних у великих розподілених сховищах даних (РСД).

**Предметом дослідження** є математичні та імітаційні моделі процесу реплікації даних у РСД, методи забезпечення узгодженості.

**Методами дослідження** в цій дисертаційній роботі є математичний базис теорії графів та теорії ймовірності, програмне моделювання для імітаційних моделей розподіленого сховища даних, де використовувались інструменти мови UML та програмне забезпечення для проведення експериментальних досліджень на РСД.

**Наукова новизна отриманих результатів** полягає в такому:

- зазнала подальшого розвитку стохастична модель розподіленої системи за рахунок забезпечення формалізації виконання CAP-гарантій, яка, на відміну від існуючих моделей, об'єднує формалізовані визначення для CAP-гарантій в одну цілісну структуру і дає підставу для оцінки гарантій на іншому рівні;
- вперше запропоновано систему статистичних метрик для оцінювання виконання CAP-гарантій у РСД, що базується на формалізації визначень CAP-гарантій, і замість бінарної оцінки гарантій РСД пропонується оцінювати систему гарантій за допомогою ймовірнісних величин, що дає підставу для розвитку метрик для оцінки рівня CAP-гарантій;
- удосконалено методи обчислення ймовірностей втрати даних на основі запропонованої моделі, що дає можливість оцінки CAP-гарантій і формування метрик для оцінки з більшою точністю, з врахуванням реальних умов таких систем, і, таким чином, це дає підставу для формування методу регулювання CAP-характеристик на рівні, потрібному конкретній розподіленій системі;
- зазнали подальшого розвитку методи забезпечення CAP-вимог шляхом розробки нової стратегії контрольованого балансування узгодженості у формі спеціального методу контрольованого балансування узгодженості, що може бути інтегрований з методами реплікації у РСД.

**Практичне значення.** На основі запропонованих методів та моделей реалізовано метод регулювання CAP-гарантій у формі алгоритму для покриття такого рівня CAP-гарантій, які будуть необхідні конкретному РСД. Цей метод може бути застосований на етапі проектування системи РСД для запобігання помилок і неоптимальних рішень побудови розподіленої мережі, а також невірному виборі

або застосування методів забезпечення CAP-гарантій, тобто зменшення ризиків, які коштуватимуть дорожче на етапі застосування такої системи. Також такий метод містить елементи, які мають працювати на етапі застосування системи, а саме регулювання рівня узгодженості з метою утримування її на необхідному рівні. Розроблені в дисертаційній роботі моделі, метрики та методи можуть бути використані як компоненти систем моніторингу РСД з метою контролювання CAP-гарантій і оперативного реагування на інциденти в системі. Тож, такі методи використовуються для підвищення ефективності роботи РСД, мережі в яких налічують сотні і тисячі вузлів.

**Особистий внесок здобувача.** Наукові результати, викладені в дисертаційній роботі, одержано здобувачем особисто. У роботах, які написано в співавторстві, особистий внесок здобувача полягає у наступному: сформовані елементи мови для визначення моделі годинника, що може використовуватися у розподіленій системі для обчислення точного часу при визначенні версій реплік і визначення, яка репліка найновіша, з метою коректної роботи операції заміни репліки на новішу (див. [1], [5]); зроблено аналіз моделі якості розподіленого сховища даних, що дало можливість оцінити, які саме характеристики більш важливі для такої системи, формалізовано визначення для CAP-гарантій [6]; розвинено метрики для оцінки рівня гарантії узгодженості у РСД та доведено гіпотезу про час узгодженості у РСД, що дало підставу для розвитку оцінки часу для виконання повної реплікації, тобто відновлення системи до повністю узгодженої ([7]); розширено модель оцінки CAP-гарантій за рахунок розвинення метрики оцінки часу для виконання повної реплікації в реальних умовах розподіленої системи ([3]); застосовано метод балансування рівня CAP-гарантій залежно від потреб системи [4].

**Апробація результатів дисертації.** Основні положення і результати дисертаційних досліджень пройшли апробацію та представлені на 9<sup>th</sup> International Conference on ICT in Education, Research, and Industrial Applications ICTERI (Херсон, 2013), 11<sup>th</sup> International Conference on ICT in Education, Research, and Industrial Applications ICTERI (Львів, 2015), XIV Міжнародній науково-практичній конференції «Теоретичні та прикладні аспекти побудови програмних систем» ТАAPSD'2017 (Київ, 2017), а також на VI міжнародній науково-технічній конференції «Комп'ютерне моделювання та оптимізація складних систем» (Дніпро, 2020).

**Публікації.** За результатами дисертаційної роботи опубліковано 8 наукових праць, у тому числі 3 статті у наукових фахових виданнях України з технічних наук, 1 стаття в іноземному спеціалізованому виданні (наукометрична база «Scopus») і 4 тези доповідей на фахових міжнародних наукових конференціях.

**Структура та обсяг дисертації.** Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел з 102 найменувань на 11 сторінках та 7 додатків на 13 сторінках, а також містить 54 рисунка. Загальний обсяг роботи складає 138 сторінок, включаючи 110 сторінок основного тексту.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** наведено загальну характеристику роботи: актуальність теми та аналіз пов'язаних з темою робіт, предмет, мета і задачі дослідження, об'єкт дослідження, наукова новизна та практичне значення отриманих результатів, а також наведений список публікацій за темою дисертаційної роботи.

У **першому розділі** проведений аналіз сучасних протоколів поширення реплік, забезпечення консенсусу, механізмів забезпечення CAP-гарантій даних та сучасних реалізацій реплікації у розподілених базах даних. Розглянуто розподілені бази даних, вказано їх актуальність та класифікація, зроблено опис основних парадигм в архітектурі таких баз даних, описано модель якості баз даних та характеристики і принципи їх надійності, проаналізовано механізми внутрішньої роботи обробки запитів до БД. Також зроблена порівняльна характеристика баз даних з відкритим кодом і їх архітектури та зроблені висновки про можливість внеску нових змін, що плануються у реалізації БД з відкритим кодом.

Окрім того, перший розділ містить опис CAP-гарантій та CAP-теорему і порівняльну характеристику розподілених баз даних. Згадано, що CAP-теорема вводить певні обмеження: не можна одночасно задовольнити всі три гарантії одночасно, які є основними для масштабованого надійного РСД – узгодженість, доступність, стійкість до розділень. Зважаючи на ці обмеження, робиться порівняльна характеристика рішень, що були запроваджені у різних реалізаціях РСД для забезпечення тих чи інших CAP-гарантій, і порівняльна характеристика гарантій, що забезпечує та чи інша система РСД.

Такий детальний аналіз показує, що нема структурованого стандартизованого рішення, що могло б послужити механізмом регулювання CAP-гарантій та механізмом прийняття рішень, що дасть можливість знайти баланс між гарантіями, залежачи від конкретної специфікації потреб системи РСД, що і стало підставою для даного дослідження.

**Другий розділ** присвячено розробці математичних моделей процесів обробки даних у розподілених базах даних з метою аналізу розподілених баз даних та розвитку методів забезпечення CAP-гарантій на потрібному рівні. Побудова таких моделей дає можливість формалізувати метрики відповідності CAP-гарантіям і розробити методи забезпечення цих гарантій на необхідному рівні для розподіленого сховища. Враховуючи складність об'єкту дослідження, запропоновано комплекс моделей для РСД, який складається з двох базових моделей: моделі комп'ютерної мережі, моделі управління даними, і методу управління репліками, що описує забезпечення узгодженості даних в решті-решт.

Модель мережі. Математичною моделлю комп'ютерної мережі є простий неорієнтований граф  $G = (N, E)$ , де елементи множини  $N$  – вузли, що відповідають серверам даних, елементи множини  $E$  – комунікаційним каналам. Кожне ребро  $e \in E$  є двоелементною множиною вузлів, які відповідають двом серверам, що комунікують через відповідний канал, представлений цим ребром.

До моделі мережі додатково включено такі три функції:

$$p_a : N \rightarrow [0,1], \quad (1)$$

$$p_d : E \rightarrow [0,1], \quad (2)$$

$$c : E \rightarrow R_+, \quad (3)$$

перша з яких визначає ймовірність доступності відповідного вузла, а друга і третя – відповідно ймовірність доставки повідомлення через відповідний комунікаційний канал та мережеву ціну цієї доставки.

Модель управління даними. Ця модель розвиває централізовану модель управління даними, для опису якої фіксуються дві множини:  $K$  – зліченна множина ключів і  $V$  – домен (множина можливих) значень.

Станом даних називається скінченна підмножина

$$s \subset K \times V \quad (4)$$

яка задовольняє умові

$$\forall k \in K, \forall v', v'' \in V, (k, v') \in s \rightarrow (k, v'') \rightarrow v' = v'', \quad (5)$$

що забезпечує функціональність залежності значення від ключа.

Така скінченна множина  $s$  моделює сукупність одиниць даних, кожна з яких є парою  $(k, v)$ , де  $k \in K$  є ключем, що посилається на значення  $v \in V$ , відома в теоретичній інформатиці як “словник”. Зазначимо, що ця модель даних ідентифікує одиницю даних  $(k, v)$  з її ключем  $k$ .

Позначимо через  $D(K, V)$  сім'ю всіх словників, ключі яких належать множині  $K$ , а значення - множині  $V$ . Зрозуміло, що ця сім'я містить порожній словник, тобто словник, що відповідає порожній підмножині множини  $K \times V$ . Для словників сім'ї  $D(K, V)$  можна визначити оператори

$$write : K \times \bar{V} \rightarrow D(K, V) \rightarrow D(K, \bar{V}), \quad (6)$$

$$read : K \rightarrow D(K, V) \rightarrow \bar{V} \quad (7)$$

у такий спосіб

$$write(k, none)s = s \setminus (\{k\} \times V), \quad (8)$$

$$write(k, v)s = write(k, none) \cup \{(k, v)\}, v \in V;$$

$$read(k)s = v, (k, v) \in s, \quad (9)$$

$$read(k)s = none, (\{k\} \times V) \cap s = \emptyset.$$

Тут і далі використовується позначення  $\bar{V}$  для посилання на множину значень  $V$ , розширену елементом *none*.

Простим, але важливим з точки зору коректності моделі даних, є наступне твердження.

**Твердження 1.** Результатом застосування оператора  $write(k, v)$  для будь-яких  $k \in K, v \in \bar{V}$  до будь-якого словника  $s \in D(K, \bar{V})$  є словник.

Для переходу до моделі управління розподіленими даними враховується, що у РСД одиниці даних розподілені по вузлах мережі як репліки. Тобто, моделлю розподілу даних у сховищі є відображення  $s : N \rightarrow D(K, \bar{V})$ .

Однак, ця модель може бути зведена до словника у такий спосіб: ототожнимо  $s : N \rightarrow D(K, \bar{V})$  і множину  $S = \{(n, k, v) \in N \times K \times \bar{V} \mid (k, v) \in s(n)\}$ .

*Твердження 2.* Множина  $S$  є словником з ключами з множини  $N \times K$  і значеннями з множини  $\bar{V}$ . Навпаки, якщо  $S \in D(N \times K \times \bar{V})$ , тоді  $s : N \rightarrow D(K, \bar{V})$  є РСД.

Враховуючи цей результат, далі використовуються словники з сім'ї  $D(N \times K, \bar{V})$  як моделі розподілених сховищ даних. При цьому пара  $(k, v) \in K \times \bar{V}$ , яка задовольняє умові:  $\exists n \in N, (n, k, v) \in S$ , називається одиницею даних, що зберігає РСД у стані  $S$ . Дві такі одиниці даних, які мають спільний ключ  $k \in K$ , називаються репліками одна одної. Слід зазначити, що визначення операторів *write* і *read* не є тривіальним для РСД. Це, наприклад, можна зробити у такий спосіб

$$\begin{aligned} \text{write}(k, \text{none})S &= S \setminus (\{(n, k)\} \times \bar{V}), \\ \text{write}(k, v)S &= (\text{write}(k, \text{none})S \cup \{(n, k, v)\}); \end{aligned} \quad (10)$$

$$\begin{aligned} \text{read}(k)S &= (n, k, v) \in S, \\ \text{read}(k)S = \text{none}, (\{n, k\} \times \bar{V})S &= \emptyset, \end{aligned} \quad (11)$$

де  $n \in N$  вибирається випадково.

Проблемою цього визначення є залежність результату виконання операторів від вибору  $n \in N$ . Тому слід зазначити, що будь-яке визначення операцій, узгоджене з їх визначенням для централізованих сховищ даних, веде до неоднозначності, яка називається неузгодженістю (inconsistency). Для уникнення неузгодженості вводиться наступна вимога. Припустимо, що поточним станом РСД є  $S \in D(N \times K, \bar{V})$ . Будемо казати, що  $S$  узгоджено за ключем  $k \in K$ , якщо для будь-яких вузлів  $n' \neq n'' \in N$  і значення  $v \in \bar{V}$  умови  $(n', k, v) \in S$  і  $(n'', k, v) \in S$  виконуються чи не виконуються одночасно. Стан  $S$  РСД називають узгодженим, якщо він є узгодженим за всіма ключами. Позначимо множину узгоджених станів РСД через  $D \subset (N \times K, \bar{V})$ . Зрозуміло, що для узгодженого стану РСД результат виконання оператора *read* не залежить від вибору вузла мережі РСД. Опишемо можливі варіанти визначення оператора *write*, для яких узгодженість даних є інваріантом.

$$\begin{aligned} \text{write} : \{(n, k, v) : K \times \bar{V} \rightarrow D(K, \bar{V}) \rightarrow D(K, \bar{V})\}, \\ \text{read} : \{(n, k) \mid (n, k, v) : N \times K \rightarrow D(K, \bar{V}) \rightarrow \bar{V}\} \end{aligned} \quad (12)$$

У централізованій моделі управління даних є одне джерело істини, тому виконується умова (5) у будь-який момент часу. Щоб дослідити виконання умови (5) на розподіленій моделі даних, введемо твердження абсолютної узгодженості. Репліки на вузлах узгоджені повністю, коли на будь-який запит на одиницю даних отримані значення співпадають

$$C = \forall n \in N, k \in K, \forall v', v'' \in \bar{V}, s_n(k, v') \in s \wedge (k, v'') \in s \rightarrow v' = v''. \quad (13)$$

Будемо називати ймовірністю стану узгодженості  $\text{Pr}(C)$  ймовірність виконання такої умови.

На основі визначених теоретичних елементів, опису розподіленої моделі даних та сформованої моделі РСД представлено базовий метод розповсюдження даних з метою остаточного узгодження даних у РСД, що описує процес



розповсюдження реплік, тобто є процесом реплікації. Тригером реплікації є факт успішного виконання операції *write*. Ця операція у РСД – найскладніша з описаних. За умови роботи такого методу система знову стане узгодженою. Вирішення конфліктів та забезпечення узгодженості на етапі паралельних запитів на запис на різні сервери забезпечується іншими методами, такими як алгоритми консенсусу та вирішення конфліктів за допомогою часових міток кожної репліки, що інтегруються у систему у вигляді окремих модулів-надбудов над методами реплікації. Метод, що описує процес реплікації, описаний нижче у формі такого алгоритму, поданого на Рис. 1.

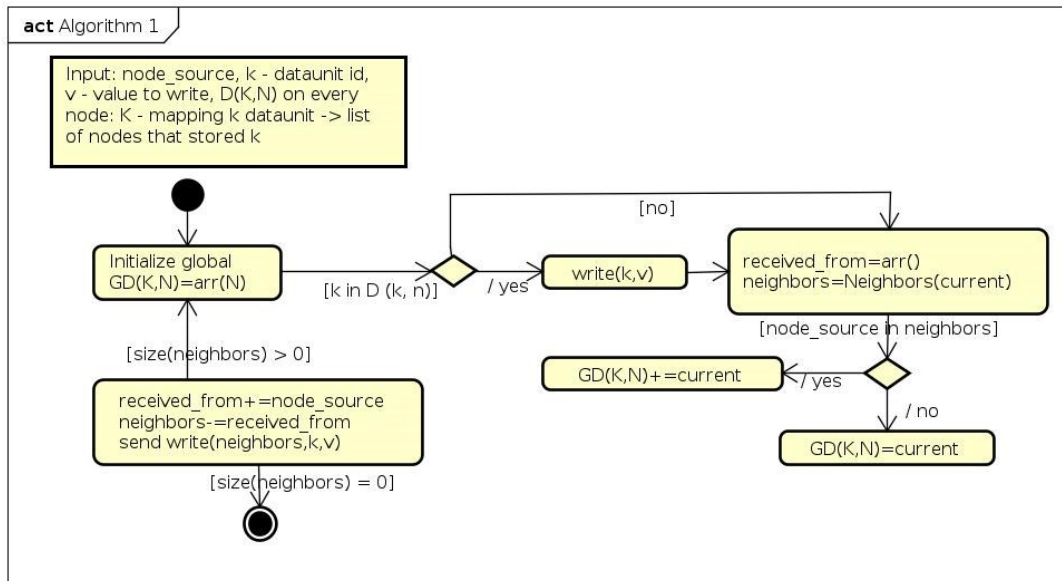


Рис. 1. Алгоритм процесу реплікації в РСД

Таким чином, послідовно побудовано базові моделі для розподіленого сховища для опису еволюції РСД та описано базовий метод поширення реплік. Ці моделі та метод формують собою теоретичний базис для подальшого розвитку методів забезпечення CAP-гарантій та формування алгоритму прийняття рішень, що забезпечить CAP-гарантії на рівні, який потрібен для встановлених потреб такої системи залежно від певної її специфікації.

У **третьому розділі** розглядаються CAP-гарантії для розподіленого сховища даних. Формалізовано визначення для кожної з CAP-гарантій на основі теоретичної бази, сформованої у другому розділі, які надають підставу для розробки метрик CAP-гарантій, зокрема, метрики для оцінки рівня узгодженості. Також у цьому розділі наводяться і обґрунтовуються припущення щодо середньої ціни відновлення узгодженості у РСД, тобто середньої ціни процесу реплікації для однієї події запису у РСД.

Метрика узгодженості. Узгодженість у РСД гарантує, що версії реплік на всіх вузлах співпадають у будь-який момент часу. Нагадаємо, що виокремлюють повну узгодженість, коли задовольняється ця умова, та узгодженість врешті-решт, коли через деякий період часу всі репліки в системі РСД будуть співпадати. Формальне визначення метрики для оцінювання рівня цієї гарантії РСД запропоновано нижче.

Нехай  $\sigma_t \in \{0,1\}$  – випадкова величина, яка демонструє одну з таких подій у момент часу  $t \geq 0$ :

$\sigma_t = 0$  – відповідає події «існує одиниця даних, яка неузгоджена у момент часу  $t$ », тобто не виконується умова (13);

$\sigma_t = 1$  – відповідає події «всі одиниці даних узгоджені у момент часу  $t$ », тобто виконується умова (13).

Метрика доступності. Доступність полягає у тому, що кожен запит повинен отримати коректну відповідь. Формальне визначення метрики для оцінювання такої CAP-гарантії для РСД виглядає наступним чином.

Нехай  $\tau_t$  – часовий інтервал між подією отримання запиту у момент часу  $t$  та відповіддю на цей запит. Тоді середній час доставки визначається за формулою

$$T(t) = E[\tau_t]. \quad (14)$$

Метрика стійкості до розділень. Стійкість до розділень гарантує, що попри мережеві розділення на дві чи більше ізольовані частини, система спроможна коректно відповідати на запити і оперувати, незважаючи на втрати даних внаслідок розділень.

Нехай  $\zeta_t \in \{0,1\}$  – випадкова величина, яка демонструє одну з таких подій у момент часу  $t \geq 0$ :

$\zeta_t = 0$  – відповідає події «існує одиниця даних, недосяжна з хоча б одного вузла у момент часу  $t$ »;

$\zeta_t = 1$  – відповідає події «всі одиниці даних досяжні з будь-якого вузла у момент часу  $t$ ». Тоді метрика стійкості до розділень визначається за формулою

$$P(t) = \Pr(\zeta_t = 1) \quad (15)$$

На основі наданих метрик сформовані наступні твердження.

*Твердження 3.* Нехай  $\epsilon$  РСД з  $N$  множиною вузлів. Нехай  $N_d$  – множина вузлів, що має одиницю даних  $d$ , а  $n_c$  – множина вузлів, що мають одну версію репліки. Тоді ймовірність того, що з множини  $N_d$  два вузли мають одну й ту ж репліку (тобто вони узгоджені між собою) дорівнює

$$p_c = \frac{n_c}{l(N_d)} \cdot \frac{n_c - 1}{l(N_d) - 1} = \frac{n_c \cdot (n_c - 1)}{l(N_d) \cdot (l(N_d) - 1)} \quad (16)$$

Якщо досліджувана одиниця даних представлена на  $N = N_1, \dots, N_k$ , то на цій множині  $N$  репліки можуть бути однаковими або можуть розходитись. Будемо розглядати у даному випадку ймовірність неузгодженості між репліками на представлених вузлах. Тоді формула ймовірності загальної неузгодженості між репліками на представлених вузлах

$$I(d) = 1 - \sum_{k=1}^K \frac{N_k (Y_k - 1)}{N(N - 1)}. \quad (17)$$

Для теоретичного доведення розглянуто систему з 5 вузлами і доведено правильність формули для такої системи. Після цього проведено експериментальне дослідження для 100 вузлів і отримано числа, кожна з яких лежить близько

величини, яка порахована згідно з формулою, з невеликою похибкою (див Рис. 2, Рис. 3).

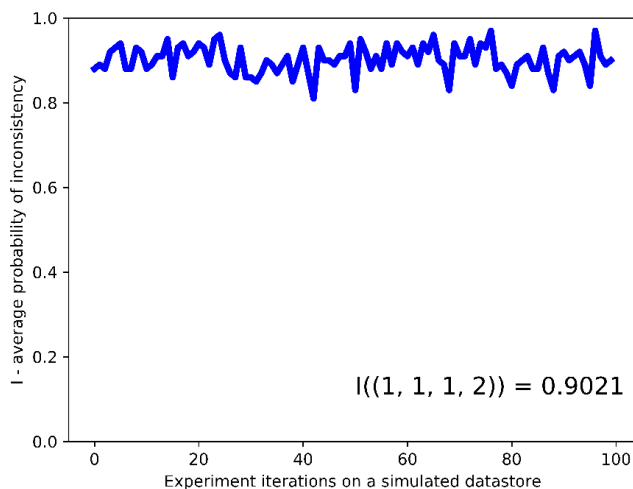


Рис. 2. Експеримент узгоджених розділень на РСД з 5 вузлами

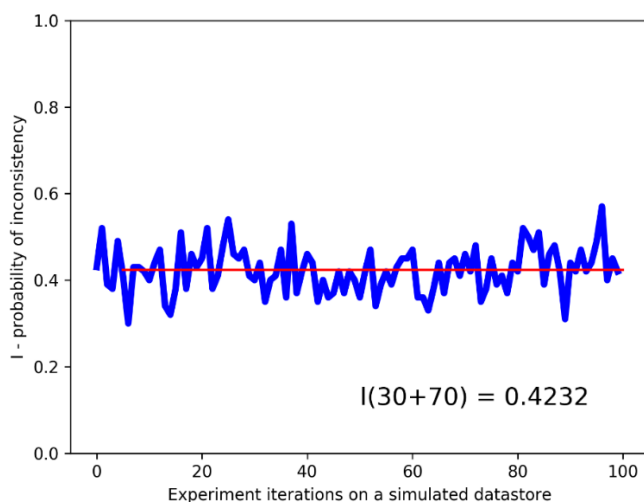


Рис. 3. Експеримент узгоджених розділень на РСД із 100 вузлами

*Твердження 4.* Припустимо, що у РСД немає втрат даних, та всі зв'язки – стійкі до розділень. Припустимо, що надійшло одне повідомлення. Тоді ціна відновлення повної узгодженості буде не більше за діаметр графу, який представляє топологію мережі сховищ.

На Рис. 4 показано, що для кожного РСД з різним діаметром (вісь Y) кількість часових слотів (вісь X) не перебільшує діаметру графу, що представляє РСД.

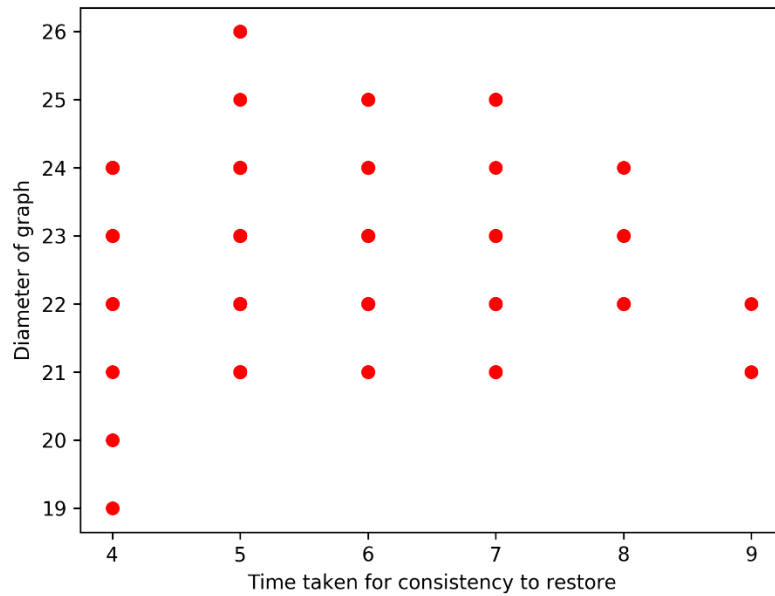


Рис. 4. Залежність ціни відновлення узгодженості до діаметра графа

Далі це дослідження отримало розвиток як доповнені характеристики РСД, які відповідають реальним умовам, а саме коли є ймовірність втрат даних, збоїв на зв'язках або вузлах. Таким чином, з'являється такий термін як ймовірність доставки. Обчисливши ймовірність доставки і використовуючи формули середнього – математичного сподівання, можна обчислити усереднений час доставки  $\tau$  за формулою

$$\begin{aligned}
 P_k &= P_{k1} \cdot \dots \cdot P_{kn} = \prod_{j=1}^n P_{kj}, \\
 P(\text{delivery}) &= 1 - \prod_{k=1}^n p_k(\text{loss}) = 1 - \left( \prod_{i=1}^n (1 - \prod_{j=1}^m p_{ij}) \right), \\
 \tau_{ij} &= \sum_{k=1}^n t_k \cdot \left( 1 - \frac{P_k}{P(\text{delivery})} \right).
 \end{aligned} \tag{18}$$

Такі формули дали підставу для формування практичних рекомендацій для проектування РСД, що є елементом алгоритму прийняття рішень.

Повернемося до оцінки рівня узгодженості даних. Маючи метрики для оцінювання поточного рівня узгодженості і можливість оцінювати узгодженість не бінарним, а статистичним шляхом, з'являється підстава для методу моніторингу та регулювання рівня узгодженості у РСД. Для цього метод регулювання узгодженості має бути інтегрований у формі алгоритму в метод розповсюдження даних, загальний опис якого подано у другому розділі, а також інтегрований в метод відповіді на запит на читання.

Метод реплікації. Метод регулювання узгодженості складається з двох частин: перша – для запису даних, друга – для відповіді на запит на читання. Перша частина методу ініціює процес, що дозволяє контролювати узгодженість на заданому рівні. Задача першої частини полягає в тому, щоб динамічно змінювати глобальний словник  $CD(K, N)$ , який має інформацію про те, які вузли узгоджені

між собою. Першу частину методу, що забезпечує контроль узгодженості в процесі запису даних, подано у формі алгоритму на Рис. 5.

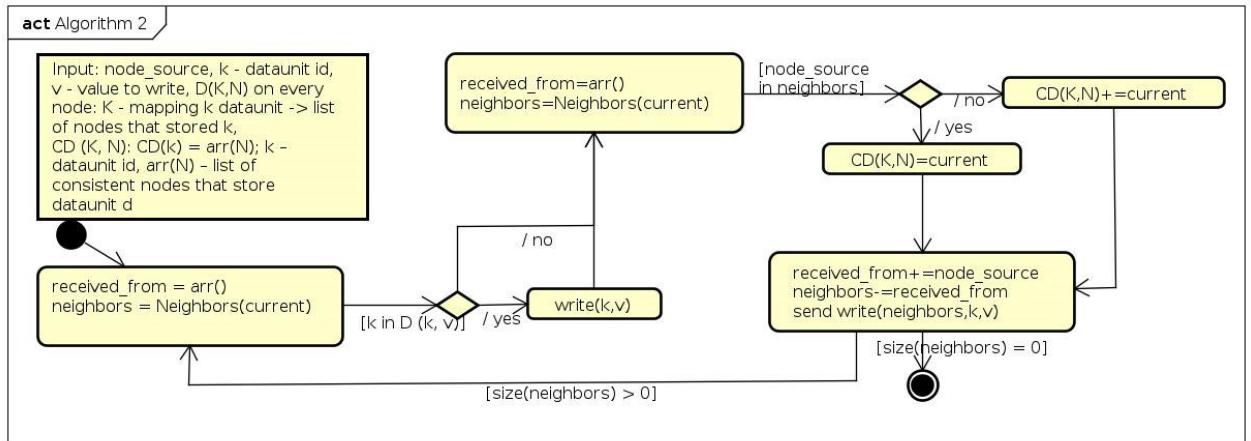


Рис. 5. Удосконалений алгоритм процесу реплікації

Друга частина методу, яка дозволяє контролювати рівень узгодженості, повинна бути інтегрована з алгоритмом відповіді на запит на читання. Такий алгоритм отримує інформацію зі спільного словнику, який має інформацію про те, які вузли узгоджені між собою, і який динамічно змінюється в процесі реплікації. Така інформація використовується на етапі відповіді на запит на читання за умови, якщо кількість узгоджених між собою вузлів не нижче за заданий поріг. У протилежному разі є ризик сильного навантаження на систему у поточному її стані, і краще відповісти неузгодженою інформацією, ніж втратити бажаний рівень доступності системи РСД. Алгоритм подано нижче на Рис. 6.

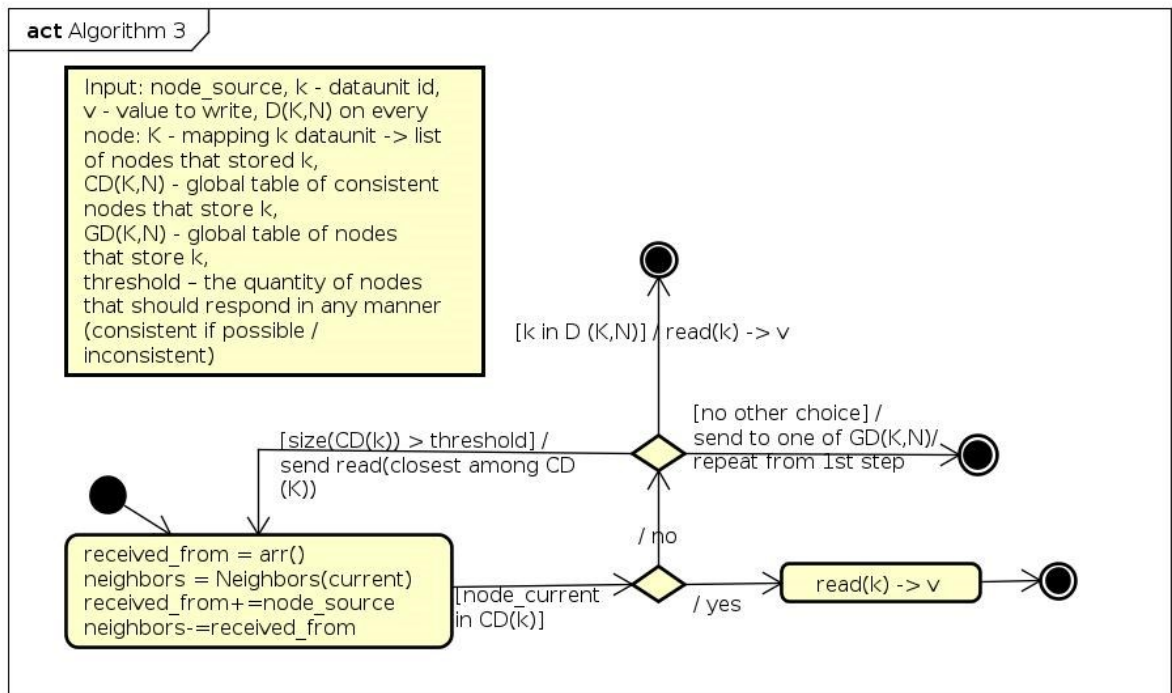


Рис. 6. Алгоритм процесу відповіді на запит на читання

Наявність метрик для обчислення рівня середньої ціни доставки і інші запропоновані метрики для оцінки рівня CAP-гарантій дають можливість

сформувати метод для прийняття рішень, який може бути використаний на етапі побудови системи РСД для оптимальної її роботи у контексті CAP-гарантій. Метод подано у вигляді практичних рекомендацій щодо порядку проектування РСД.

Крок 1. Сформувати граф, що представляє топологію мережі РСД з початковими параметрами: для кожного вузла – ймовірність, що вузол може коректно відповісти на запит; для кожного зв'язку – ймовірність того, що пакет буде доставлений через даний зв'язок; для кожного зв'язку – ціна доставки. Обчислити середню ціну доставки між будь-якими вузлами за допомогою формули середньої ціни доставки.

Крок 2. Якщо ціна доставки не задовольняє поточні потреби системи, відсортувати список вузлів у системі за важливістю (вузол, який має більшу кількість сусідів, вузол, на який за наданою статистикою надходить більше запитів тощо) у низхідному порядку.

Крок 3. Застосувати підвищення якості доставки для вузлів, починаючи з найбільш важливого. Прибрати зі списку найважливіші вузли, для яких ціна доставки задовольняє потребам. Якщо такий список непустий, підвищити якість для вузлів зі списку, і виконати Крок 1. Повторювати такі дії, доки потрібна якість не задовольнить пороги рівнів середніх цін доставки, які встановлені потребами системи РСД.

Крок 4. Якщо необхідне балансування між рівнем доступності і узгодженості, інтегрувати алгоритм на Рис. 5 та Рис. 6 в поточні методи реплікації та відповіді на запит на читання, встановивши наперед необхідний поріг – кількість вузлів, які мають бути доступними у будь-якому разі.

**Четвертий розділ** містить опис програмних моделей, які реалізують комплекс експериментальних досліджень.

Метою першого експериментального дослідження є визначення рівня узгодженості і перевірка метрики для оцінки рівня узгодженості у системі РСД (див. Твердження 3). Кожний експеримент проводиться на певній кількості узгоджених розділень вузлів (підмножини вузлів, в якій репліки на вузлах співпадають), тобто, для  $n$  вузлів на кожному з експериментів беруться  $m$  розділень, і роздивляються різні комбінації для таких розділень. За цими розділеннями експеримент проводиться таким чином: з множини вузлів з кількістю  $n$  беруться випадково два вузли і проводиться тест на випадок, чи є вони елементами одного розділення чи різних, і в першому випадку вони вважаються узгодженими. Таким чином, ця операція повторюється для кожного розділення. Множина операцій для всіх розділень називається ітерацією, для кожного експерименту проведено 100 ітерацій на кількості вузлів від 5 до 100. Шляхом багатьох повторень такої операції над випадковою подією експеримент доводить правильність формули для метрики узгодженості і дозволяє побудувати для відображення графіки, які відображають рівень узгодженості.

Метою наступного експериментального дослідження є оцінка часу, за який процес реплікації в РСД відновить узгодженість в системі за умов відсутності втрат даних. Програмна модель для експерименту симулює розповсюдження однієї репліки сховищем і вимірює ціну розповсюдження всім сховищем, залежно від

ціни кожного зв'язку, що є елементом шляхів, якими мандрує репліка. Ціна кожного зв'язку («link cost») – це величина, виміряна заздалегідь за допомогою емпіричних оцінок характеристик мережі, термін, що впливає на час, який знадобиться для відправлення повідомлення даним зв'язком. Отже, шляхом проведення великої кількості експериментів, симулюючи поширення одного повідомлення сховищем, показано, що ціна розповсюдження репліки буде не більше за діаметр за умови відсутності втрат даних і дозволяє перевірити правильність Твердження 2.

Згодом, метою останнього експериментального дослідження є оцінка ефективності методів контролю узгодженості та доступності, що подані алгоритмами на Рис. 5 та Рис. 6. Реалізовано декілька симуляційних моделей архітектури РСД з різними комбінаціями застосування декількох технологій, таких як балансувальник навантаження, рішення для глобального словнику і т.ін. Розглядаються ці рішення у контексті задач, які можна покласти на те чи інше рішення, що дало можливість зробити порівняльну характеристику цих архітектур, оцінити ризики і вибрати більш раціональне. Детальніше програмна модель для цього дослідження є у додатках. У результаті експеримент відображає ефективність методів контролю узгодженості та доступності шляхом проведення багатьох симуляцій, які реалізовано за допомогою програмних моделей РСД; дозволяє виконати заміри витраченого протягом симуляції часу та пам'яті.

У додатках подано схеми програмних моделей для експериментів, результати досліджень у вигляді графіків, діаграми послідовності та діаграми компонентів для експериментів.

## ВИСНОВКИ

У дисертаційній роботі поставлена та вирішена актуальна науково-практична задача забезпечення контрольованого рівня узгодженості реплік у РСД на етапі його проектування, яка складається з моделей, методів та інструментальних засобів для запобігання ризиків, що можуть з'явитися як наслідок неоптимального проектування архітектури системи РСД. Внаслідок виконаних досліджень отримано такі наукові та практичні результати:

- досліджено існуючі технології та протоколи реплікації, які забезпечують виконання CAP-вимог (CA, AP, CP), що дозволило запропонувати статистичний підхід до задачі, що розглядається;
- побудовано математичні і імітаційні моделі процесів реплікації даних у РСД, що дозволило формалізувати їх основні операції з даними і властивості;
- формалізовано статистичні метрики для оцінки рівня виконання кожної з CAP-вимог, що дозволило побудувати модель для оцінки рівня узгодженості реплік у РСД з мережевими розділеннями, втратами даних при передачі і можливими збоями вузлів;
- розроблено методи управління рівнем узгодженості за умови забезпечення заданих рівнів інших CAP-вимог;
- запропоновано і обґрунтовано метод балансування узгоджених вузлів з метою забезпечення заданого рівня доступності і стійкості до розділення на узгоджених вузлах РСД;

- метод балансування доведено до алгоритмічної реалізації;
- отримано теоретичні оцінки часової і просторової складності для методу балансування, які підтверджені результатами імітаційних експериментів;
- на основі практичних результатів, поданих вище, сформульовано загальні рекомендації щодо проектування РСД зі збалансованими рівнями виконання CAP-вимог в залежності від технічного завдання.

Достовірність отриманих результатів забезпечується математично коректними доведеннями основних тверджень, а також шляхом аналізу результатів комп'ютерних експериментів.

Результати роботи використані у освітньому процесі при підготовці фахівців рівня «бакалавр» і «магістр» за спеціальностями 122 – комп'ютерні науки і 126 – інформаційні системи і технології в Запорізькому національному університеті та Херсонському державному університеті.

#### СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

[1] Zholtkevych G., Mallet F., Zaretska I., Zholtkevych G. Two Semantic Models for Clock Relations in the Clock Constraint Specification Language. *Communications in Computer and Information Science*. 2013. Vol. 1000, P. 190–209.

[2] Zholtkevych, G. Metrics for evaluating consistency in distributed datastores. *Сучасний стан наукових досліджень та технологій в промисловості*. 2020. № 2(12). С. 40–48.

[3] Rukkas, K., Zholtkevych, G. Probabilistic model for estimation of CAP-guarantees for distributed datastore. *Сучасні інформаційні системи*. 2020. Т. 4 (№ 2). С. 47–50.

[4] Rukkas, K., Zholtkevych, G. Load balancing consistency in a distributed datastore. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 2020. Т. 2 (№ 60). С. 95–100.

[5] Zaretska, I., Zholtkevych, G., Zholtkevych, G., Mallet, F. Clocks Model for Specification and Analysis of Timing in Real-Time Embedded Systems. *9th International Conference on ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer*, 9–22 June 2013. Vol. 1356, P. 475–489.

[6] Rukkas, K., Zholtkevych, G. Distributed Datastores: Towards Probabilistic Approach for Estimation of Dependability. *11th International Conference on ICT in Education, Research, and Industrial Applications: Integration, Harmonization, and Knowledge Transfer*, 14–16 May 2015. Vol. 1356, P. 523–534.

[7] Rukkas, K., Zholtkevych, G. Distributed Datastores: Consistency Metric for Distributed Datastores. *XV Міжнародна науково-практична конференція «Теоретичні та прикладні аспекти побудови програмних систем» ТААПСД'2017*, 4–8 грудня 2017. С. 89–100.

[8] Zholtkevych, G. PROBABILISTIC MODEL FOR COMPUTING TIME DELIVERY IN DISTRIBUTED DATASTORE NETWORK. *VI міжнародна науково-технічна конференція «Комп'ютерне моделювання та оптимізація складних систем»*, 4–6 листопада 2020 р. С. 33–35.



## АНОТАЦІЯ

Жолткевич Г.Г. Моделювання процесів реплікації даних у розподілених сховищах. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.02 – математичне моделювання та обчислювальні методи. – Інститут проблем машинобудування ім. А.М. Підгорного НАН України, Харків. – 2021.

У дисертаційній роботі вирішується актуальна науково-практична задача наявності загального методу прийняття рішень, на який можливо спиратись, будуючи мережу для розподіленого сховища даних. Це дозволяє не думати про ліміт запитів та можливу кількість кластерів та їх розташування безперервно, а мати конкретні критерії для формування сховища за умов різноманітних потреб конкретної розподіленої системи, та збалансувати CAP-гарантії залежно від таких потреб. Для вирішення цієї задачі розглянуто існуючі методи забезпечення CAP-гарантії та процеси реплікації у розподіленій системі, розглянуто існуючі рішення та протоколи, що застосовуються у забезпеченні CAP-гарантії. Сформовано моделі для розподіленого сховища, що описують дані та їх зберігання, централізований та розподілений підхід, моделі для розподіленої мережі, що є базою для формування стохастичних метрик і подальших досліджень. Логічним продовженням є розвинена стохастична модель для оцінки різних компонентів CAP-гарантії та метрик, які впливають на ці значення. Це впливає на формування загальних рекомендацій для проектування розподіленого сховища зі збалансованими рівнями виконання CAP-гарантії в залежності від наданого технічного завдання; дає розвиток методам управління узгодженістю на потрібному користувачу рівні, і дає змогу контролювати рівні CAP-гарантії. Створено та доведено до алгоритмічної реалізації метод балансування між узгодженими вузлами для забезпечення необхідного рівня узгодженості.

Ключові слова: розподілене сховище даних, CAP-теорема, математичні моделі, стохастичні метрики, метод балансування, алгоритм прийняття рішень.

## ABSTRACT

Zholtkevych G. Modeling data replication processes in distributed datastores. – Manuscript.

Thesis for candidate of technical science on the specialty 01.05.02 – mathematical modeling and computing methods. – A. Pidhornyi Institute of Mechanical Engineering problems of NASU. – 2021.

Actual scientific-applied problem has been posed in the thesis – the problem of existence of general decision-making method, that would be the one of the base components on the stage of distributed datastore design, that allows to make larger size of distributed network and not to think permanently about possible clusters and nodes allocation, but to have specific criteria for building distributed datastore in the conditions of various system requirements and balance CAP-requirements depending on these needs. To solve the problem, existing methods of fulfilling CAP-requirements and distributed database replication processes are investigated. Models for distributed datastore that describe centralized and distributed data storing approach, distributed network were

formed. This is the basis for the further research: it is developed the stochastic model to evaluate different components of CAP-requirements and metrics that impact these values. This gives the reason to form general recommendations for designing distributed data store with balanced degrees of CAP-guarantees depending on the given business and system requirements; gives the ability to evolve methods of managing required degree consistency value (one of CAP-guarantees) on the needed level and gives possibility to monitor CAP-guarantees. Also, for that the method for balancing between consistent nodes is created and the algorithm is implemented.

Keywords: distributed data store, CAP-theorem, mathematical models, stochastic metrics, balancing method, decision-making algorithm.

#### АННОТАЦИЯ

Жолткевич Г.Г. Моделирование процессов репликации данных в распределенных хранилищах. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 01.05.02 – математическое моделирование и вычислительные методы. – Институт проблем машиностроения им. А.Н. Подгорного НАН Украины, Харьков. – 2021.

В диссертационной работе решается актуальная научно-практическая задача существования общего метода принятия решений, на который можно опираться на этапе построения и дизайна распределенного хранилища данных. Это позволяет не задумываться постоянно об ограничениях на запросы и расположение кластеров и узлов в распределенной сети, а иметь четко поставленные критерии для формирования хранилища в условиях разнообразных требований конкретной реализации распределенной системы, и сбалансировать CAP-гарантии независимо от таких требований. Для решения этой задачи рассмотрены существующие методы обеспечения CAP-гарантий и процессы репликации в распределенной системе, исследованы существующие решения и протоколы, применяемые в обеспечении CAP-гарантий. Сформированы модели для распределенного хранилища, которые описывают данные и их хранение, централизованный и распределенный подход, модели для распределенной сети, что является базой для формирования стохастических метрик и дальнейших исследований. Логическим продолжением является развитие стохастической модели для оценки разных компонентов CAP-гарантий и метрик, которые влияют на эти значения. Это влияет на формирование общих рекомендаций для проектирования распределенного хранилища со сбалансированными порогами выполнения CAP-гарантий в зависимости от технического задания; дает развитие методам управления согласованностью на нужном пользователю уровне и дает возможность контроля уровня, на котором выполняются CAP-гарантии. Создан и доведен до алгоритмической реализации метод балансировки между согласованными узлами для обеспечения необходимого уровня согласованности.

Ключевые слова: распределенное хранилище данных, CAP-теорема, математические модели, стохастические метрики, метод балансирования, алгоритм принятия решений.