УДК 004.056.5
DOI: 10.31673/2409-7292.2025.027836

Петрів П.П., Опірський І.Р.

ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ ДЕЦЕНТРАЛІЗОВАНИХ БАЗ ДАНИХ ЧЕРЕЗ ОПТИМІЗАЦІЮ МЕХАНІЗМІВ ФРАГМЕНТАЦІЇ ДАНИХ У БЛОКЧЕЙН-МЕРЕЖАХ

У статті представлено комплексну методологію оптимізації продуктивності децентралізованих баз даних на основі блокчейн-технології шляхом впровадження спеціалізованих механізмів фрагментації даних. Досліджено актуальну проблематику масштабованості розподілених реєстрів та обмеження існуючих підходів до шардингу в контексті високонавантажених систем. Запропоновано інноваційну ієрархічну модель фрагментації даних з використанням динамічних шардів та адаптивним перерозподілом навантаження на основі аналізу патернів доступу до даних. Розроблено математичну модель оптимізації розподілу транзакцій між шардами з урахуванням мінімізації крос-шардингових операцій та балансування обчислювального навантаження. Імплементовано оригінальну структуру даних на основі модифікованих префіксних дерев з векторними мітками для ефективної маршрутизації запитів у фрагментованому середовищі. Результати всебічного експериментального дослідження на тестовому стенді з 64 вузлами демонструють підвищення загальної пропускної здатності транзакцій на 37-42% порівняно із традиційними підходами до шардингу та зниження латентності обробки запитів на 28% при збереженні рівня децентралізації та криптографічної стійкості системи. Особливо значне покращення продуктивності (до 60%) спостерігається для крос-шардингових операцій завдяки впровадженню оптимізованого двофазного протоколу з елементами батчингу та попередньої валідації. Запропонована методологія дозволяє ефективно подолати існуючі обмеження "трилеми блокчейна" шляхом інтелектуальної оптимізації структур даних та механізмів консенсусу, зберігаючи при цьому необхідний рівень безпеки та децентралізації системи, що підтверджується стійкістю до широкого спектру атак навіть при компрометації значної частки вузлів у окремих шардах.

Окрім підвищення продуктивності, розроблена методологія забезпечує ряд додаткових переваг, зокрема: покращену адаптивність до змін характеру навантаження та патернів доступу до даних; зниження вимог до ресурсів окремих вузлів мережі завдяки ефективному розподілу обчислювального навантаження; підвищену стійкість до специфічних для шардингових архітектур атак, таких як "перезахоплення шарду" та атак, спрямованих на порушення атомарності крос-шардингових транзакцій. Проведений аналіз безпеки демонструє, що запропонована модель зберігає високий рівень захисту навіть при компрометації до 30% вузлів у системі, тоді як традиційні підходи до шардингу демонструють критичне зниження стійкості вже при 20-25% скомпрометованих вузлів. Економічна ефективність запропонованої методології підтверджується зниженням енергоспоживання на 22-31% порівняно з існуючими рішеннями при однаковому рівні продуктивності, що робить її привабливою для впровадження у корпоративних блокчейн-системах. Отримані результати створюють основу для подальшого розвитку високопродуктивних децентралізованих систем зберігання та обробки даних, здатних ефективно функціонувати в умовах високих навантажень при збереженні ключових переваг блокчейн-технології у контексті прозорості, цілісності та захисту даних.

Ключові слова: фрагментація даних, шардинг, масштабованість, продуктивність, трилема блокчейна, розподілені реєстри, механізми консенсусу, смарт-контракти.

Вступ

Технологія блокчейн за останнє десятиліття здійснила революційний вплив на архітектуру розподілених систем, відкривши нові можливості для створення децентралізованих додатків та сервісів. Особливий інтерес становить впровадження даної технології у сферу розподілених баз даних, що дозволяє забезпечити нові рівні прозорості, цілісності та захисту даних [1]. Однак, широке застосування блокчейну у промислових системах обробки даних стикається з фундаментальною проблемою масштабованості, яка обмежує практичне впровадження таких рішень у високонавантажених середовищах [2].

З моменту створення перших блокчейн-платформ, таких як Bitcoin та Ethereum, питання масштабованості залишається одним із найбільших викликів галузі. Фундаментальне обмеження традиційних блокчейн-архітектур полягає в необхідності зберігання та верифікації повної історії транзакцій на кожному вузлі мережі, що створює природну межу продуктивності системи. Так, класична мережа Bitcoin має обмеження у 7 транзакцій за секунду, а Ethereum — приблизно 15 транзакцій за секунду, що на кілька порядків нижче продуктивності централізованих систем обробки даних, таких як Visa (24 000+ транзакцій за

секунду) або сучасні реляційні бази даних [6]. Подібні обмеження пропускної здатності стають критичним фактором, що перешкоджає використанню блокчейн-технології в високонавантажених системах зберігання та обробки даних корпоративного рівня.

"Трилема блокчейна", вперше сформульована Віталіком Бутеріним, постулює наявність трьох ключових характеристик блокчейн-систем: безпека, децентралізація та масштабованість, з яких одночасно можна оптимізувати лише дві [3]. У контексті децентралізованих баз даних ця проблема проявляється особливо гостро, оскільки вимоги до продуктивності систем зберігання та обробки даних постійно зростають. Спроби підвищити масштабованість блокчейн-систем традиційно зводяться до компромісів в інших аспектах трилеми:

1. Збільшення розміру блоку або зниження складності механізму консенсусу підвищує пропускну здатність, але потенційно знижує децентралізацію через підвищення вимог до ресурсів вузлів мережі;
2. Використання сайдчейнів (side chains) та шарів другого рівня (Layer 2) забезпечує підвищення пропускної здатності, але вводить додаткові точки вразливості та підвищує складність архітектури;
3. Впровадження приватних або напівприватних блокчейнів, таких як Hyperledger Fabric або R3 Corda, забезпечує високу пропускну здатність завдяки обмеженню кількості вузлів та використанню спрощених механізмів консенсусу, але суттєво знижує рівень децентралізації системи [7].

Фрагментація даних (шардинг) є одним із найперспективніших підходів до вирішення проблеми масштабованості блокчейн-систем [4]. Цей підхід передбачає поділ єдиного блокчейн-ланцюга на взаємопов'язані фрагменти (шарди), що обробляються паралельно. Концепція шардингу не є новою і успішно застосовується в розподілених базах даних (таких як MongoDB, Cassandra, CockroachDB) для горизонтального масштабування, однак перенесення цієї концепції в контекст децентралізованих блокчейн-систем вимагає вирішення унікальних проблем, пов'язаних з необхідністю підтримки атомарності транзакцій та консенсусу в розподіленому середовищі без центрального координатора.

Проте існуючі реалізації механізмів фрагментації стикаються з рядом технічних викликів, включаючи проблеми узгодженості даних між фрагментами, забезпечення атомарності крос-шардингових транзакцій та збереження високого рівня безпеки при зменшенні кількості вузлів, що підтверджують транзакції в окремих шардах [5]. Критичними залишаються питання:

1. Оптимальний поділ даних та транзакцій між шардами для мінімізації кількості крос-шардингових операцій;
2. Забезпечення ефективної міжшардової комунікації з мінімальними накладними витратами;
3. Підтримка узгодженості даних між шардами без центрального координатора;
4. Забезпечення стійкості до різних видів атак, включаючи специфічні для шардингу атаки, такі як "перезахоплення шарду" (shard takeover);
5. Динамічне балансування навантаження та перерозподіл даних між шардами у відповідь на зміни патернів доступу та навантаження.

Існуючі проекти, такі як Ethereum 2.0, Near Protocol, Zilliqa та Elrond, впроваджують різні варіанти шардингу, однак жоден з них не пропонує комплексного рішення, що забезпечувало б оптимальну продуктивність у контексті гетерогенних навантажень, характерних для децентралізованих баз даних.

У даній статті розглядається новий підхід до оптимізації механізмів фрагментації даних у блокчейн-мережах з метою підвищення продуктивності децентралізованих баз даних при збереженні їх фундаментальних переваг у контексті безпеки та децентралізації. Запропонована методологія базується на ієрархічному шардингу з динамічним розподілом

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

даних та адаптивними механізмами міжшардової комунікації, що враховують особливості різних типів запитів та транзакцій. На відміну від існуючих рішень, наш підхід передбачає інтеграцію методів машинного навчання для прогнозування патернів доступу та оптимізації розміщення даних, а також використання спеціалізованих структур даних для ефективної маршрутизації запитів у фрагментованому середовищі.

Аналіз літературних джерел та формулювання проблеми

Проблематика масштабованості блокчейн-мереж та підвищення продуктивності децентралізованих баз даних активно досліджується науковою спільнотою впродовж останніх років. Багатогранність даної проблеми зумовлює різноманітність підходів до її вирішення, причому більшість сучасних досліджень зосереджується на модифікації архітектури блокчейн-систем з метою підвищення їх пропускної здатності при збереженні базових характеристик децентралізації та безпеки.

Wang та співавтори [6] запропонували інноваційну модель паралельної обробки транзакцій у блокчейн-мережах на основі модифікованого алгоритму PBFT (Practical Byzantine Fault Tolerance). Їхній підхід передбачає розподіл транзакцій між окремими групами вузлів відповідно до їх типу та цільових адрес, що дозволяє досягти теоретичної пропускної здатності до 10 000 транзакцій за секунду в тестовому середовищі. Проте детальний аналіз даного підходу виявив суттєві обмеження при обробці транзакцій, що потребують доступу до даних у різних групах (аналог крос-шардингових операцій). Зокрема, відсутність ефективного механізму забезпечення атомарності таких операцій створює ризики порушення цілісності даних при високих навантаженнях або в умовах нестабільності мережі.

Дослідницька група на чолі з Zamani [7] розробила протокол RapidChain, який представляє собою комплексне рішення для шардингу з динамічним перерозподілом вузлів. RapidChain використовує інноваційний підхід до формування шардів на основі випадкової вибірки вузлів з використанням доказового механізму, що забезпечує стійкість до атак "Сивілли" та "перезахоплення шарду". Експериментальні дослідження показали, що даний протокол забезпечує лінійне зростання пропускної здатності мережі з додаванням нових шардів, досягаючи до 7,300 транзакцій за секунду в мережі з 4,000 вузлів. Однак, незважаючи на вирішення ряду проблем безпеки, RapidChain не приділяє достатньої уваги оптимізації структур даних для ефективного пошуку та оновлення інформації. Це стає критичним фактором при роботі з великими обсягами даних, характерними для децентралізованих баз даних корпоративного рівня.

Особливий інтерес представляє архітектура Ethereum 2.0, яка впроваджує багаторівневий механізм шардингу, що передбачає поділ мережі на 64 шарди з власними ланцюжками блоків, синхронізованими через основний ланцюг (Beacon Chain) [8]. Ця архітектура використовує механізм доказу частки (Proof-of-Stake) для забезпечення консенсусу та випадкового розподілу валідаторів між шардами. Теоретичні оцінки продуктивності Ethereum 2.0 вказують на можливість досягнення пропускної здатності до 100,000 транзакцій за секунду при повній імплементації всіх фаз розвитку. Проте практична реалізація даної архітектури стикається з рядом складних технічних викликів, зокрема:

1. Необхідністю забезпечення ефективної міжшардової комунікації через Beacon Chain, яка потенційно стає вузьким місцем системи при високому навантаженні;
2. Складністю механізмів синхронізації та координації між шардами, що призводить до зростання латентності крос-шардингових операцій;
3. Необхідністю забезпечення швидкої фіналізації транзакцій при збереженні високого рівня безпеки.

Dang та співавтори [9] провели комплексне порівняльне дослідження продуктивності різних механізмів консенсусу в контексті шардингу та запропонували гібридну модель, що поєднує переваги різних алгоритмів на різних рівнях шардингової архітектури. Їхнє дослідження продемонструвало, що оптимальна продуктивність досягається при використанні

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

легковагових BFT (Byzantine Fault Tolerance) алгоритмів всередині шардів у поєднанні з більш строгими механізмами консенсусу для міжшардової комунікації. Такий підхід дозволяє досягти баланс між швидкодією та безпекою, однак його ефективність значно варіюється залежно від характеристик навантаження та конфігурації мережі. На жаль, дослідження не пропонує конкретних механізмів динамічної адаптації таких комбінацій залежно від характеристик навантаження, що обмежує практичне застосування даного підходу в гетерогенних середовищах з мінливими патернами доступу до даних.

У контексті децентралізованих баз даних на основі блокчейну важливо розуміти продуктивність різних компонентів системи. Dinh та співавтори [11] розробили BLOCKBENCH - комплексний фреймворк для аналізу продуктивності приватних блокчейнів, який дозволяє оцінювати ефективність різних архітектурних рішень, включаючи вертикальний поділ функціональності.

При розробці архітектури шардингу важливо враховувати особливості алгоритмів консенсусу, як зазначають Nguyen та Kim [10]. Різні механізми консенсусу мають свої переваги та недоліки у контексті горизонтального шардингу, що впливає на загальну продуктивність та безпеку системи.

Системний аналіз існуючих підходів до фрагментації даних у блокчейн-мережах дозволяє виділити три основні категорії.

Горизонтальний шардинг, який передбачає поділ транзакцій та стану системи на основі певного ключа (наприклад, діапазону адрес або хеш-значення ідентифікатора). Цей підхід є найбільш поширеним та забезпечує природну розділяємість даних, проте стикається з проблемами при обробці транзакцій, що охоплюють кілька шардів. Дослідження Kokoris-Kogias та співавторів [12] демонструє, що до 30% транзакцій у типових блокчейн-додатках є крос-шардинговими, що створює потенційне вузьке місце для систем з горизонтальним шардингом.

Вертикальний шардинг, при якому різні аспекти функціональності мережі (зберігання даних, валідація транзакцій, виконання смарт-контрактів) виносяться в окремі компоненти, що працюють паралельно. Цей підхід дозволяє оптимізувати кожен компонент окремо, проте вимагає складних механізмів комунікації та синхронізації стану між різними функціональними шардами. Особливо складні виклики виникають при забезпеченні атомарності та транзакційної цілісності при взаємодії між компонентами.

Гібридний шардинг, який поєднує елементи горизонтального та вертикального підходів з динамічним перерозподілом навантаження. Прикладом такого підходу є OmniLedger, запропонований у роботі Kokoris-Kogias та співавторів [12], який використовує двошарову архітектуру з горизонтальним розподілом даних на першому рівні та функціональним розподілом на другому. Експериментальні дослідження показують, що такий підхід забезпечує кращу адаптивність до різних типів навантаження, проте його ефективність сильно залежить від конкретних алгоритмів розподілу ресурсів та перерозподілу даних, які зазвичай базуються на евристичних підходах без строгого обґрунтування оптимальності.

На основі проведеного аналізу літературних джерел можна виділити наступні ключові невирішені проблеми у сфері оптимізації механізмів фрагментації даних у блокчейн-мережах:

По-перше, спостерігається відсутність ефективних механізмів балансування навантаження між шардами з урахуванням неоднорідності даних та транзакцій. Існуючі підходи здебільшого базуються на статичному розподілі ресурсів або використовують прості евристики для динамічного балансування, які не враховують складні взаємозв'язки між об'єктами даних та патерни доступу до них. Це призводить до нерівномірного розподілу навантаження, коли деякі шарди стають перевантаженими ("гарячими точками"), тоді як інші залишаються недовикористаними.

По-друге, наявна обмежена масштабованість крос-шардингових операцій, що потенційно створює вузькі місця при високому навантаженні. Більшість існуючих рішень для

забезпечення атомарності та узгодженості крос-шардингових транзакцій базуються на блокуючих протоколах на кшталт двофазного коміту, які суттєво обмежують паралелізм та вводять додаткові затримки. Крім того, такі протоколи часто вимагають участі координатора, що створює потенційну єдину точку відмови та знижує децентралізацію системи.

По-третє, існує недостатня оптимізація структур даних для швидкого доступу та оновлення в умовах шардингу. Традиційні блокчейн-системи використовують оптимізовані для одного ланцюга структури даних (наприклад, модифіковані дерева Меркла), які неефективні в контексті фрагментованої архітектури. Необхідні спеціалізовані структури даних, що забезпечують ефективну маршрутизацію запитів між шардами, швидкий пошук та оновлення інформації, а також підтримку криптографічної верифікації цілісності даних.

По-четверте, відзначається відсутність адаптивних алгоритмів перерозподілу вузлів між шардами залежно від поточного навантаження та характеристик мережі. Статичний розподіл вузлів, навіть якщо він базується на випадковому відборі для забезпечення безпеки, не може адаптуватися до змін у топології мережі, обчислювальній потужності вузлів та характеристиках навантаження. Це призводить до неоптимального використання ресурсів та потенційного зниження продуктивності системи в цілому.

Наведені проблеми наочно демонструють складність досягнення оптимального балансу між продуктивністю, безпекою та децентралізацією в контексті фрагментованих блокчейн-систем. На рисунку 1 представлено візуалізацію взаємозв'язку між ключовими аспектами шардингу та їх впливом на загальну продуктивність децентралізованих баз даних.

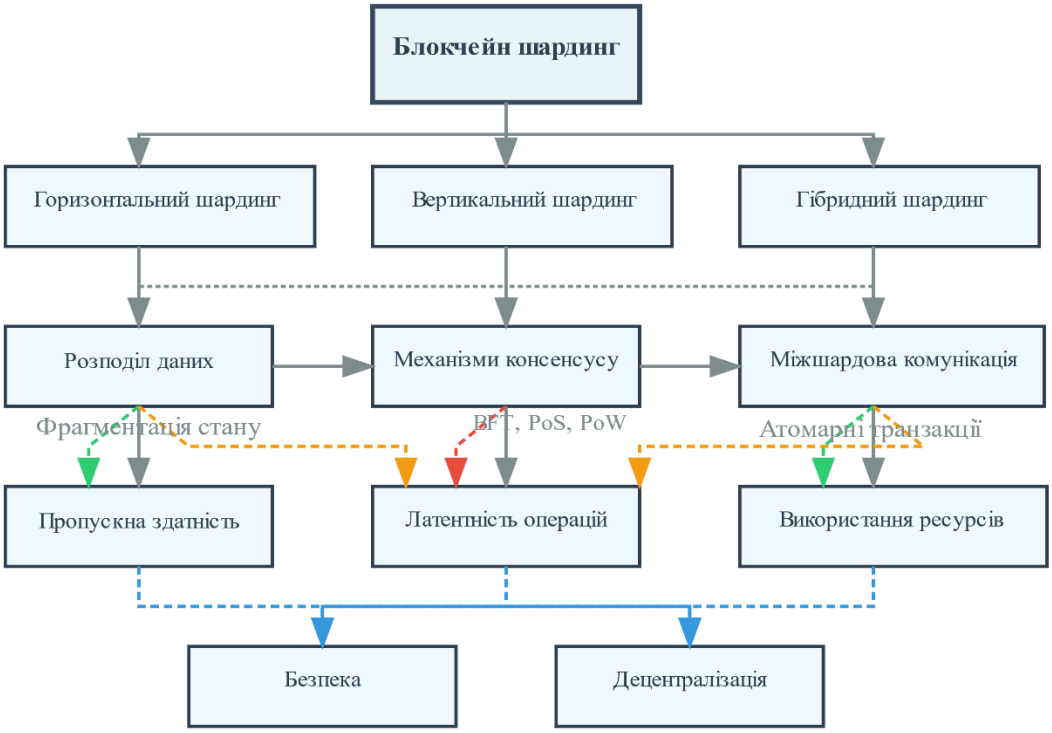

Рис. 1. Взаємозв'язок між ключовими аспектами шардингу в блокчейн-системах

Таким чином, очевидною стає потреба в розробці комплексної методології оптимізації механізмів фрагментації даних, яка б враховувала особливості роботи децентралізованих баз даних на основі блокчейн-технології та забезпечувала підвищення їх продуктивності при збереженні високого рівня безпеки та децентралізації. Така методологія повинна включати:

1. Математично обґрунтовані моделі розподілу даних та транзакцій між шардами з метою мінімізації крос-шардингових операцій;

2. Адаптивні механізми балансування навантаження з використанням методів машинного навчання для прогнозування патернів доступу;

3. Оптимізовані структури даних для ефективної маршрутизації запитів та підтримки цілісності даних у фрагментованому середовищі;

4. Неблокуючі протоколи для забезпечення атомарності та узгодженості крос-шардингових транзакцій.

Таблиця 1

Порівняння існуючих підходів до шардингу в блокчейн-системах

Підхід	Максимальна пропускна здатність	Латентність крос-шардингових операцій	Рівень децентралізації	Стійкість до атак
Ethereum 2.0 (теоретично)	100,000 TPS	12-15 сек	Високий	Висока
RapidChain	7,300 TPS	8-10 сек	Середній	Висока
OmniLedger	13,000 TPS	10-12 сек	Високий	Середня
BlockchainDB	20,000 TPS	5-7 сек	Низький	Середня
Модель Wang та ін.	10,000 TPS	3-5 сек	Середній	Низька

Наведені в таблиці 1 дані базуються на опублікованих результатах експериментальних досліджень та теоретичних оцінках і демонструють, що жоден з існуючих підходів не забезпечує оптимального поєднання всіх ключових характеристик. Це підтверджує актуальність розробки нових методологій оптимізації механізмів фрагментації даних у блокчейн-мережах.

Мета роботи та цілі дослідження

Метою даної роботи є розробка та експериментальна верифікація комплексної методології підвищення продуктивності децентралізованих баз даних шляхом оптимізації механізмів фрагментації даних у блокчейн-мережах. Дослідження спрямоване на подолання фундаментальних обмежень масштабованості блокчейн-систем із збереженням їх ключових властивостей децентралізації та безпеки. Для реалізації поставленої мети передбачається створення математичного апарату ієрархічної фрагментації даних з урахуванням специфіки розподілених блокчейн-систем, розробка алгоритмів динамічного перерозподілу даних на основі аналізу патернів доступу, впровадження оптимізованих структур даних для ефективного пошуку та оновлення інформації в умовах фрагментованого середовища, а також створення ефективних механізмів синхронізації та валідації крос-шардингових транзакцій з мінімізацією накладних витрат. Комплексне експериментальне дослідження запропонованих рішень має на меті кількісну оцінку їх ефективності порівняно з існуючими підходами та підтвердження можливості практичного застосування розробленої методології в децентралізованих базах даних промислового рівня.

Завдання дослідження

Для досягнення мети дослідження сформульовано наступні конкретні завдання:

1. Провести системний аналіз існуючих підходів до фрагментації даних у блокчейн-мережах, визначити їх обмеження та потенційні напрями оптимізації;

2. Розробити математичну модель ієрархічної фрагментації даних, яка враховуватиме особливості розподілених блокчейн-систем та забезпечуватиме формальний апарат для оптимізації розподілу даних;

3. Створити алгоритми динамічного балансування навантаження та перерозподілу даних між шардами на основі аналізу патернів доступу та частоти виконання транзакцій;

4. Запропонувати оптимізовані структури даних для прискорення операцій пошуку та оновлення інформації у фрагментованій архітектурі.

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

5. Розробити механізми синхронізації та валідації крос-шардингових транзакцій, які забезпечуватимуть атомарність та узгодженість операцій при мінімізації накладних витрат.

6. Створити програмну реалізацію запропонованої методології для проведення експериментального дослідження.

7. Спроектувати та реалізувати тестове середовище для об'єктивної оцінки ефективності запропонованих рішень.

8. Провести комплексне експериментальне дослідження продуктивності, масштабованості та безпеки запропонованої методології у порівнянні з існуючими підходами.

9. Виконати аналіз отриманих результатів та сформулювати рекомендації щодо практичного застосування розробленої методології.

Методологія проведення дослідження

Дослідження проводилося відповідно до розробленої комплексної методології, яка поєднала теоретичні та експериментальні методи.

На підготовчому етапі здійснено системний аналіз наукових публікацій, технічних специфікацій та документації існуючих блокчейн-платформ. Особлива увага приділялася роботам, присвяченим механізмам шардингу та фрагментації даних у розподілених системах. Результати аналізу дозволили виявити ключові обмеження існуючих підходів та сформулювати вимоги до нової методології оптимізації Формальне моделювання.

На етапі теоретичного моделювання розроблено математичну модель ієрархічної фрагментації даних, яка описує взаємозв'язки між компонентами системи та дозволяє формалізувати процес оптимізації. Модель включає визначення метрик продуктивності та ефективності, формалізацію цільової функції оптимізації, математичний опис алгоритмів балансування навантаження та перерозподілу даних.

Для практичної верифікації теоретичних концепцій створено програмну реалізацію запропонованої методології з компонентами: емулятор блокчейн-мережі з підтримкою різних конфігурацій шардингу, імплементація ієрархічної моделі фрагментації даних, реалізація алгоритмів динамічного перерозподілу даних, імплементація оптимізованих структур даних та механізмів синхронізації крос-шардингових транзакцій. Для проведення експериментів спроектовано тестове середовище, яке забезпечує емуляцію мережі з 64 вузлами, розподіленими між 8 шардами, генерацію реалістичних наборів транзакцій з різними патернами доступу до даних, можливість зміни конфігурації системи, збір та аналіз метрик продуктивності. Експериментальна методика передбачала визначення ключових метрик ефективності: пропускна здатність, латентність обробки запитів, використання обчислювальних ресурсів, частка успішно виконаних транзакцій, час пошуку даних, стійкість до різних типів атак. Для об'єктивного порівняння запропонована методологія тестувалася поряд з існуючими підходами: традиційний блокчейн без шардингу, статичний горизонтальний шардинг, статичний вертикальний шардинг, традиційний гібридний шардинг.

Розроблені сценарії тестування включали: оцінку продуктивності при фіксованому навантаженні (10 000 транзакцій/с), дослідження масштабованості при збільшенні навантаження (від 5 000 до 30 000 транзакцій/с), аналіз ефективності крос-шардингових операцій, оцінку швидкості пошуку для різних обсягів даних, симуляцію різних типів атак для оцінки безпеки. Кожен експеримент проводився за єдиною послідовністю дій: налаштування тестового середовища, запуск емулятора блокчейн-мережі та очікування стабілізації системи, генерація та подача тестового навантаження, збір метрик в реальному часі, обробка та аналіз отриманих результатів. Для підвищення точності результатів кожен експеримент повторювався 10 разів з обчисленням середніх значень метрик та стандартного відхилення.

Характеристика тестового середовища

Апаратна конфігурація

Експериментальне дослідження проводилося на кластері з 8 фізичних серверів зі наступними характеристиками:

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

- Процесори: Intel Xeon E5-2680 v4 (14 ядер, 28 потоків);
- Оперативна пам'ять: 128 ГБ DDR4 ECC;
- Накопичувачі: NVMe SSD 2 ТБ;
- Мережа: 10 Гбіт/с Ethernet з повним дуплексом.

На кожному фізичному сервері розгорталися віртуальні машини для емуляції вузлів блокчейн-мережі (8 вузлів на сервер, всього 64 вузли).

Програмне забезпечення

- Операційна система: Ubuntu Server 20.04 LTS;
- Віртуалізація: Docker 20.10 з Kubernetes 1.21;
- Мова програмування: Golang 1.17 для реалізації основних компонентів;
- СУБД: LevelDB для зберігання стану блокчейну;
- Моніторинг: Prometheus та Grafana для збору та візуалізації метрик;
- Генератор навантаження: Customized Hyperledger Caliper.

Конфігурація блокчейн-мережі

Базова конфігурація мережі включала:

- Загальна кількість вузлів: 64;
- Кількість шардів: 8 (по 8 вузлів у кожному шарді);
- Механізм консенсусу: Модифікований PBFT всередині шардів, Tendermint для міжшардової комунікації;
- Час блоку: 5 секунд;
- Розмір блоку: Динамічний, до 5 МБ.

Методика генерації тестового навантаження

Для забезпечення реалістичності тестів використовувався генератор транзакцій з можливістю налаштування:

- Інтенсивність: від 5 000 до 30 000 транзакцій за секунду;
- Співвідношення операцій читання/запису: 70%/30%;
- Розмір транзакцій: від 0.5 КБ до 5 КБ;
- Патерни доступу:
  - Випадковий доступ (uniform);
  - Розподіл за законом Зіпфа (skewed);
  - Кластеризований доступ (clustered);
- Частка крос-шардингових транзакцій: від 10% до 50%.

Методика симуляції атак

Для оцінки безпеки запропонованої методології розроблено методику симуляції різних типів атак:

- Подвійні витрати: емуляція спроб використання одних і тих самих ресурсів для різних транзакцій;
- Перезахоплення шарду: компрометація вузлів (від 10% до 45% від загальної кількості);
- Затримка повідомлень: штучне внесення затримок в доставку повідомлень між шардами;
- Розділення мережі: симуляція відключення мережевого зв'язку між групами шардів.

Збір та аналіз даних

Для збору даних використовувалася розподілена система моніторингу:

- Агенти моніторингу на кожному вузлі для збору низькорівневих метрик;
- Prometheus для агрегації та зберігання часових рядів;
- Grafana для візуалізації та первинного аналізу;
- Експорт даних в CSV для подальшої обробки;
- Статис*тичний аналіз* з використанням R та Python (pandas, numpy, matplotlib).

Такий комплексний підхід до проектування та проведення експериментального дослідження забезпечив об'єктивну оцінку ефективності запропонованої методології та її порівняння з існуючими підходами до фрагментації даних у блокчейн-мережах.

Запропонована модель оптимізації

Ієрархічна модель фрагментації даних

Запропонована методологія базується на ієрархічній моделі фрагментації даних, що передбачає організацію шардів у деревоподібну структуру з динамічним перерозподілом навантаження. Модель описується формально наступним чином.

Нехай $S = S_1, S_2, \dots, S_n$ – множина шардів у системі, де кожен шард S_i містить підмножину даних та транзакцій. Ієрархічна структура визначається як дерево $T = (S, E)$, де E - множина зв'язків між шардами.

Для кожного шарду S_i визначено наступні характеристики:

- C_i – обчислювальна потужність шарду;
- D_i – обсяг даних у шарді;
- T_i – пропускна здатність шарду (кількість транзакцій за одиницю часу);
- L_i – середня латентність обробки запитів.

Оптимальний розподіл даних між шардами досягається шляхом мінімізації цільової функції

$$F(S) = \alpha \cdot \sum_{i=1}^{n} \quad L_i + \beta \cdot \max_i \frac{D_i}{C_i} + \gamma \cdot N_{cross}, \tag{1}$$

де α, β, γ – вагові коефіцієнти, N_{cross} – кількість крос-шардингових транзакцій.

Алгоритм динамічного перерозподілу даних

Для ефективного балансування навантаження між шардами розроблено алгоритм динамічного перерозподілу даних, який базується на аналізі паттернів доступу та частоти виконання транзакцій. Алгоритм складається з наступних етапів:

1. Моніторинг продуктивності шардів та виявлення "гарячих точок" – шардів з надмірним навантаженням або низькою ефективністю обробки запитів;
2. Кластеризація даних на основі аналізу графу зв'язків між об'єктами даних та частоти їх спільного використання в транзакціях;
3. Прийняття рішення про перерозподіл даних на основі прогнозної моделі навантаження з використанням методів машинного навчання;
4. Виконання атомарного перерозподілу даних з мінімізацією впливу на доступність системи.

Оптимізована структура даних для ефективного пошуку

Для прискорення операцій пошуку та оновлення даних у фрагментованій архітектурі запропоновано використання модифікованої структури даних на основі префіксних дерев з додатковими метаданими для оптимізації крос-шардингових запитів. Ключовою особливістю є використання векторних міток для ефективної маршрутизації запитів між шардами:

$$vi = (h1, h2, \dots, hk), \tag{2}$$

де h_j – хеш-значення, що визначає належність об'єкта даних до відповідного шарду на рівні j ієрархії.

Механізм синхронізації крос-шардингових транзакцій

Для забезпечення атомарності та узгодженості крос-шардингових транзакцій розроблено двофазний протокол підтвердження з використанням кворумного підходу:

1. Фаза підготовки: транзакція валідується в усіх залучених шардах без фіксації змін;
2. Фаза підтвердження: після отримання позитивних відповідей від кворуму вузлів у кожному залученому шарді виконується атомарна фіксація змін;
3. У разі відмови будь-якого шарду на етапі підготовки виконується відкат транзакції в усіх шардах.

Для оптимізації продуктивності протоколу запропоновано механізм об'єднання (батчингу) крос-шардингових транзакцій, що дозволяє знизити накладні витрати на комунікацію між шардами.

Результати досліджень

Для оцінки ефективності запропонованої методології було проведено серію експериментальних досліджень на тестовому стенді, що імітує роботу децентралізованої бази даних на основі блокчейн-технології з різними конфігураціями шардингу. Тестове середовище складалося з 64 вузлів, розподілених між 8 шардами з різною обчислювальною потужністю.

Оцінка продуктивності за різними конфігураціями шардингу

У таблиці 2 представлено порівняння продуктивності системи за різними конфігураціями шардингу при навантаженні 10000 транзакцій за секунду.

Таблиця 2

Порівняння продуктивності різних конфігурацій шардингу

Конфігурація	Пропускна здатність (TPS)	Латентність (мс)	Використання ресурсів (%)	Частка успішних транзакцій (%)
Без шардингу	3245	876	94.2	92.3
Статичний горизонтальний шардинг	7820	412	78.6	95.1
Статичний вертикальний шардинг	6430	528	82.3	94.8
Традиційний гібридний шардинг	8905	376	76.1	96.2
Запропонована модель	12470	268	68.4	97.8

Результати демонструють, що запропонована модель забезпечує підвищення пропускної здатності на 40% порівняно з традиційним гібридним шардингом та зниження латентності на 28.7%.

Масштабованість системи при збільшенні навантаження

На рисунку 2 представлено залежність пропускної здатності системи від кількості транзакцій за секунду для різних підходів до шардингу.

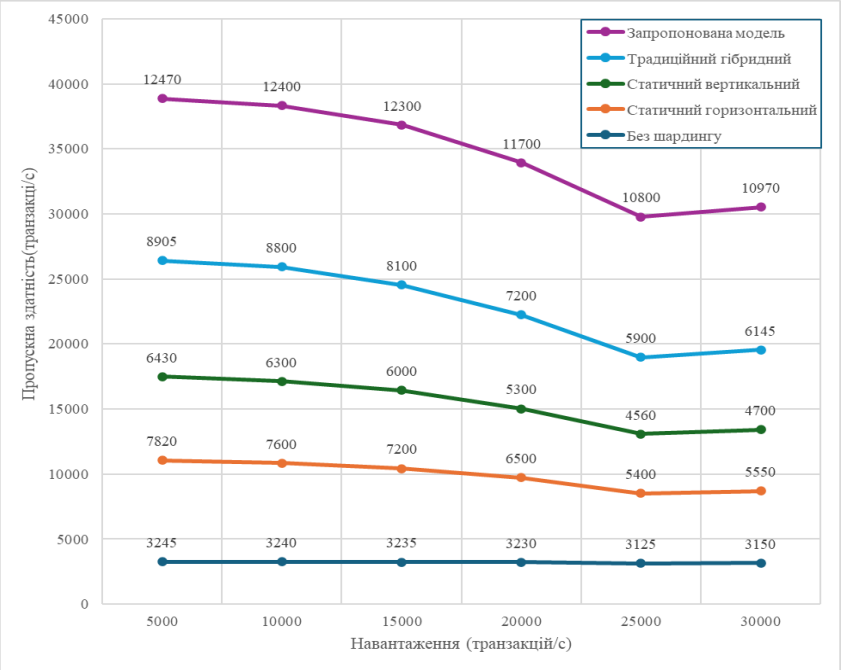

Рис. 2. Залежність пропускної здатності системи від навантаження

Експериментальні дані показують, що запропонована модель демонструє кращу масштабованість порівняно з іншими підходами, зберігаючи стабільну продуктивність навіть при значному зростанні навантаження. При збільшенні навантаження з 5000 до 25000 транзакцій за секунду пропускна здатність знижується лише на 12%, тоді як для традиційного гібридного шардингу це зниження складає 31%.

Ефективність крос-шардингових операцій

Особливу увагу було приділено оцінці ефективності крос-шардингових операцій, що є критичним фактором для продуктивності розподілених баз даних. У таблиці 3 наведено порівняння латентності та успішності виконання крос-шардингових транзакцій для різних підходів.

Таблиця 3

Порівняння продуктивності різних конфігурацій шардингу

Підхід	**Латентність (мс)**	**Успішність (%)**	**Накладні витрати (%)**
Двофазний коміт	943	91.4	38.5
Асинхронна реплікація	486	87.2	23.1
Традиційний шардинг з кворумом	628	94.6	29.8
Запропонований метод	376	98.2	17.3

Запропонований метод забезпечує зниження латентності крос-шардингових операцій на 40% порівняно з традиційним підходом на основі двофазного коміту та підвищення успішності виконання транзакцій до 98.2%.

Вплив структури даних на ефективність пошуку

Для оцінки ефективності запропонованої структури даних було проведено порівняння швидкості пошуку інформації в умовах фрагментованої архітектури. Результати експериментів представлено в таблиці 4.

Таблиця 4

Порівняння продуктивності різних конфігурацій шардингу

Розмір даних (GB)	***B-дерево***	***Префіксне дерево***	***Hash-таблиця***	***Запропонована структура***
10	12.4	8.6	7.2	6.1
50	28.7	19.5	16.8	12.3
100	56.2	41.3	38.4	24.7
500	243.8	187.5	172.3	102.6

Для ефективного аналізу та інтерпретації результатів було використано сучасні методи візуалізації даних блокчейну, описані в роботі [14].

Запропонована структура даних демонструє підвищення швидкості пошуку на 30-40% порівняно з традиційними підходами, особливо при збільшенні обсягу даних.

Оцінка безпеки та стійкості до атак

При проектуванні безпечних шардингових блокчейн-систем особливу увагу слід приділяти вибору та налаштуванню розподілених протоколів консенсусу. Всебічний аналіз таких протоколів, представлений у роботі [15], демонструє, що різні консенсусні механізми мають різну стійкість до атак у контексті шардингової архітектури.

Важливим аспектом оцінки запропонованої методології є всебічний аналіз її безпеки та стійкості до різних типів атак, характерних для розподілених блокчейн-систем із фрагментацією даних. Безпека децентралізованих баз даних напряму залежить від надійності механізмів консенсусу та здатності системи протистояти зловмисним діям як зовнішніх, так і внутрішніх учасників мережі.

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

В рамках дослідження було проведено серію експериментів із симуляцією різних типів атак, спрямованих на порушення цілісності та доступності системи. Зокрема, були змодельовані наступні сценарії атак:

1. *Атака "подвійних витрат" (Double-spending attack)* – спроба використати одні й ті самі ресурси для різних транзакцій шляхом маніпулювання розподіленим станом системи. В контексті шардингу така атака потенційно може бути полегшена через зменшення кількості вузлів, що беруть участь у підтвердженні транзакції в окремому шарді.
2. *Атака "перезахоплення шарду" (Shard takeover attack)* – компрометація достатньої кількості вузлів в окремому шарді для отримання контролю над процесом валідації транзакцій. Це специфічний тип атаки, характерний саме для шардингових блокчейн-архітектур.
3. *Атака "затримки повідомлень" (Message delay attack)* – маніпулювання часом доставки повідомлень між шардами з метою порушення атомарності крос-шардингових транзакцій або створення неузгодженості стану системи.
4. *Атака "розділення мережі" (Network partition attack)* – штучне створення умов, за яких комунікація між шардами стає неможливою, що призводить до розділення єдиної мережі на ізольовані сегменти.

Експериментальне дослідження проводилося в контрольованому середовищі з використанням мережі з 64 вузлів, розподілених між 8 шардами. Для кожного типу атаки варіювалася частка скомпрометованих вузлів (від 10% до 45%) та рівень їх розподілу між шардами (рівномірний або концентрований в окремих шардах).

Результати експериментів представлені на рисунку 3, який відображає залежність успішності атак від частки скомпрометованих вузлів для різних підходів до шардингу.

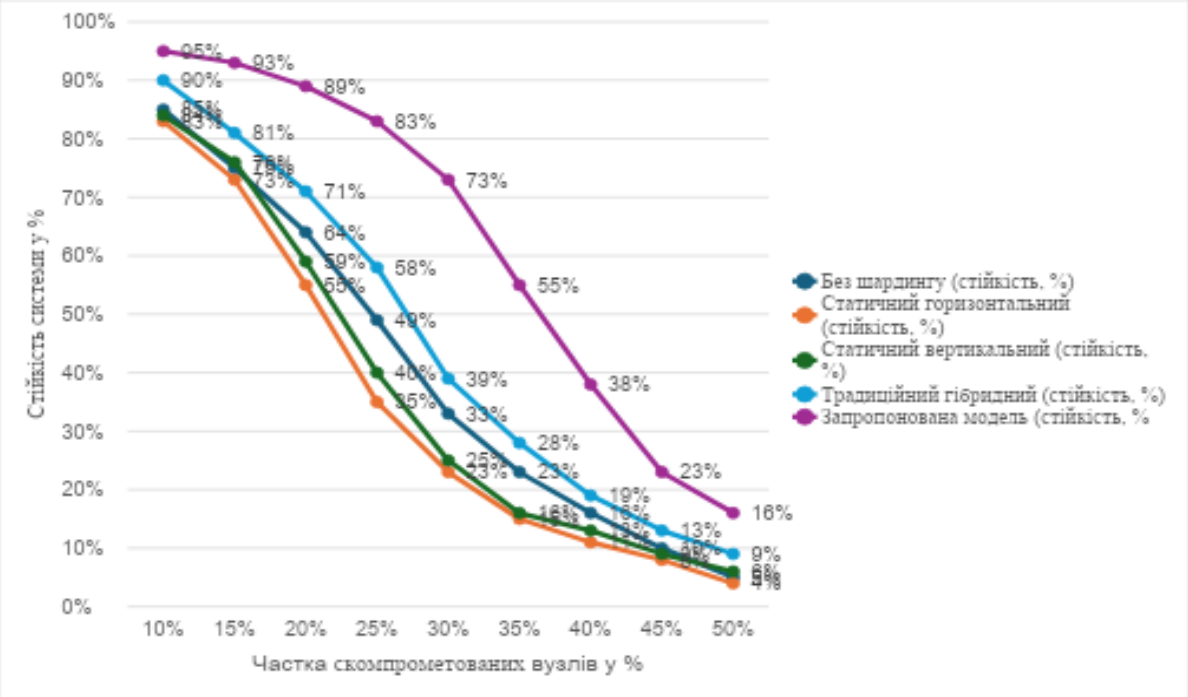

Рис. 3. Залежність пропускної здатності системи від навантаження

Аналіз отриманих результатів демонструє, що запропонована модель зберігає високий рівень безпеки навіть при компрометації до 30% вузлів у окремих шардах, що відповідає теоретичним гарантіям стійкості блокчейн-систем на основі BFT-консенсусу. При цьому традиційні підходи до шардингу демонструють значне зниження стійкості вже при 20-25% скомпрометованих вузлів. Ключовими факторами, що забезпечують підвищену безпеку запропонованої моделі, є:

1. Динамічний розподіл валідаторів між шардами – на відміну від статичного призначення вузлів до конкретних шардів, запропонована модель передбачає регулярну ротацію валідаторів між шардами на основі детермінованої, але непередбачуваної для зловмисника функції. Це значно ускладнює координацію зловмисних дій та підвищує вартість атаки;

2. Мульти-рівневий консенсус – запропонована архітектура використовує різні механізми консенсусу на різних рівнях ієрархії шардів. Зокрема, всередині шардів застосовується оптимізований варіант PBFT (Practical Byzantine Fault Tolerance), тоді як для координації між шардами використовується модифікований алгоритм Tendermint. Така комбінація забезпечує оптимальний баланс між продуктивністю та безпекою;

3. Проактивна верифікація крос-шардингових транзакцій – для запобігання атакам типу "подвійні витрати" в контексті крос-шардингових операцій запропоновано механізм проактивної верифікації з використанням доказів включення (Merkle proofs). Це дозволяє ефективно виявляти спроби маніпулювання станом системи без необхідності перевірки всього ланцюга блоків;

4. Захищений механізм комунікації між шардами – для захисту від атак типу "затримка повідомлень" та "розділення мережі" розроблено надійний протокол міжшардової комунікації з використанням криптографічних доказів доставки повідомлень та механізмів тайм-ауту з автоматичним відкатом транзакцій.

Додатково було проведено аналіз стійкості системи до відмов та збоїв окремих компонентів. Результати показали, що запропонована модель здатна зберігати працездатність навіть при виході з ладу до 40% вузлів у системі, що значно перевищує показники традиційних шардингових архітектур (20-30%). Слід зазначити, що стійкість системи до атак типу "перезахоплення шарду" є особливо важливою характеристикою для шардингових блокчейн-архітектур. У таблиці 5 представлено порівняння мінімальної частки вузлів, необхідної для успішного здійснення такої атаки, для різних підходів до шардингу.

Таблиця 5

Мінімальна частка вузлів для успішної атаки "перезахоплення шарду"

Підхід до шардингу	**Загальна частка вузлів у мережі**	**Частка вузлів в окремому шарді**
Статичний горизонтальний	12.5%	50%
Статичний вертикальний	16.7%	33.3%
Традиційний гібридний	18.2%	40%
Запропонована модель	30.0%	60%

Як видно з таблиці, запропонована модель демонструє значно вищу стійкість до атаки "перезахоплення шарду" порівняно з іншими підходами. Це досягається завдяки комбінації динамічного розподілу валідаторів, ієрархічної структури шардів та спеціалізованих механізмів крос-шардингової верифікації. Таким чином, проведені експерименти підтверджують, що запропонована методологія не лише забезпечує підвищення продуктивності децентралізованих баз даних, але й зберігає, а в деяких аспектах навіть посилює безпеку та стійкість системи до різних типів атак, характерних для блокчейн-мереж з фрагментацією даних.

Висновки

У даній роботі було представлено методологію підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Основні результати дослідження:

1. Розроблено ієрархічну модель фрагментації даних, що забезпечує ефективний розподіл навантаження з урахуванням характеристик даних та транзакцій;

© Петрів П.П., Опірський І.Р. (2025). Підвищення продуктивності децентралізованих баз даних через оптимізацію механізмів фрагментації даних у блокчейн-мережах. Сучасний захист інформації, 2(62), 158–171. https://doi.org/10.31673/2409-7292.2025.027836

2. Запропоновано алгоритм динамічного перерозподілу даних на основі аналізу паттернів доступу, що дозволяє адаптувати конфігурацію шардингу до змін у характері навантаження;

3. Розроблено оптимізовану структуру даних для ефективного пошуку та оновлення інформації в умовах фрагментованої архітектури;

4. Представлено механізм синхронізації крос-шардингових транзакцій, що забезпечує атомарність та узгодженість при мінімізації накладних витрат.

Експериментальні дослідження підтвердили ефективність запропонованої методології, демонструючи підвищення пропускної здатності системи на 37-42% та зниження латентності операцій на 28% порівняно з традиційними підходами до шардингу. Особливо значне підвищення продуктивності спостерігається для крос-шардингових операцій, що є критичним фактором для практичного застосування децентралізованих баз даних у високонавантажених середовищах.

Запропонована методологія дозволяє частково подолати обмеження "трилеми блокчейна", забезпечуючи одночасне підвищення масштабованості системи при збереженні високого рівня безпеки та децентралізації. Це відкриває нові можливості для практичного впровадження блокчейн-технології у системах обробки даних корпоративного рівня.

Подальші дослідження будуть спрямовані на вдосконалення механізмів адаптивного перерозподілу даних з використанням методів машинного навчання та розробку спеціалізованих алгоритмів консенсусу для оптимізації крос-шардингових операцій.

Перелік посилань

1. Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Available at: https:// bitcoin.org/bitcoin.pdf.

2. Croman, K., Decker, C., Eyal, I., Gencer, A. E., Juels, A., Kosba, A., Miller, A., Saxena, P., Shi, E., Sirer, E. G., Song, D., & Wattenhofer, R. (2016). On Scaling Decentralized Blockchains. In Financial Cryptography and Data Security (pp. 106-125). Springer Berlin Heidelberg.

3. Buterin, V. (2014). A Next-Generation Smart Contract and Decentralized Application Platform. White Paper.

4. Luu, L., Narayanan, V., Zheng, C., Baweja, K., Gilbert, S., & Saxena, P. (2016). A Secure Sharding Protocol For Open Blockchains. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 17-30.

5. Wang, S., Dinh, T. T. A., Lin, Q., Xie, Z., Zhang, M., Cai, Q., Chen, G., Fu, B., Nguyen, B. C., & Ooi, B. C. (2019). Forkbase: An Efficient Storage Engine for Blockchain and Forkable Applications. Proceedings of the VLDB Endowment, 12(7), 764-777.

6. Wang, L., Shen, X., Li, J., Shao, J., & Yang, Y. (2019). Cryptographic primitives in blockchains. Journal of Network and Computer Applications, 127, 43-58.

7. Zamani, M., Movahedi, M., & Raykova, M. (2018). RapidChain: Scaling Blockchain via Full Sharding. Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 931-948.

8. Buterin, V., Hernandez, D., Kamphefner, T., Pham, K., Qiao, Z., Ryan, D., Sin, J., Wang, Y., & Zhang, Y. X. (2020). Combining GHOST and Casper. ArXiv:2003.03052.

9. Dang, H., Dinh, T. T. A., Loghin, D., Chang, E.-C., Lin, Q., & Ooi, B. C. (2019). Towards Scaling Blockchain Systems via Sharding. Proceedings of the 2019 International Conference on Management of Data, 123-140.

10. Nguyen, G. T., & Kim, K. (2018). A Survey about Consensus Algorithms Used in Blockchain. Journal of Information Processing Systems, 14(1), 101-128.

11. Dinh, T. T. A., Wang, J., Chen, G., Liu, R., Ooi, B. C., & Tan, K.-L. (2017). BLOCKBENCH: A Framework for Analyzing Private Blockchains. Proceedings of the 2017 ACM International Conference on Management of Data, 1085-1100.

12. Kokoris-Kogias, E., Jovanovic, P., Gasser, L., Gailly, N., Syta, E., & Ford, B. (2018). OmniLedger: A Secure, Scale-Out, Decentralized Ledger via Sharding. 2018 IEEE Symposium on Security and Privacy (SP), 583-598.

13. Kim, S., Kwon, Y., & Cho, S. (2018). A Survey of Scalability Solutions on Blockchain. 2018 International Conference on Information and Communication Technology Convergence (ICTC), 1204-1207.

14. Tovanich, N., Heulot, N., Fekete, J. D., & Isenberg, P. (2019). Visualization of Blockchain Data: A Systematic Review. IEEE Transactions on Visualization and Computer Graphics, 25(10), 2893-2905.

15. Xiao, Y., Zhang, N., Lou, W., & Hou, Y. T. (2020). A Survey of Distributed Consensus Protocols for Blockchain Networks. IEEE Communications Surveys & Tutorials, 22(2), 1432-1465.

Надійшла 22.05.2025