

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Методика проектування та розробки розподілених
систем зберігання даних

(тема)

Виконав:

студент II курсу, групи СПМ-22-3
Білоконь А.С.
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системне програмування
(повна назва освітньої програми)

Керівник: доц. Федорченко В.М.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерної інженерії та управління _____

Кафедра _____ електронних обчислювальних машин _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 123 «Комп'ютерна інженерія» _____
(код і повна назва)

Тип програми _____ освітньо-наукова _____
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Системне програмування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту _____ Білоконю Антону Сергійовичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи Методика проєктування та розробки розподілених систем зберігання даних

затверджена наказом по університету від “ 01 ” квітня 2024 р. № 257 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 15 червня 2024 р.

3. Вхідні дані до роботи _____
методи обробки даних _____

зберігання даних _____

класифікація даних _____

4. Перелік питань, що потрібно опрацювати у роботі _____

Архітектурні рішення побудови систем зберігання даних _____

Параметри та характеристики РСЗ _____

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 16 слайдів

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання та аналіз літератури	01.04.2024 – 06.04.2024	
2	Огляд існуючих рішень та методів	07.04.2024 – 12.04.2024	
3	Розробка методики	13.04.2024 – 18.04.2024	
4	Вибір програмних засобів	19.04.2024 – 25.04.2024	
5	Програмна реалізація	26.04.2024 – 02.05.2024	
6	Аналіз отриманих результатів	03.05.2024 – 16.05.2024	
7	Оформлення записки	17.05.2024 – 14.06.2024	
8	Представлення роботи в ЕК	15.06.2024	

Дата видачі завдання 01 квітня 2024 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

доц. Федорченко В.М.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 58 с., 26 рис., 1 табл., 1 дод., 12 джерел.

РОЗПОДІЛЕНІ СИСТЕМИ ЗБЕРІГАННЯ ДАНИХ, МЕХАНІЗМ, ОБРОБКА ДАНИХ.

Метою кваліфікаційної роботи є аналіз методики проектування та розробки розподілених систем зберігання даних.

У ході виконання кваліфікаційної роботи розглянуто методику проектування розподілених систем для зберігання даних. Розробка розподілених систем зберігання даних є перспективною для сучасних реалій, оскільки вона забезпечує високу надійність, ефективність доступу до даних і масштабованість. Впровадження механізмів реплікації з відповідною стратегією (синхронна, асинхронна, географічно розподілена) забезпечує високу доступність та надійність даних, що є ключовими факторами для досягнення стабільної та ефективної роботи розподілених систем зберігання даних у великих масштабах.

ABSTRACT

Master's thesis: 58 pages, 26 figures, 1 tables, 1 appendices, 12 sources.

DISTRIBUTED DATA STORAGE SYSTEMS, MECHANISM, DATA PROCESSING.

The major goal of this thesis is to analyze the methodology of designing and developing distributed data storage systems.

In order to the qualification work the methodology of designing distributed systems for data storage was considered. The development of distributed data storage systems is promising for modern realities, as it provides high reliability, data access efficiency and scalability. The implementation of replication mechanisms with an appropriate strategy (synchronous, asynchronous, geographically distributed) ensures high availability and reliability of data, which are key factors for achieving stable and efficient operation of distributed data storage systems on a large scale.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1 АРХІТЕКТУРНІ РІШЕННЯ ПОБУДОВИ СИСТЕМ ЗБЕРІГАННЯ ДАНИХ.....	11
1.1 Існуючі архітектурні рішення.....	11
1.2 Аналіз перспективних рішень.....	14
1.3 Розробка кластера проектованої РСЗД.....	19
1.4 Висновки до глави 1.....	22
2 ПАРАМЕТРИ ТА ХАРАКТЕРИСТИКИ РСЗ.....	23
2.1 Час обробки даних у підсистемі кодування інформації.....	23
2.2 Раннє діагностування відключення жорсткого диска в РСЗД	29
2.3 РСЗД як багатоканальна система масового обслуговування із відмовами	36
2.4 Відповідність структури РСЗД відомому профілю навантаження.....	41
ВИСНОВКИ.....	47
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	48
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	50

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

ПЗ – програмні засоби

РСЗД – розподілена система зберігання даних

СЗД – система зберігання даних

ВСТУП

Останніми роками економіці притаманний інноваційний характер, тобто. управління інтелектуальним капіталом підприємства відіграє ключову роль. Завдяки інформаційним технологіям відбувається глибока інтеграція різних галузей діяльності: дослідження, виробництво, комерціалізація, маркетинг, управління, логістика тощо. Зараз інформаційні технології без яких неможливо досягти успіху, набувають глобального характеру практично у всіх областях. Грамотне впровадження інформаційних технологій може призвести до покращення позиції на ринку та збільшити конкурентоспроможність підприємства.

Інформаційні технології сьогодні стають єдиною сферою діяльності та життєво необхідною галуззю. У великих підприємствах Росії стала проявляється проблема організації інформаційного взаємодії, оскільки характерні виробництва організаційні структури – холдинги, підприємства міста і т.д. передбачають складні взаємозв'язки між окремими, часто фінансово незалежними підприємствами.

Однією з основних та важливих проблем при розробці та впровадженні інформаційних систем управління виробничо-економічною діяльністю завжди було значне збільшення обсягів інформації з високим ступенем важливості, яку звичайно слід ретельно відбирати та акумулювати, досить надійно зберігати та оперативно обробляти. Розробка інформаційних систем від всесвітньо відомих вендорів: HP, IBM, BMC, Oracle, SAP, 1C, Microsoft, стала вирішенням вище зазначеної проблеми.

Однак, у ході розвитку економіки, заснованою на вибудовуванні взаємодії багатьох виробників і постачальників у рамках систем управління життєвим циклом продукції та відповідних ланцюгів поставок, створення на цій основі груп підприємств (ДПР), актуалізувалася проблема обробки безлічі різних інформаційних потоків та часто неформалізованої інформації

як усередині такої групи, так і при взаємодії із зовнішніми підприємствами. Зазначене призвело до того, що аналізувати, зберігати ці дані та приймати правильні ефективні управлінські рішення стає важко через надлишок даних та неструктурованість інформації.

Зараз перехід до «цифрової» економіки — це інвестиції в ІТ-інфраструктуру. Для інтенсивного інноваційного розвитку економіки необхідно створення ефективних комплексних інформаційних систем на національному рівні, що забезпечують високу реакцію при збиранні, аналізі та обміні даними про введення нових продуктів та процесів, а також виробничі технології.

Прогнозований загальний світовий обсяг даних, що зберігаються на 2022 рік, орієнтовно становить за нашими оцінками, близько 40 зеттабайт. Таке зростання ставить перед розробниками завдання знаходження рішень, що дозволяють здійснювати довгострокове надійне зберігання великих масивів даних із швидкою швидкістю відгуку та мінімальними затримками передачі. Якщо розглядати ДПР з безліччю підрозділів та філій, то таке об'єднання має складну інформаційну систему, в якій повинні зберігатися та оброблятися всі дані, що забезпечують управління виробничими процесами, матеріальними та грошовими потоками, кадрами, електронним документообігом та іншими необхідними для виробництва процесами. Але саме ефективність обміну інформаційними даними є основою інноваційного розвитку – від швидкості виведення ринку якісної та сучасної продукції залежить конкурентоспроможність економіки загалом. Рекомендована структура управління такими складними групами підприємств має бути горизонтальною. Розробка інноваційних видів продукції є зараз результатом роботи багатьох науково-виробничих об'єднань.

На даний момент вже існує чітке розуміння важливості ефективної організації процесів збору, обробки та аналізу інформації.

Метою роботи є створення методики проектування та розробки розподілених систем зберігання даних із заданими характеристиками.

Для досягнення поставленої мети було розроблено систему взаємопов'язаних завдань дослідження:

аналіз основних характеристик РСЗД, що впливають на її продуктивність;

розробка методик проектування РСЗД для заданого профілю навантаження;

розробка методик розрахунку характеристик службового інформаційного потоку з метою визначення коефіцієнта корисності інформаційного обміну в РСЗД;

Практична значимість полягає у розробці та вдосконаленні моделей, методів та створення інструментарію розробки та проектування РСЗД. Викладено нові моделі, методи та методики проектування РСЗД для групи підприємств, що має суттєве значення для економіки.

1 АРХІТЕКТУРНІ РІШЕННЯ ПОБУДОВИ СИСТЕМ ЗБЕРІГАННЯ ДАНИХ

1.1 Існуючі архітектурні рішення

Існуючі на сьогодні вимоги до проектування та розробки систем зберігання даних (СЗД), що забезпечують прийнятну надійність та продуктивність, зумовлюються об'єктивною тенденцією високого зростання обсягу даних, які потребують довготривалого зберігання. Приклади даних: значні обсяги неструктурованої інформації, що отримується від датчиків, систем спостереження та реєстрації тощо; архівні та статистичні дані тощо, що регулярно поповнюються в процесі роботи різних мережевих сервісів. Прогнозовані величини обсягу та інтенсивності інформаційного обміну пред'являють конкретні вимоги до показників надійності та продуктивності у використуваних та проєктованих СЗД.

Агентством Gartner був запропонований термін Web-Scale IT для опису технології побудови найбільш перспективних СЗД, а пізніше агентством було спрогнозовано, що цей підхід до побудови безпосередньо архітектури СЗД буде використано більш ніж 50% всіх основних компаній. Сьогодні цей підхід до розробки СЗД всесвітньо прийнято та практично випробувано багатьма провідними компаніями у цій галузі. Пояснимо, що під терміном Web-Scale IT мається на увазі концепція проектування досить великої обчислювальної інфраструктури (аналогічної за масштабами найбільшим світовим провайдерам IT-сервісів), яка надає організаціям можливість хмарного сервісу всередині корпоративної мережі.

Фахівці багатьох організацій зазначають, що Web-Scale IT включає реалізацію різних підходів до побудови IT-інфраструктури компаній і має наступні переваги: за короткий термін виведення нових продуктів на ринок, можливість проведення значної кількості внесення змін у стислий термін, використання відкритого апаратного та програмного забезпечення відповідно

до відомих та застосовуваних методологій Agile та DevOps. Основною відмінністю Web-Scale IT є реалізація стійкості до відмови на рівні програмного забезпечення і вимагає оптимізації коду під апаратне забезпечення. Пояснимо, що відмова від будь-яких технологій віртуалізації вносить ще один шар взаємодії з апаратним забезпеченням (і, цілком імовірно, відповідні затримки). Апаратне забезпечення розробляється з урахуванням архітектури програмного забезпечення. Апаратне забезпечення може бути інтегрованим або розподіленим на компоненти. Т.ч., архітектури, побудовані за Web-Scale IT, високо масштабуються, оскільки рівні продуктивності досягаються великою кількістю ефективних неспеціалізованих вузлів та використанням типових серверів. Сукупні ресурси багатьох цих вузлів об'єднуються в пул для створення «обчислювальної матриці» (computing fabric) та розподілу відповідно до потреби. Обладнання Web-масштабу – відносно недороге та доступне. При відмові вузол просто замінюється. Керує цією мережею вузлів інтегроване програмне забезпечення, яке – на відміну від апаратних засобів управління – може «перепрограмувати» систему, щоб вона відповідала мінливій динаміці. Така система по суті своїй еластична, оскільки дані, метадані та операції розподіляються на велику кількість малих вузлів, що замінюються, і, можливо, навіть на кілька центрів обробки даних. Ключовими елементами архітектурного рішення СЗД з Web-Scale IT є такі:

- орієнтація на використання центрів обробки даних (ЦОД);
- використання веб-орієнтованої або мікросервісної архітектури надання послуг;
- наявність програмованого управління СЗД;
- можливість налаштування будь-яких бізнес-процесів, орієнтованих на високу продуктивність;
- застосування методології «корпоративна культура організації, що навчається», сфокусована на інноваціях та безперервному навчанні.

Також зазначимо, що ряд фахівців, наприклад, згідно, виділяють два

різні підходи до забезпечення масштабованості. Вертикально-масштабовані системи (scale-up), яким властиве збільшення кількості доступних ресурсів за рахунок збільшення потужності застосовуваних обчислювальних засобів, та горизонтально-масштабовані системи (scale-out), що характеризуються наявністю значної кількості невеликих серверів, об'єднаних у кластер, на яких розподіляється поставлена замовником прикладне завдання. Незважаючи на різницю, scale-up та scale-out зазначені підходи можуть комбінуватися в рамках однієї СЗД.

Однак масштабування до великої кількості вузлів, у тому числі з можливим розподілом по географічно рознесеним ЦОД, може призвести до проблеми забезпечення доступності та консистентності даних.

В даний час найбільш характерними представниками систем, що горизонтально-масштабуються, спроектованих згідно Web-Scale IT, є СЗД компаній EMC Isilon і IBM XIV. Кластер горизонтально-масштабованої системи, зазначених рішень складається з серверів, що включають процесори, кеш-пам'ять, диски, Ethernet-порти засобів підключення. Крім даних пристроїв, вузол EMC Isilon містить також порти Infiniband RDMA, за допомогою яких сервери об'єднуються у внутрішню мережу з метою забезпечення стійкості до відмови кластера. Відповідно до архітектура кластера зазначених СЗД відповідає схемі, наведеній на рисунку 1.1. Зазначимо, що загальним для цих горизонтально-масштабованих систем є одиниця масштабування, що є апаратним модулем з фіксованою кількістю приєднаних до нього дисків. Основні відмінності цих систем полягають в особливостях організації доступу до даних і способу їх зберігання. В одному випадку забезпечується доступ до даних через організацію файлової системи, в іншому випадку використовується блоковий пристрій.

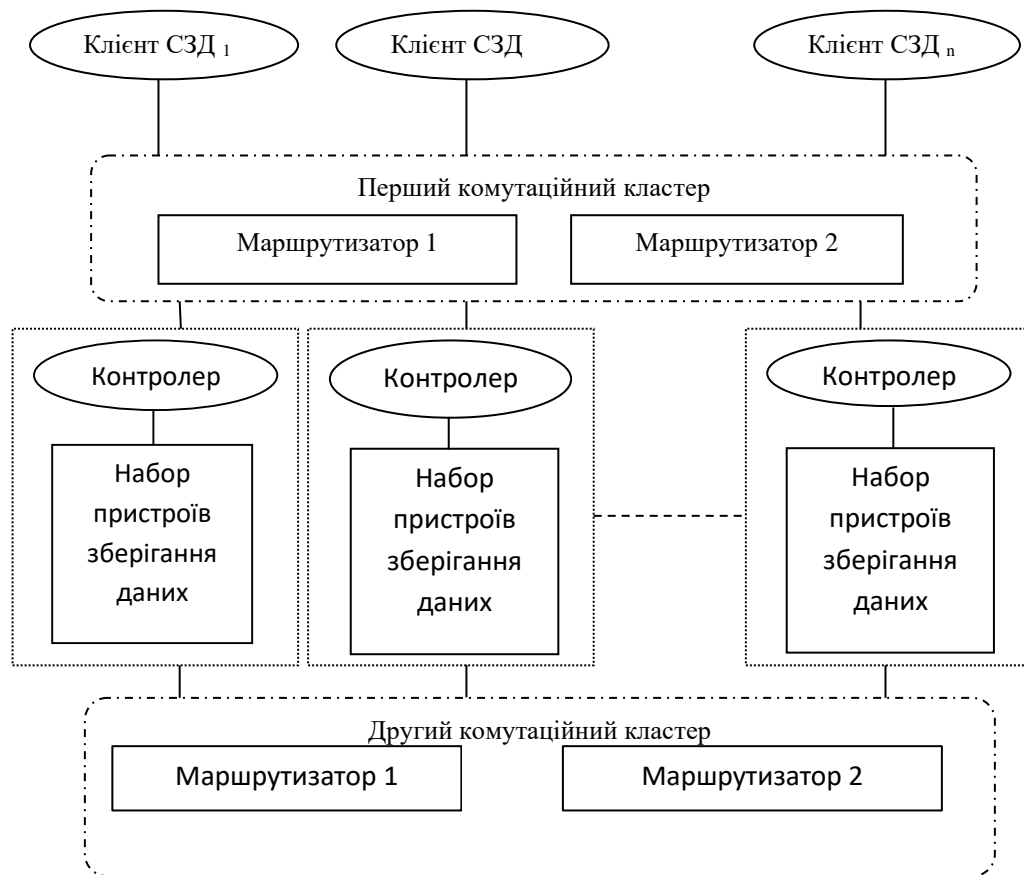


Рисунок 1.1 – Архітектура кластера СЗД EMC Isilon та IBM XIV

1.2 Аналіз перспективних рішень

На основі аналізу СЗД від компанії EMC і IBM, зокрема Isilon, а також автором на основі власного досвіду, зроблено висновок, що всі дані розподілені між усіма вузлами даних СЗД, і будь-який вузол може на рівних обслуговувати запити до будь-якого файлу, незалежно від того, на дисках або кеші якого вузла знаходяться дані. Тим самим досягається висока продуктивність роботи з будь-яким файлом. Також варто зазначити, що ресурси об'єднані у пул, архітектура паралельна. У цій концепції при використанні SSD (Solid-State Drive) повинна досягатися максимальна ефективність утилізації SSD дисків. Крім зазначеного, має бути передбачена можливість використання SSD для метаданих. При файловому доступі значну частину операцій становлять саме операції з метаданими, частка

часто, зазвичай, перевищує 50%. Наведемо приклад, перш ніж почати читати блоки файлу з мережного диска, необхідно пробігтися по дереву каталогів, отримати атрибути файлу, отримати список його фізичних блоків. Вказане вимагає певних операцій та додає затримки. У цьому випадку безліч невеликих звернень до метаданих моментально перенаправляється на обслуговування на SSD, а не втрачає час разом з великими запитами до даних у черзі до повільних дисків. Т.ч., можна додати в кластер всього 1-2% ємності на SSD (наприклад, по одному SSD у кожен вузол) і отримати значний приріст продуктивності файлових операцій (до 30% на близьких до випадкових навантажень при вдвічі меншому часі відгуку. У традиційних системах потрібно стежити за завантаженням контролерів, портів, дискових груп, що стає практично неможливим із зростанням обсягів та числа систем. В описаній ситуації передбачається лише один найбільш ефективний і прийнятний підхід, а саме: рівномірно розподілити дані та навантаження, що дозволить досягти високої ефективності як у сенсі завантаження всіх компонентів, так і в сенсі корисної ємності.

Говорячи про коефіцієнт корисної ємності, часто враховують лише накладні витрати на RAID (Redundant Array of Independent Disks) та ФС (файлова система), забуваючи про другу складову загальної ефективності. ПТК (програмно-технічний комплекс) завдяки своїй схемі захисту даних та єдиної ФС може забезпечити реальний коефіцієнт корисної ємності понад 80% .

Класичні платформи зберігання SAN та NAS, хоч і критично важливі для корпоративних додатків, але не призначені для роботи з сучасними хмарними додатками та не відповідають вимогам масштабування у хмарі. Крім того, постійне зростання обсягів неструктурованих даних вимагає створення більш простої архітектури зберігання, яка призначена для управління мільярдами та трильйонами файлів, прискорює розробку хмарних та мобільних додатків та додатків для великих даних і при цьому дозволяє знизити витрати, пов'язані із зберіганням, та витрати в цілому . Щоб

вирішити ці завдання, ІТ-служби та постачальники послуг почали тестувати та впроваджувати економічні типові та відкриті інфраструктури. Типові компоненти та відкриті технології на базі стандартів знижують витрати на зберігання, але окремо забезпечують нижчі показники продуктивності та надійності. Також слід пам'ятати, що інфраструктура зручна в обслуговуванні, якщо ІТ-фахівці замовника мають достатній досвід її експлуатації.

Безпрецедентне зростання обсягів даних (структурованих та неструктурованих) призводить до того, що файли великих обсягів (наприклад, зображення, відео тощо) зберігаються у дорогих ізольованих СЗД. У класичній інфраструктурі, зазвичай, часто дані ізолюється, ускладнюючи цим обмін і управління ними і збільшуючи витрати. Зазначений фактор не підтримує ефективне та економічне масштабування, що змушує замовників шукати рішення, яке поєднує в собі переваги публічної та приватної хмари. Рішення дозволяють будь-якій організації об'єднати кілька систем зберігання та архівів змісту в єдиний глобально доступний та ефективний репозиторій змісту, який може обслуговувати велику кількість програм .

Покладена в основу рішення архітектура типу «активний-активний» з підтримкою кількох майданчиків, єдиний глобальний простір імен та універсальна доступність забезпечують доступ до змісту з будь-якої точки для будь-якої програми чи пристрою. Завдяки розподілу контейнерів даних між майданчиками операції запису та читання можуть виконуватися в будь-якій точці світу. Рішення забезпечують високий рівень семантичної узгодженості, що дозволяє спростити розробку додатків та отримувати доступ до даних із будь-якої точки. Також у рішенні є функція геокешування, яка визначає схеми доступу з декількох майданчиків та кешує дані на майданчику, де ці дані використовуються найчастіше .

У загальноприйнятих класичних інфраструктурах зберігання даних, як правило, запровадити системи бізнес-аналітики може бути непросто. Так,

дані часто розподілені між кількома складними системами, і ускладнює та збільшує витрати на доступ та управління. Організації змушені отримувати дані від операційних систем і потім завантажувати їх у виділений кластер для аналізу. В основі ефективної аналітики лежать точність та своєчасність даних. Рішення надає замовникам можливість видобувати практично значущі дані бізнес-аналітики з великих обсягів розподіленого змісту, не використовуючи робочого процесу ETL (Extract, Transform, Load). Рішення можуть підтримувати озеро даних, надаючи можливості доступу та управління будь-яким змістом організації за допомогою кількох дистрибутивів Hadoop .

При використанні звичних СЗД зі зростанням обсягів даних неминуче зростає кількість контролерів, дискових груп, томів, файлових систем, так чи інакше з'являються «острівці» різнорідних систем, стає все складніше та складніше підтримувати ефективність їх використання. Нарощування систем вимагає попереднього аналізу, складних робіт з конфігурації та міграції даних .

В останні роки стає актуальним завдання аналізу великих даних. Тобто, необхідна система, яка дозволить отримати цінні результати з тих великих обсягів інформації, які вже створювалися різними системами та процесами, але які було неможливо, або важко обробити використовуючи класичні методи .

Передбачається, що система має дозволяти:

- захоплювати та зберігати вихідні неструктуровані дані у відносно недорогій та масштабованій платформі;
- зберігати дані різних типів та з різних джерел в єдиному репозиторії;
- виконувати операції з трансформації даних у рамках єдиної системи;
- визначати структуру даних у момент, коли цю інформацію необхідно використовувати, тобто. структура даних може, якщо необхідно, визначатися схемою читання;

- легко впроваджувати та застосовувати нові механізми та інструменти з аналізу даних;
- досить ємно виконувати унікальні, під «одиничне» конкретне завдання, аналітичні запити;
- надавати досить широкі можливості доступу до даних для різноманітного аналізу.

При проектуванні рішень СЗД, зазначених у главі 1.1, їх загальна продуктивність визначається продуктивністю модуля управління, вплинути який із боку персоналу (адміністраторів) мало представляється можливим. А ось перспективнішим варіантом управління продуктивністю є виділення двох рівнів масштабування: масштабування за ресурсами, яке виконується управлінням кількості апаратно-керованих масивів жорстких дисків, і масштабування за продуктивністю - управлінням кількістю апаратних блоків підготовки та обробки інформації. Згідно з вищевикладеним, пропонується архітектурне рішення на основі проаналізованого підходу, представлене на рисунку 1.2.

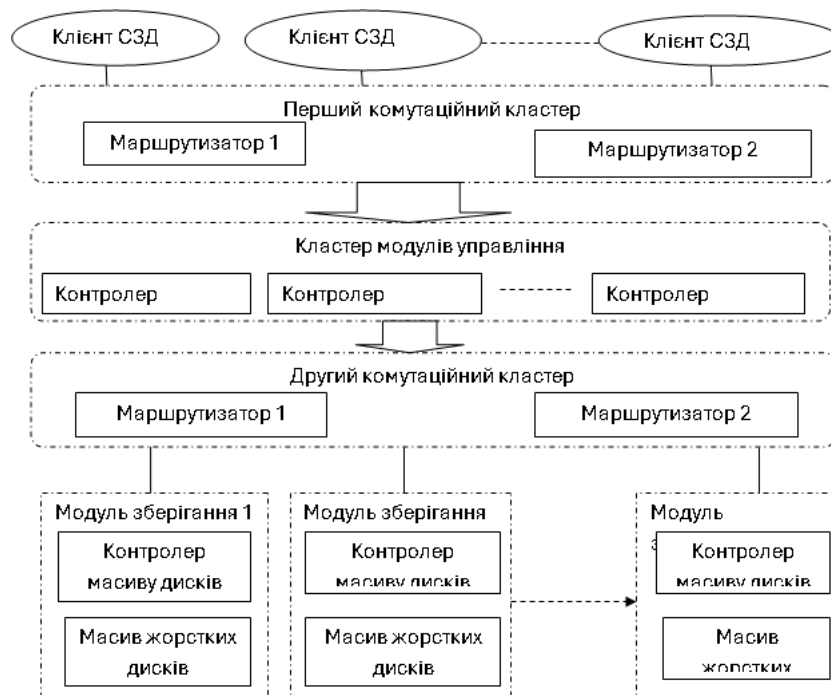


Рисунок 1.2 – Пропонована архітектура кластера проектованої СЗД

Зазначимо, що апаратний модуль зберігання даних містить у собі, крім контролера управління та інтерфейсу обміну даними з масивом жорстких дисків, також і резервовану область пам'яті зі збереженою інформацією про адресний простір підключеного масиву жорстких дисків. Тому до складу програмного забезпечення апаратного модуля підготовки та обробки інформації, крім модуля доступу та обміну даними, а також блоку підготовки та розподілу даних, входять репліковані області пам'яті, в які здійснюється прийом інформації та зберігання ідентифікаційних записів про структури, що зберігаються у СЗД.

1.3 Розробка кластера проектованої РСЗД

На основі вищевикладеного та практичного досвіду виконано розробка архітектурного рішення кластера РСЗД, представленої на рисунку 1.3. Згідно з розробленим рішенням, кластер включає в себе мінімум один модуль керування та один або більше модулів зберігання.

Необхідно відзначити, що модуль управління складається з модуля, забезпечує зв'язок із клієнтом (S3 Сервер), модуля управління нереляційною базою даних (MongoDB), що забезпечує роботу з метаданими та модуля, що здійснює кешування, розподіл блоків та зв'язок з модулями зберігання (Контролер). Кожен модуль зберігання є програмно-апаратним комплексом, що включає сервер, операційну систему та програми для роботи з модулем управління (А точніше, програмний модуль «TLV-клієнт» для зв'язку з модулем управління; програмний модуль White DB для зберігання метаданих модуля зберігання).

Крім цього, слід зазначити, що в даній архітектурі для клієнта використовується виключно блоковий доступ. Клієнтом є програмно-апаратний комплекс (Сервер), що має у своєму складі операційну систему, яка організовує файлову систему на блоковому пристрої від РСЗД. Пояснимо, що блоковий доступ – ресурс, що представляється клієнту СЗД як

підключеного жорсткого диска. Як і зі звичайним жорстким диском, з ним можна виконувати операції розбиття на розділи, форматування, установки операційної системи. Підключений блоковий ресурс може бути використаний для завантаження робочих станцій бездисків. У разі блокового доступу використовується протокол iSCSI (Internet Small Computer System Interface). Найменша одиниця зберігання інформації – блок. Розмір блоку задається на етапі налаштування блокового ресурсу. Протокол блочного доступу надає значно більшу швидкість доступу у зв'язку з меншою кількістю проміжних шарів при доступі до дискового масиву.

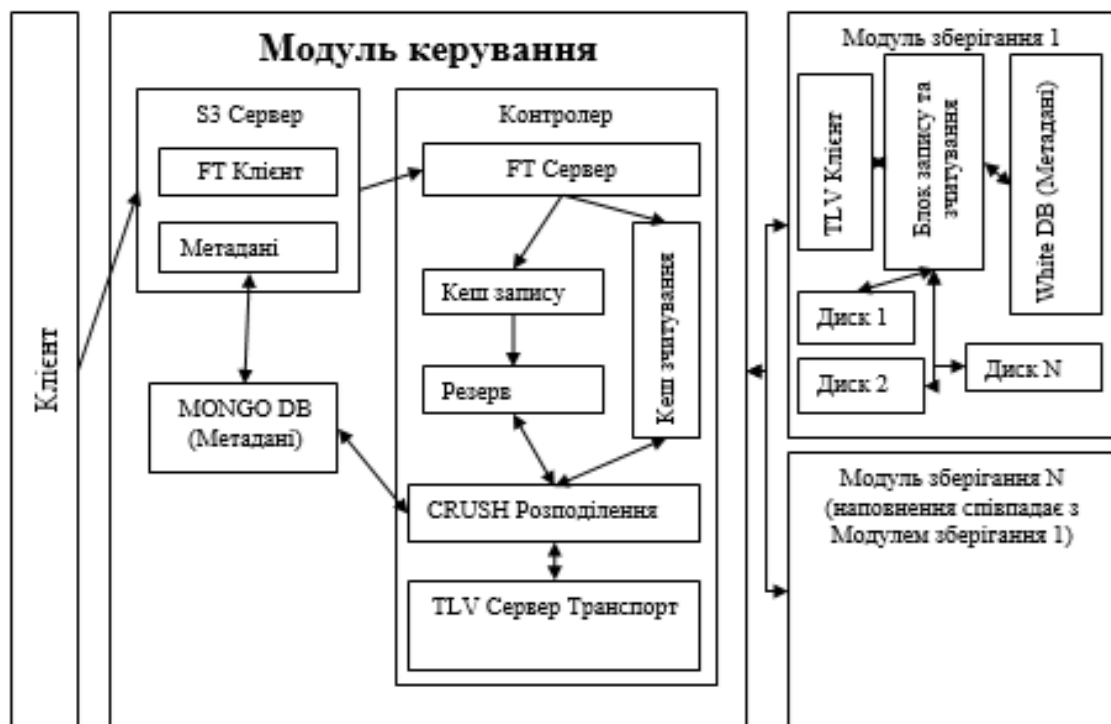


Рисунок 1.3 – Архітектурне рішення кластера проектової РСЗД

Модулі зберігання РСЗД утворюють високопродуктивний кластер, асинхронно та паралельно взаємодіє з модулями управління РСЗД. Модуль керування готує дані для кеш запису, а модулі зберігання у свою чергу паралельно забирають дані для запису. За рахунок масового

паралелізму забезпечується висока пропускна спроможність системи.

Як показує дослідження та практичний досвід, в даний час на ринку СЗД найбільш затребуване рішення Isilon Dell.

У Таблиці 1.1 наводяться параметри Isilon Dell.

Таблиця 1.1 – Характеристики Isilon Dell

Характеристика	Існуюче рішення
Файловий доступ до даних	так
Блочний доступ до даних	ні
Об'єктний доступ до даних	ні
Кількість вузлів системи зберігання на кластер	До 144
Кількість вузлів системи керування на кластер	До 144
Ємність кластеру	До 30,2 Пбайт
Час додавання вузла	До 60 сек.
Реплікація даних	Асинхронна реплікація даних на рівні файлів
Шифрування даних	Застосування самошифруючихся дисків (SED)
Дедуплікація даних	ні

Важливо помітити, що запропоновані нами структурні рішення побудови кластера горизонтально-масштабованої СЗД є перспективнішими у розвитку технології Web-Scale IT, і дозволяють забезпечити надійність і високу продуктивність СЗД у реальних умовах значно швидкого зростання обсягу даних, необхідні довгострокового зберігання. Проектована нами СЗД передбачається, що найефективніша у створенні інформаційного обміну всередині корпорацій, мають географічно рознесені ЦОД .

1.4 Висновки до глави 1

Концепція Web-Scale IT у технології побудови сучасних інформаційних систем, у тому числі і СЗД найбільш перспективна, завдяки стиснутому терміну виведення нових продуктів на ринок, можливості проведення значної кількості внесення змін у стислий термін, використання відкритого апаратного та програмного забезпечення відповідно до методологіями Agile та DevOps.

За наявності вимог до гнучкості щодо продуктивності та масштабування СЗД найбільш перспективними є scale-out (РСЗД) рішення, застосування яких тільки починається в даний час.

Підвищення гнучкості РСЗД може бути здійснено за рахунок поділу модулів управління та модулів зберігання в одному рішенні.

Пропонуються горизонтально-масштабовані системи (scale-out), що характеризуються наявністю великої кількості серверів, об'єднаних у кластер, на якому розподіляється прикладне завдання.

Концепція Web-Scale IT визначає заміну мережевої технології FiberChannel технологією Ethernet, що має більш високу пропускну здатність каналів зв'язку та інтерфейсів структурних елементів, а також надає можливість забезпечення пріоритетності обслуговування найважливіших інформаційних процесів.

Для запропонованого архітектурного рішення потрібно розробити відсутні методики аналізу нововведених у схему елементів і розробити алгоритм синтезу РСЗД.

2 ПАРАМЕТРИ ТА ХАРАКТЕРИСТИКИ РСЗД

2.1 Час обробки даних у підсистемі кодування інформації

До складу РСЗД для реалізації вище заявлених технологій включена підсистема кодування інформації, яка відповідає за форматування отриманої від клієнта інформації до структур даних, прийнятих у самій РСЗД. До підсистеми пред'являються вимоги щодо надійності та продуктивності при можливому відновленні даних без зниження працездатності бізнес-процесів .

Прийнятний рівень надійності відновлення даних у РСЗД забезпечується за рахунок 1) генерації певного обсягу надлишкової інформації та 2) використання алгоритмів локалізації та виправлення помилок, що виникають при передачі даних або виході з ладу носіїв інформації, що зберігається.

Сьогодні найбільш поширеним підходом до організації зберігання інформації в РСЗД є кодування інформації циклічним кодом і подальший розподіл сформованих блоків даних по всьому доступному масиву носіїв інформації. Зазначене дозволяє здійснити захист інформації від помилок та дозволяє виконати кодування із заздалегідь визначеними коригуючими властивостями. Алгоритм розподілу сформованих блоків даних по всьому доступному масиву носіїв інформації, що зберігається, як правило, вибирається виходячи з вимог алгоритмів кодування інформації в розділі відновлення інформації з урахуванням можливого виходу з ладу групи пристроїв з масиву носіїв інформації, що зберігається.

При проектуванні РСЗД планується використання недвійкового циклічного розширеного коду Ріда-Соломона на полі Галуа GF, тобто. із здійсненням кодування байтів інформації. Розроблено «алгоритм кодування блоку інформації, що надійшла від клієнта, який полягає в побудові матриці з фіксованим розміром стовпця, який, у свою чергу, логічно розбивається на

дві області, одна з яких визначає розмір кодованого блоку інформації, що надійшла, а друга зарезервована під необхідну кількість надлишкової інформації. Розміром цієї області визначаються коригувальні властивості кодування. Блок інформації, що надійшла розподіляється в матриці в області значущої інформації, після чого по кожному стовпцю виконується кодування інформації для подальшої передачі на носій інформації, що зберігається ». Пояснимо, що матриця має фіксоване число рядків, а кількість стовпців визначається розміром блоку інформації, що надійшла, і розміром області значущої інформації стовпця матриці.

Крім того, відзначимо, що процедура формування та кодування матриці є однією з найбільш трудомістких у процесі перетворення, і, зниження витрат на цій процедурі призведе до підвищення загальної продуктивності РСЗД в цілому.

Т.к. довжина шпальти фіксована, то середній час обробки одного шпальти матриці, при наявній обчислювальній потужності, можна вважати величиною постійної. Отже, можна дійти невтішного висновку, що загальний час операції перетворення пропорційно кількості стовпців.

Слідую способу формування матриці: збільшуючи розмір логічної області значущої інформації, можна домогтися розміщення блоку інформації, що надійшла від клієнта, в матриці меншого розміру, хоча при цьому буде знижений рівень захисту інформації. На рисунку 2.1 показано залежність кількості стовпців матриці від розміру області значущої інформації. На рисунку 2.2 наведено збільшену ділянку цього графіка. Проілюстровані графіки отримані при розмірі блоку інформації, отриманого від клієнта, що дорівнює 14 Мбайт.

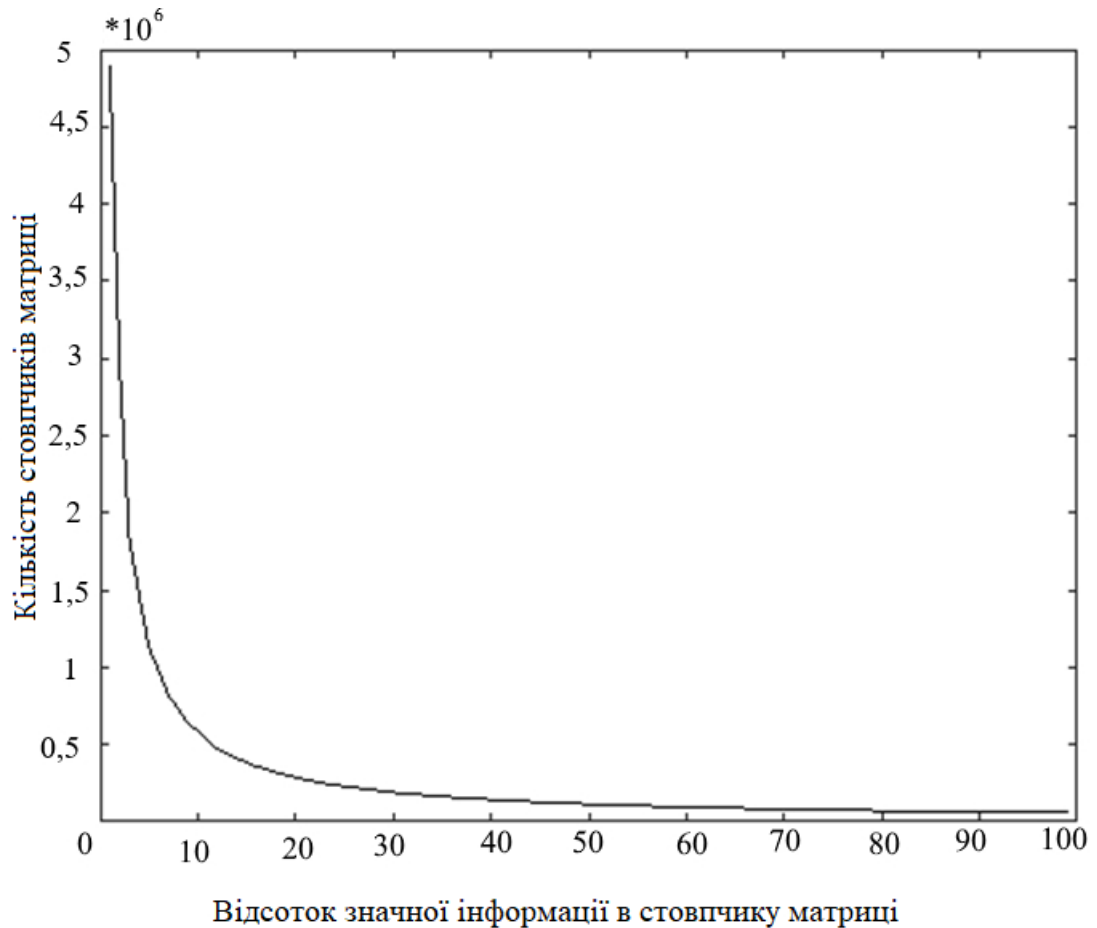


Рисунок 2.1 – Графік залежності розмірності матриці від обсягу області значної інформації у стовпчику матриці

Проаналізувавши отриманий графік, можна дійти невтішного висновку, що залучення логічних областей значної інформації мінімальної розмірності недоцільно, т.к. це призводить до різкого збільшення розмірів матриці підсистеми кодування інформації та, відповідно, до різкого збільшення часу на її обробку .

Однак, якщо вдасться підібрати архітектурне рішення РСЗД таким чином, щоб без шкоди для можливості відновлення даних виявилася допустимою можливість зниження рівня захисту інформації (і, як наслідок, можливість збільшення розміру логічної галузі значущої інформації), то таким рішенням можна досягти зниження часу обробки

матриці підсистеми кодування інформації та цим підвищити загальну продуктивність РСЗД загалом.

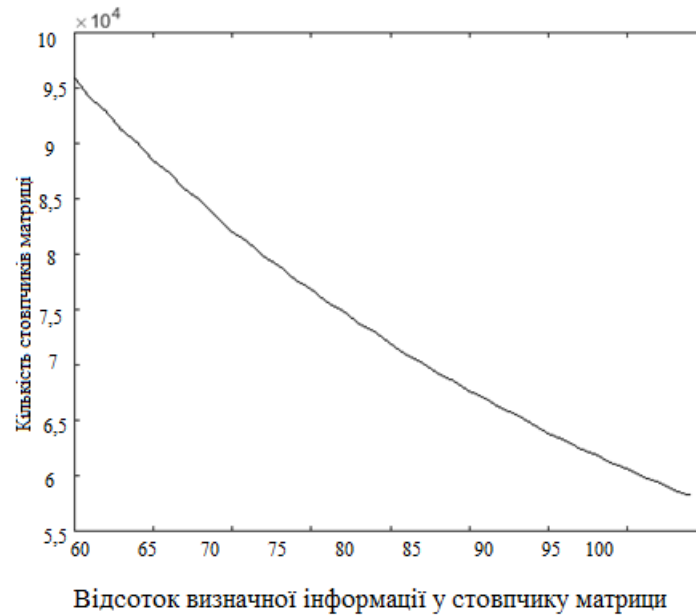


Рисунок 2.2 – Ділянка графіка, необхідний проведення дослідження залежності розмірності матриці від обсягу області значної інформації, у стовпці матриці

Вибір рішень для архітектури РСЗД, наведеної в розділі 1.1., зводиться до цільового підбору певної кількості модулів зберігання даних та масиву носіїв інформації, що зберігається, при яких вихід з ладу одного з носіїв збереженої інформації та/або в т.ч. модуль зберігання не вплине на можливість відновлення інформації.

На рисунку 2.3 представлений графік залежності кількості блоків даних, включених при розподілі матриці підсистеми кодування інформації на носій у масиві носіїв інформації, що зберігається, від розміру масиву носіїв інформації, що зберігається. У моделюванні використовувалися такі параметри: - розмір блоку інформації, отриманого від клієнта, 14 Мбайт; Обсяг логічної галузі значної інформації 75%. Розподіл даних зроблено за алгоритмом CRUSH. Пунктирна пряма на

Рисунках 2.3, 2.4 показує максимальну кількість байт у стовпці матриці підсистеми кодування інформації, втрата яких не завадить декодування вихідного блоку інформації користувача .

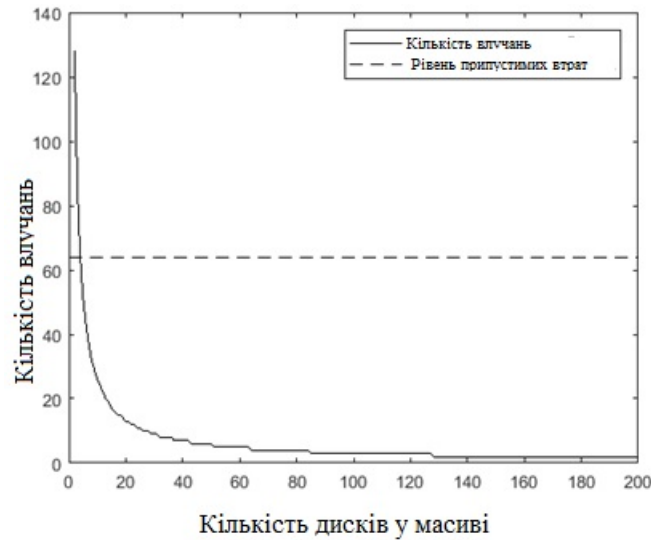


Рисунок 2.3 – Частота влучення даних з 1-го стовпця матриці на 1-ий диск масиву дисків, при розподілі за алгоритмом CRUSH, залежно від кількості дисків у масиві

На Рисунках 2.4, 2.5 наведено графіки залежності кількості блоків даних, включених при розподілі одного стовпця матриці підсистеми кодування інформації на носії одного модуля зберігання даних, від кількості носіїв інформації, що зберігається.

Проаналізувавши отриманий графік, наведений на рисунку 2.4, можна дійти невтішного висновку, що обсяг логічної області значної інформації 75% є недостатнім задля забезпечення захисту інформації, т.к. втрати даних при виході з ладу, наприклад, одного модуля зберігання перевищують можливості коригувального алгоритму відновлення інформації .

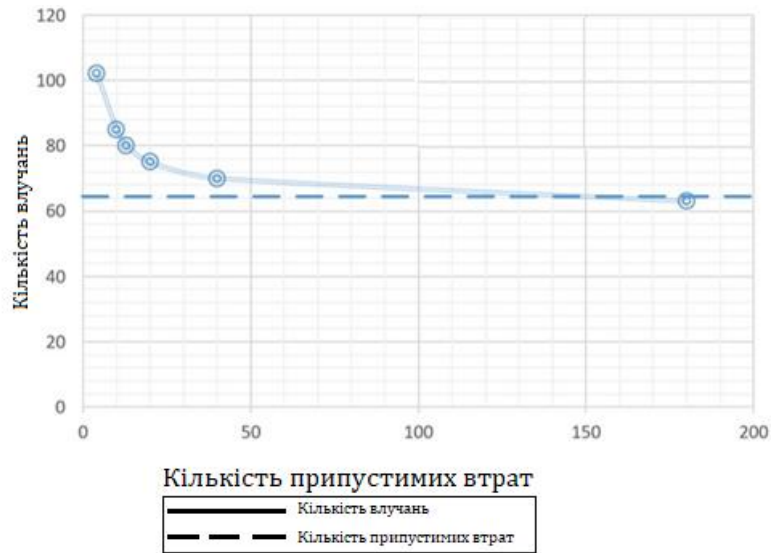


Рисунок 2.4 – Частота влучення даних з 1-го стовпця матриці на диски 1-го модуля зберігання даних, при розподілі за алгоритмом CRUSH, в системі із чотирма модулями зберігання даних

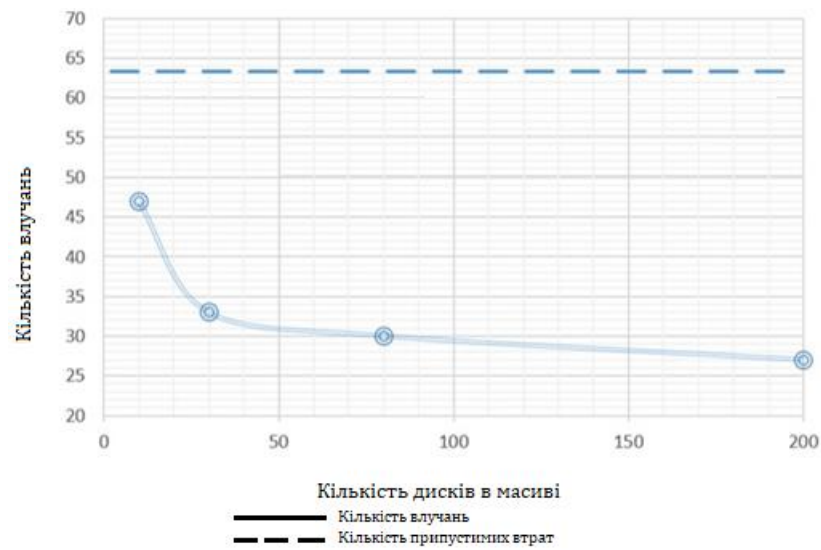


Рисунок 2.5 – Частота влучення даних з 1-го стовпця матриці на диски 1-го модуля зберігання даних, при розподілі за алгоритмом CRUSH, в системі з десятьма модулями зберігання даних

Зазначимо, що з архітектурі РСЗД з понад десять модулів зберігання даних, розмір логічної області значної інформації 75% є заниженим, і цілком допустимо його підвищення. Однак це призведе до зниження рівня допустимих втрат, хоча не позначиться істотно на можливостях алгоритму відновлення інформації, що коректує. Одночасно, підвищення обсягу області значимої інформації призведе до зниження розмірності матриці підсистеми кодування інформації, що, своєю чергою, знизить витрати на її обробку.

Шляхом розрахунку (моделювання) певної архітектури розподіленої системи зберігання даних можна домогтися зниження витрат часу на кодування блоку інформації, отриманого від клієнта і тим самим підвищити загальну продуктивність РСЗД

2.2 Раннє діагностування відключення жорсткого диска в РСЗД

Важливою та обов'язковою складовою сучасних систем зберігання та обробки даних є підсистема діагностування, до функції якої входить виявлення збоїв у роботі компонентів системи, в цілому, та своєчасне оповіщення персоналу, що виконує адміністрування системи.

Через характерні властивості РСЗД розподіл збереженої інформації здійснюється рівномірно по всьому доступному масиву жорстких дисків . У зв'язку із цим стає досить актуальним питання своєчасного виявлення дефектних дисків у масиві. Під дефектністю жорсткого диска, в РСЗД, розуміється або неможливість обміну даними між диском і контролером, що його опитує, або виникнення тривалих неспровокованих затримок в обміні даними з жорстким диском.

У РСЗД всі інформаційні блоки, адресовані дефектному жорсткому диску, будуть загублені і їх відновлення буде потрібно запуск відповідних процедур, які у переважній більшості випадків характеризуються великим

обсягом обчислювальних операцій. В результаті, поява дефектного жорсткого диска в масиві жорстких дисків РСЗД неминуче призводить до втрат продуктивності системи в цілому, тим більшим, чим довше виявлення дефектного диска в системі.

Звичайна процедура виявлення дефектного диска передбачає очікування відповіді жорсткого диска протягом максимального часу, а також деякої кількості повторних спроб звернення до диска. Без підтвердження обміну даними з жорстким диском після закінчення зазначених вище операцій у системі проводиться зміна статусу жорсткого диска.

У ході дослідження було розроблено один із способів раннього діагностування виходу з ладу диска в РСЗД з використанням математичного апарату теорії нечітких систем.

В силу характерних властивостей РСЗД, пов'язаної з рівномірним розподілом інформаційних блоків по всьому масиву жорстких дисків, контролер Модуля Управління генерує обмін даними з контролером Модуля Зберігання пропорційно до кількості дисків, підключених до Модулю Зберігання. У разі однотипної (однорідної) архітектури Модулів зберігання, підключених до Модулю Управління, інтенсивність обміну даними Модуля Управління з кожним із Модулів Зберігання буде практично однаковою і визначатиметься максимально можливою продуктивністю підсистеми кодування та розподілу інформації Модуля Управління. З іншого боку, інтенсивність обміну даними Модуля Зберігання з підключеними до нього жорсткими дисками визначається максимальною продуктивністю жорстких дисків, яка в свою чергу може залежати від ступеня заповнення диска, часу відгуку та інших причин і, в загальному випадку, не може вважатися однаковою навіть для однорідних дисків. Крім того, сумарна продуктивність процесу обміну даними між контролером Модуля Зберігання та приєднаних до нього жорстких дисків може бути нижчою за продуктивність обміну даних між контролером Модуля Управління та контролером Модуля Зберігання. Для недопущення зниження загальної продуктивності системи,

пов'язаної з неоднорідністю процесу обміну даними з жорсткими дисками, взаємодія між контролером Модуля Управління та контролером Модуля Зберігання здійснюється через деякий буфер обміну даними.

Зазначимо, що динаміка заповнення буфера обміну даними між Модулем Управління та Модулем Зберігання в процесі запису інформації, що зберігається в РСЗД, за певних умов, може служити індикатором наявності дефектних дисків у масиві дисків, підключених до даного Модулю Зберігання.

На рисунку 2.6 наведено результати математичного моделювання кривих заповнення буфера обміну даними між Модулем Управління та Модулями Зберігання для РСЗД, що складається з одного Модуля Управління та вісімнадцяти підключених до нього Модулів Зберігання. До кожного Модулю Зберігання підключено по 4 однорідні жорсткі диски з різними показниками продуктивності обміну даними між контролером Модуля Зберігання і підключеними до нього жорсткими дисками, що визначаються ступенем заповнення жорсткого диска.

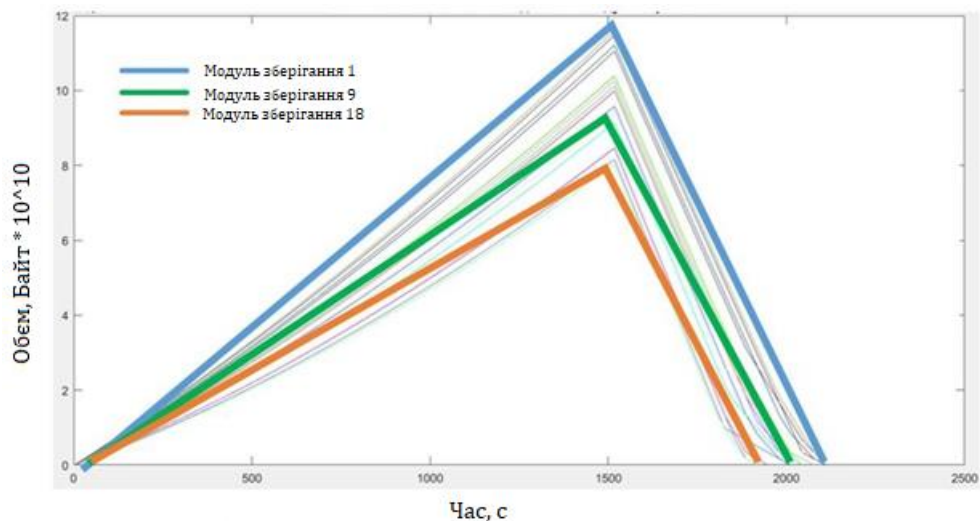


Рисунок 2.6 – Результати математичного моделювання швидкості заповнення буфера обміну даними між Модулем Управління та Модулями Зберігання

У разі появи в процесі запису дефектного диска в масиві, підключеному до Модулю Зберігання, у відповідному буфері обміну даними між Модулем Управління та Модулем Зберігання автоматично почне формуватися черга з інформаційних блоків, адресованих дефектному диску, що призведе до зміни загальної динаміки обміну даних. . На Рисунку 2.7 представлений результат моделювання ситуації, коли у Модулі Зберігання №2 один із жорстких дисків перестав відповідати на запити протягом певного проміжку часу. Динаміка заповнення буфера обміну даних між Модулем Управління та Модулем Зберігання №2 наведена порівняно з аналогічними кривими інших Модулів Зберігання. На рисунку 2.8 наведено графіки швидкості заповнення буфера обміну даних між Модулем Управління та Модулями Зберігання.

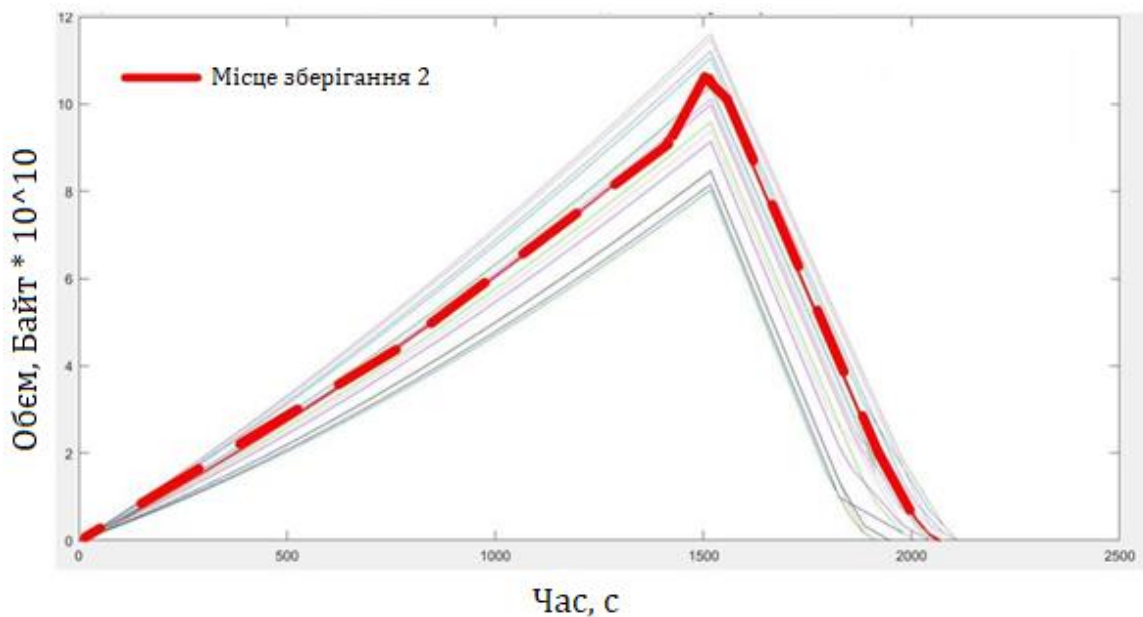


Рисунок 2.7 – Моделювання відмови HDD у Модулі Зберігання №2

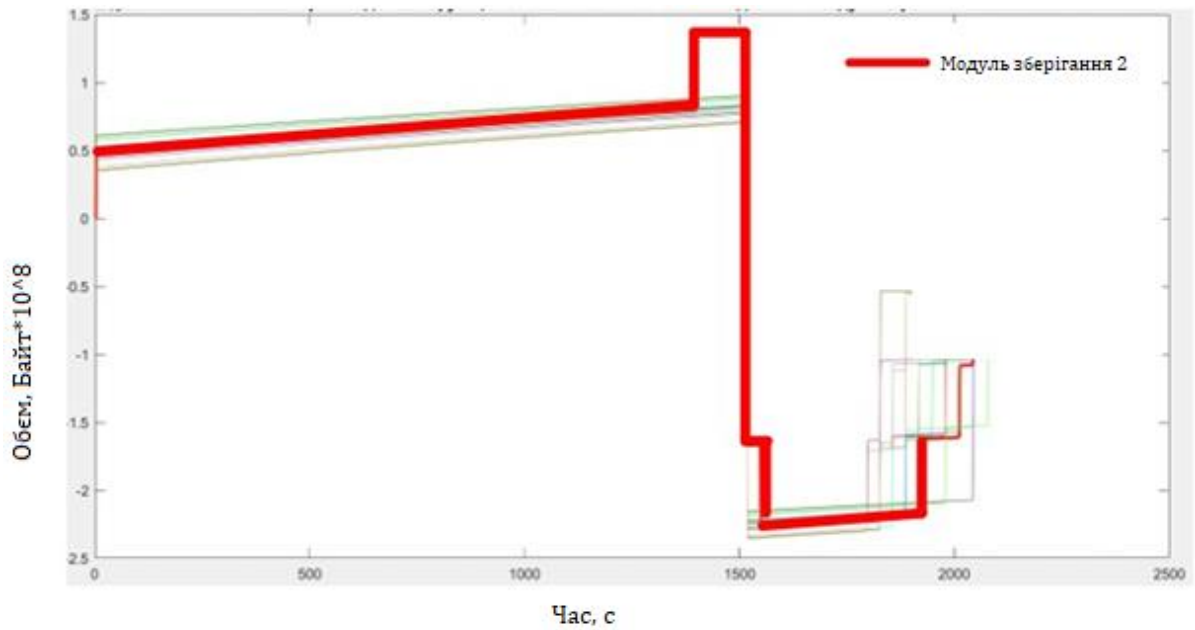


Рисунок 2.8 – Графіки швидкості заповнення буфера обміну даних між Модулем Управління та Модулем Зберігання №2

Графіки, представлені на рисунку 2.8, можуть бути індикатором ситуації появи дефектного диска в масиві жорстких дисків.

Один із численних способів розрахунку ймовірності події наявності дефектного диска в масиві використовує математичний апарат нечітких систем, реалізований, наприклад, у складі програмного продукту Matlab версії R2017b. В якості вхідних аргументів нечіткої системи визначення функції належності наявності дефектного жорсткого диска в масиві дисків, підключених до Модулю Зберігання n , візьмемо відношення величини швидкості заповнення буфера обміну даних між Модулем Управління і Модулем Зберігання n до аналогічних величин Модулей Ця величина була отримана на стенді експериментальним шляхом і дорівнює 0,058824. Графічне подання нечіткої системи наведено на рисунку 2.9. Параметри вхідних та вихідних нормувальних функцій наведено на рисунках 2.10 – 2.12.

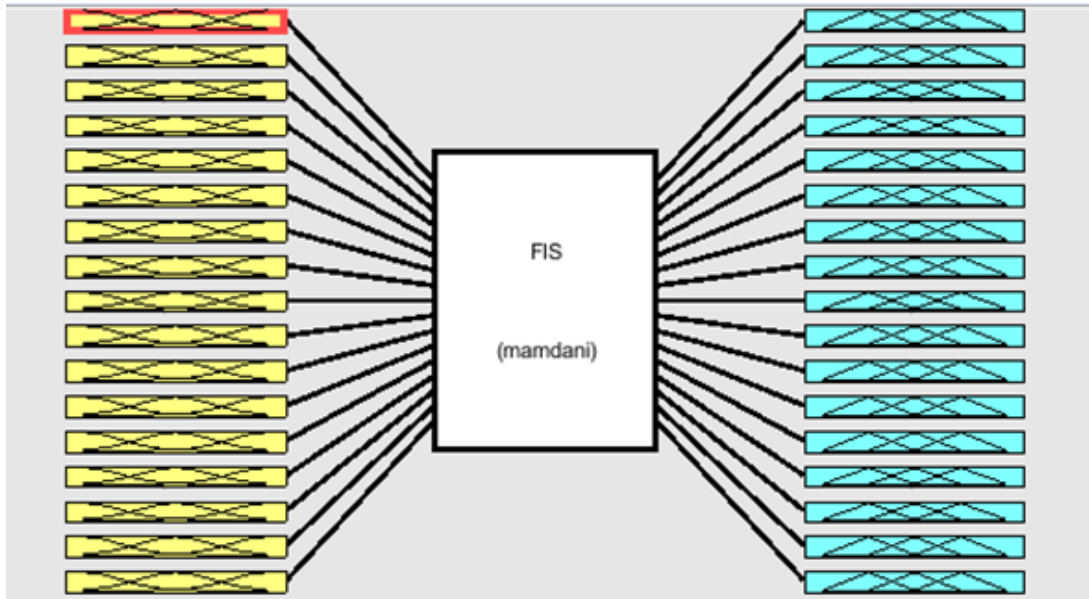


Рисунок 2.9 – Графічне подання нечіткої системи

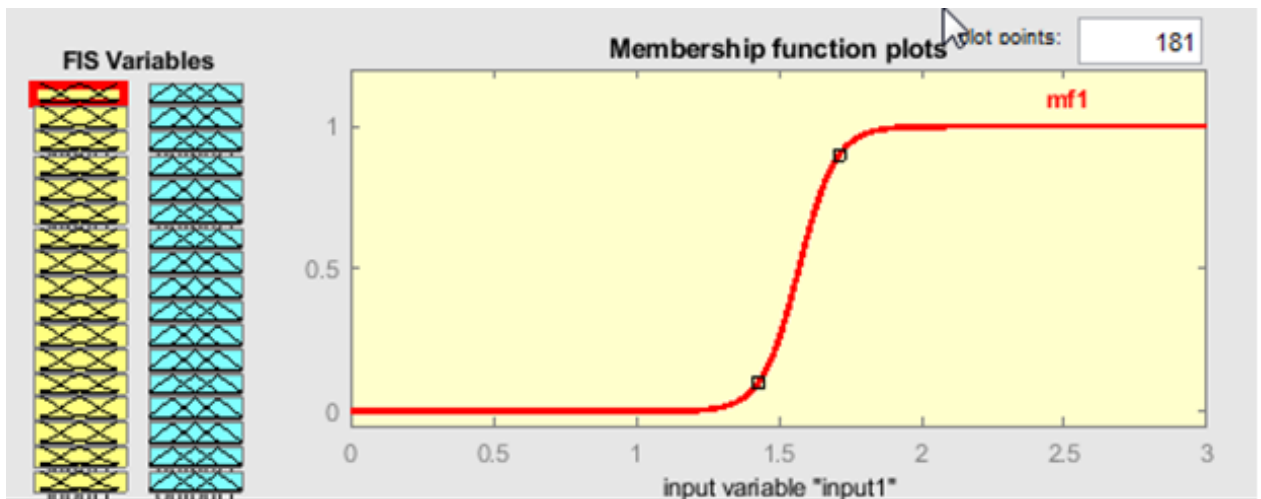


Рисунок 2.10 – Параметри вхідних нормувальних функцій

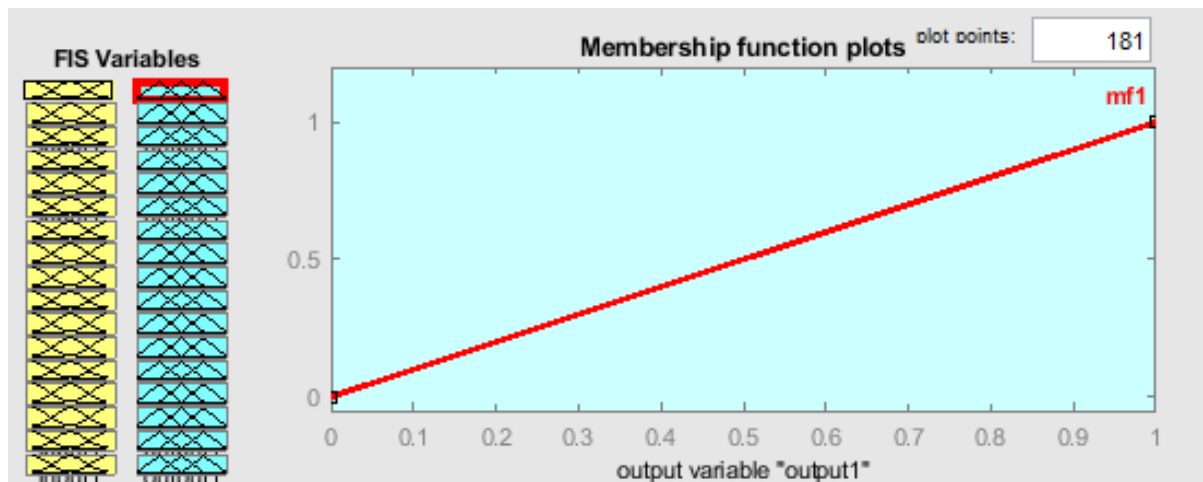


Рисунок 2.11 – Параметри вихідних нормувальних функцій

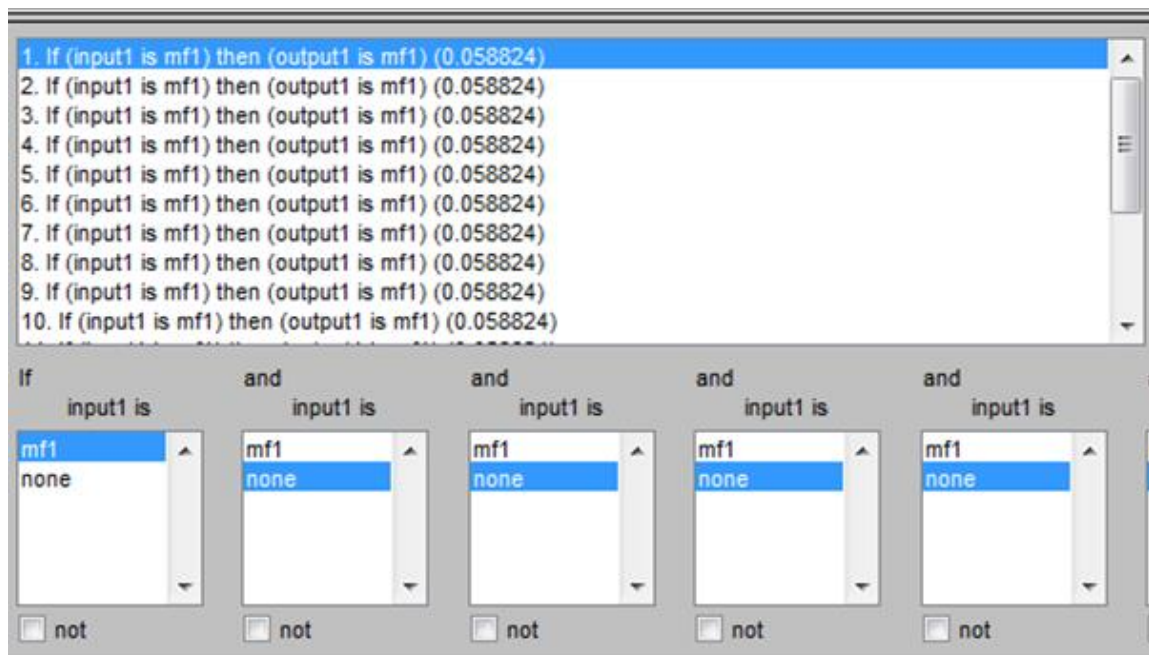


Рисунок 2.12 – Графічне представлення набору експертних правил

Підсумковим результатом є значення ймовірності наявності дефектного жорсткого диска масиві дисків, підключених до Модулю Зберігання в момент часу t . Графіки ймовірностей для прийнятих нами умов наведені на рисунку 2.13.

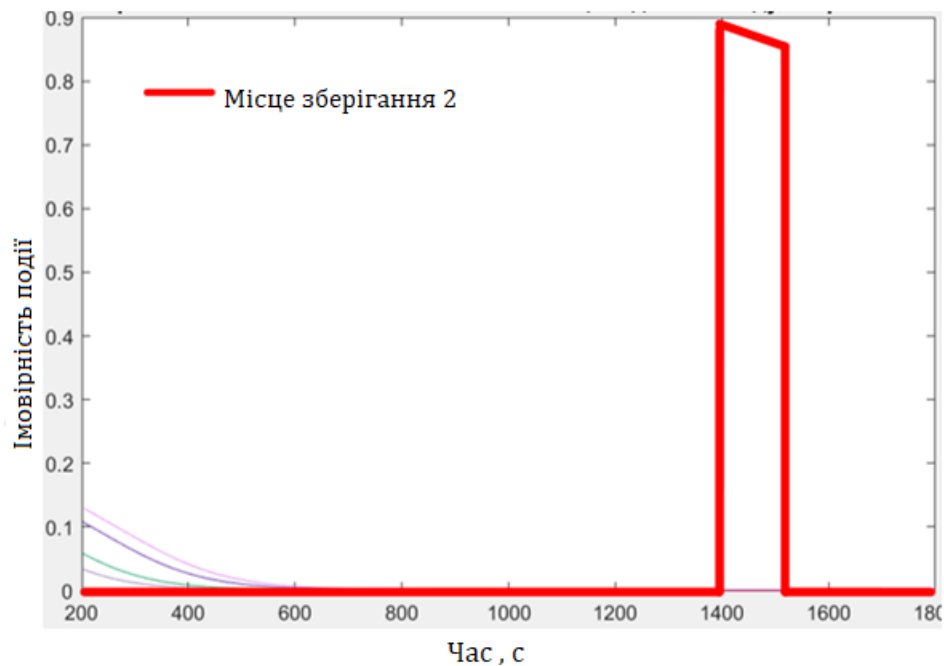


Рисунок 2.13 – Імовірність події «відмова HDD у Модулі Зберігання №2»

В результаті проведеного аналізу динаміки заповнення буферів обміну даними між Модулем Управління та Модулями Зберігання та застосування математичного апарату теорії нечітких систем з'являється можливість раннього діагностування ситуації появи дефектного диска в масиві жорстких дисків у процесі запису збереженої інформації в РСЗД в режимі реального часу, що дозволяє алгоритми мінімізації втрат загальної продуктивності РСЗД.

2.3 РСЗД як багатоканальна система масового обслуговування із відмовами

В результаті неодноразових випробувань РСЗД було зроблено висновок, що одним з основних показників РСЗД є продуктивність систем зберігання даних на профілі навантаження, що надається потенційним користувачем системи. У загальному випадку, профіль користувача навантаження являє собою інформацію про відсоткове співвідношення

операцій читання і запису з РСЗД, режиму читання/запису, розміру інформаційного блоку даних, тривалість тесту, а також містить вимоги за показниками продуктивності системи. Як правило, для забезпечення заявлених показників продуктивності потрібно виконати підбір мінімально необхідної архітектури системи зберігання даних та провести налаштування параметрів технологічних процесів обробки даних.

За способом функціонування даних відповідно до архітектури, наведеної на рисунку 1.3, розподілена система зберігання може бути класифікована як багатоканальна система масового обслуговування (СМО) з відмовами. Кількість каналів системи масового обслуговування визначається пропускнуою здатністю каналу зв'язку між клієнтом та системою зберігання даних та числом незалежних з'єднань, що встановлюються в процесі передачі даних каналом зв'язку. З чого випливає, що показники ефективності РСЗД може бути охарактеризовано через відповідні показники багатоканальної системи масового обслуговування з відмовами.

Теоретично багатоканальних систем масового обслуговування з відмовами використову.

У РСЗД з архітектурою, наведеною на рисунку 1.2, шляхом резервування певних потужностей системи можна досягти підвищення надійності інформаційного обміну потоків вищої пріоритетності за рахунок процесів із низькими показниками пріоритетності. При застосуванні цієї технології, РСЗД може розглядатися як сукупність багатоканальних систем масового обслуговування з відмовами у кожній групі пріоритетності. Варіювання параметрами потужностей, що резервуються, тобто зміною кількості каналів обслуговування системи масового обслуговування можна проводити її показники ефективності.

На рисунку 2.14 представлений графік зміни ймовірності того, що заявка отримає відмову

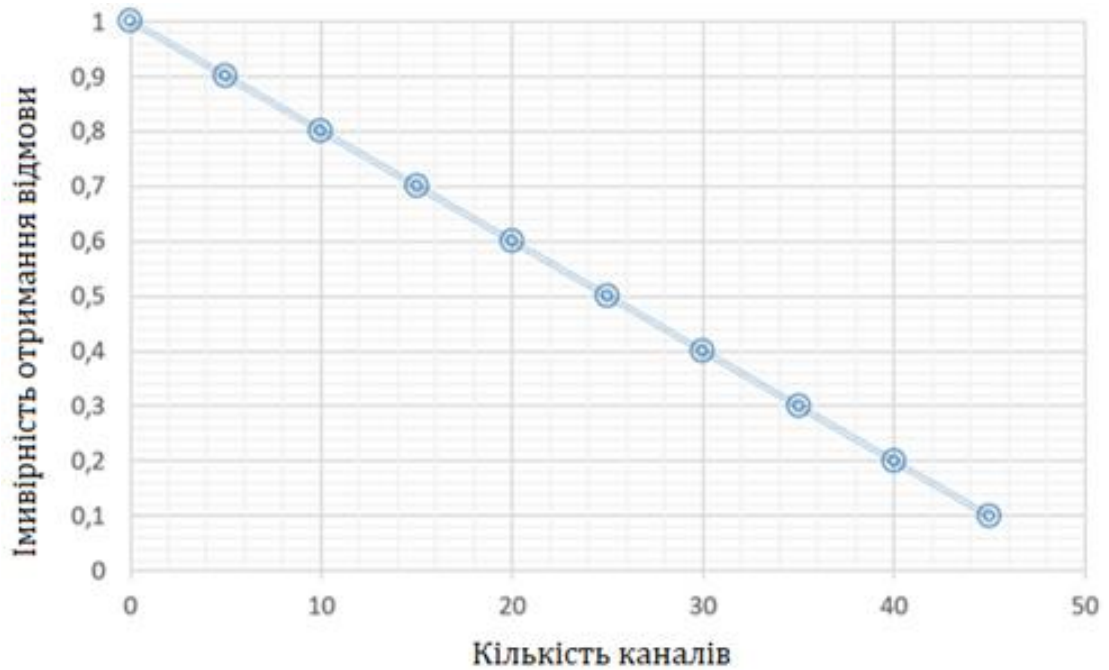


Рисунок 2.14 – Зміна параметрів ефективності системи масового обслуговування залежно від кількості каналів при фіксованих значеннях інтенсивності вхідного потоку заявок λ та інтенсивності потоку обслужених заявок, що виходить з одного каналу, μ

На рисунку 2.15 представлений графік зміни абсолютної пропускної спроможності системи масового обслуговування A , та середня кількість зайнятих каналів, k , при деяких фіксованих значеннях інтенсивності вхідного потоку заявок λ .

Абсолютна пропускна спроможність системи представлена Гбіт/с. Для перерахунку в іорс потрібно розділити розмір використовуваного блоку.

На рисунку 2.16 представлений графік зміни ймовірності відмови $^{otk}p_i$, при деякому фіксованому значенні інтенсивності вхідного потоку заявок λ для фіксованої кількості каналів n .

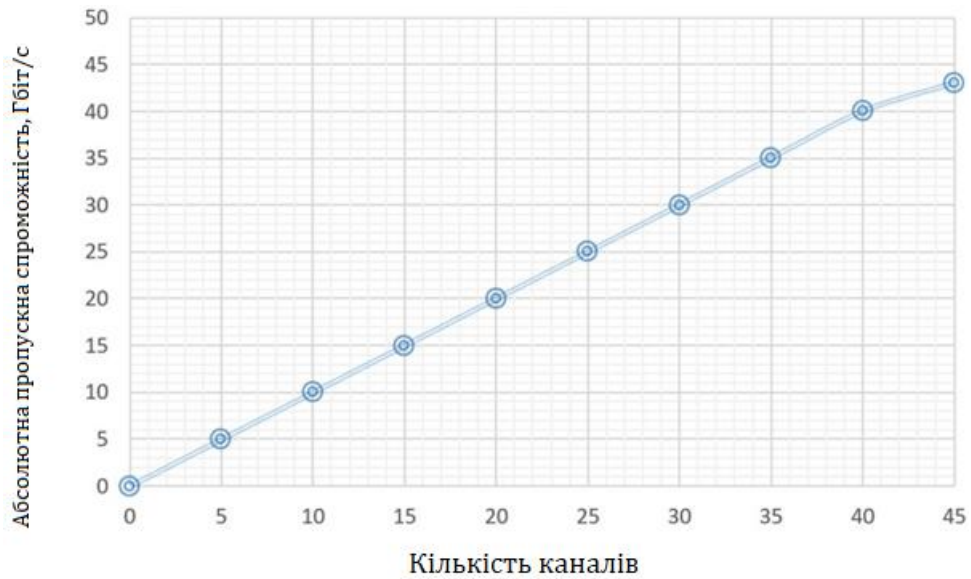


Рисунок 2.15 – Графік зміни абсолютної пропускної спроможності системи

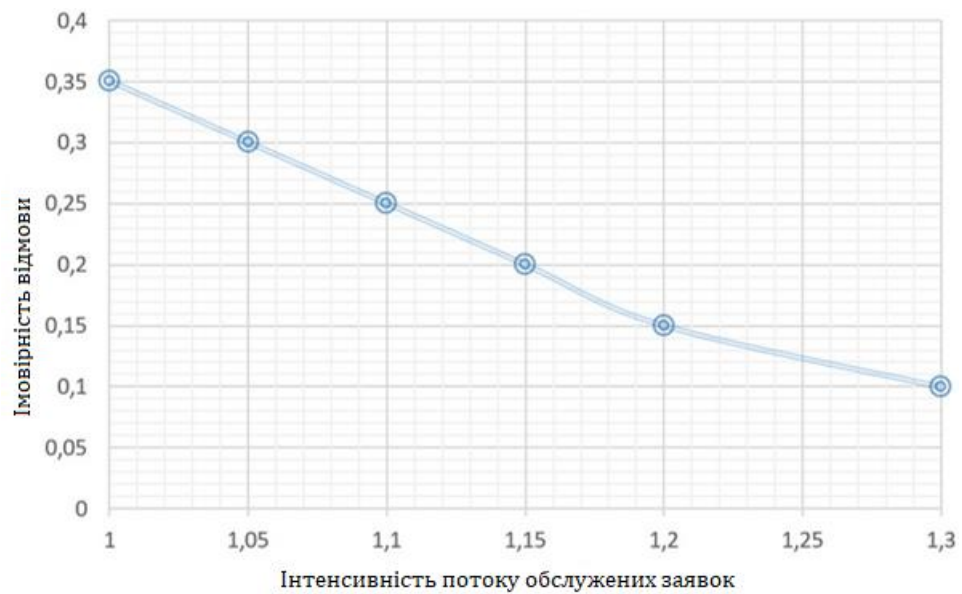


Рисунок 2.16 – Графік зміни ймовірності відмови за деякого фіксованого значення інтенсивності вхідного потоку заявок λ для фіксованої кількості каналів n

На рисунку 2.17 представлений графік зміни абсолютної пропускної спроможності системи масового обслуговування A , при деякому фіксованому значенні інтенсивності вхідного потоку заявок λ та фіксованому числі каналів n .

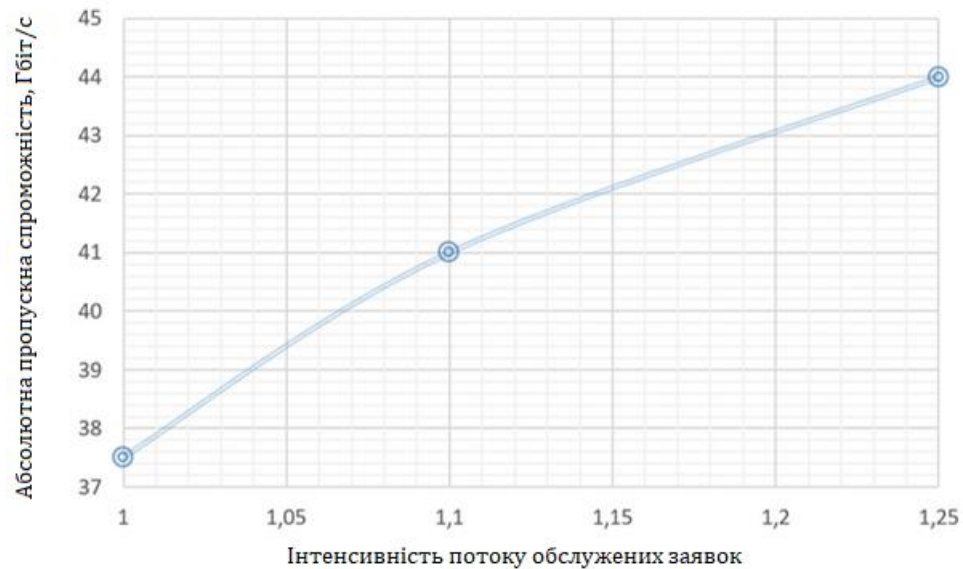


Рисунок 2.17 – Графік изменения абсолютной пропускная способности входящего потока заявок λ при фиксированном числе каналов n

Далі визначимо величину інтенсивності потоку обслужених заявок на 1 каналі, μ .

Згідно з формулою Ерланга для багатоканальної СМО з відмовами визначимо величини n_i , що визначають мінімально необхідну кількість незалежних потоків обслуговування. Виходячи з (2.3) можна обчислити за яких значеннях n_i виконується *откри*.

Визначивши відповідно до профілю користувача навантаження середню величину інтенсивності вхідного потоку заявок λ , на основі теорії багатоканальних систем масового обслуговування з відмовами можна отримати характеристики ефективності системи для певних груп пріоритетності шляхом оптимізації налаштувань пропускної спроможності

для виділеної групи пріоритетності та архітектурних рішень, що дозволяють впливати на величину інтенсивності потоку обслужених заявок.

2.4 Відповідність структури РСЗД відомому профілю навантаження

Досить часто на практиці ключовими показниками ефективності інформаційного обміну між бізнес-процесами є показники часу очікування між початком запиту на операцію та її завершенням (т.зв. та загальноприйнятий термін «latency») та сумарна продуктивність інформаційного обміну даними. Зазначені величини є базовими під час проектування конкретних бізнес-процесів. Істотні відхилення за вказаними параметрами є вкрай небажаними, оскільки можуть призвести до серйозної розбалансованості всього інформаційного обміну загалом.

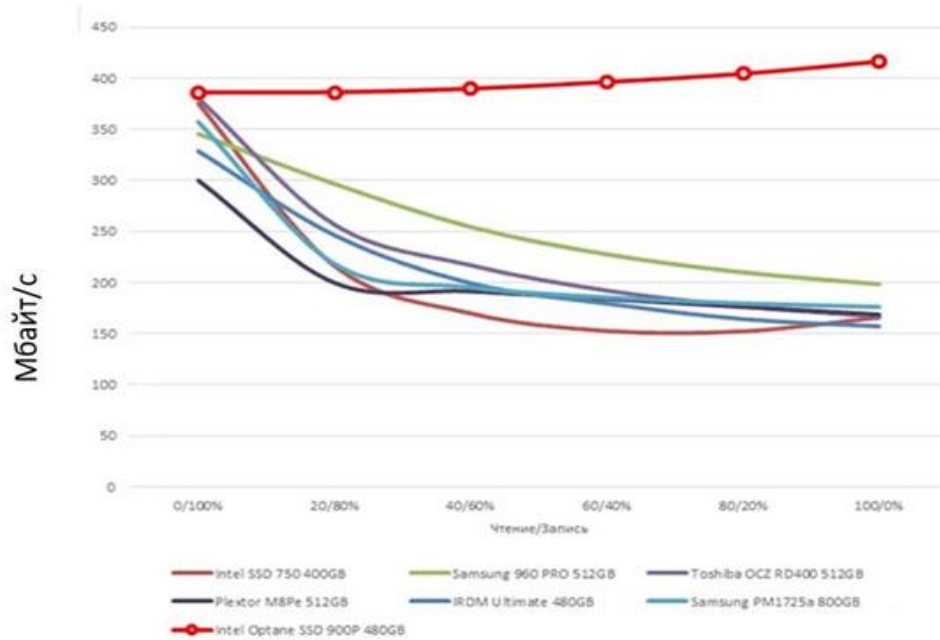
При формуванні профілю навантаження враховуються такі параметри:

- відсоткове співвідношення операцій читання;
- відсоткове співвідношення операцій запису;
- режими процесів читання/запису (послідовний чи випадковий);
- розміри інформаційного блоку даних;
- обсяг проведених операцій читання та запису.

Крім того, може визначити ряд додаткових параметрів, таких як, кількість ХОСТІВ, дедуплікацію та стиснення даних і пр. У сформованому профілі навантаження вказуються вимоги за величиною середнього та максимального часу очікування та мінімально необхідною величиною продуктивності інформаційного обміну.

Як показує n, найбільш узким місцем в інформаційному обміні високонавантаженої мережі передача даних є обмін даними з масивом жорстких дисків. При побудові РСЗД в першу чергу необхідно застосувати швидкісні ssd-диски, які в залежності від профілю навантаження мають різні показники швидкостей, відображені на рисунку

2.18. Крім цього, швидкість читання/запису на ssd-диск також залежить від кількості дисків у групі, залежно від графіка, наведеного на рисунку 2.19.



Профіль навантаження на зчитування та запису на SSD диск

Рисунок 2.18 – Показники швидкості SSD-дисків залежно від профілю навантаження

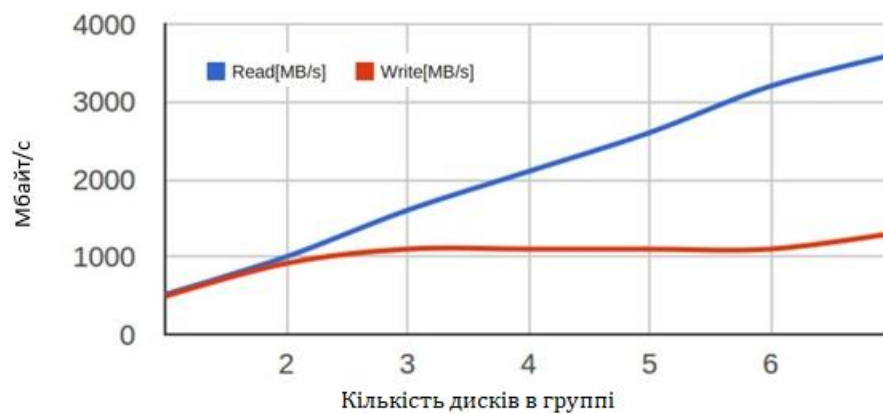


Рисунок 2.19 – Графік залежності швидкості читання/запису з ssd-диска від кількості дисків у групі

Пояснимо, що для операції запису час очікування складається з часу формування та надсилання запиту на встановлення з'єднання та часу встановлення з'єднання.

Для операції читання з РСЗД необхідною є процедура складання матриці кодування/декодування блоків даних, що зберігаються на жорстких дисках системи. Можна зробити висновок, що найбільший внесок у затримки становить час, необхідний для читання з жорстких дисків необхідної кількості блоків для складання матриці кодування/декодування. Далі, приймаючи, що з усіх дисків системи здійснюється у паралельному режимі, отримуємо пряму залежність часу очікування від кількості жорстких дисків у системі. Характер цієї залежності наведено на рисунку 2.20.

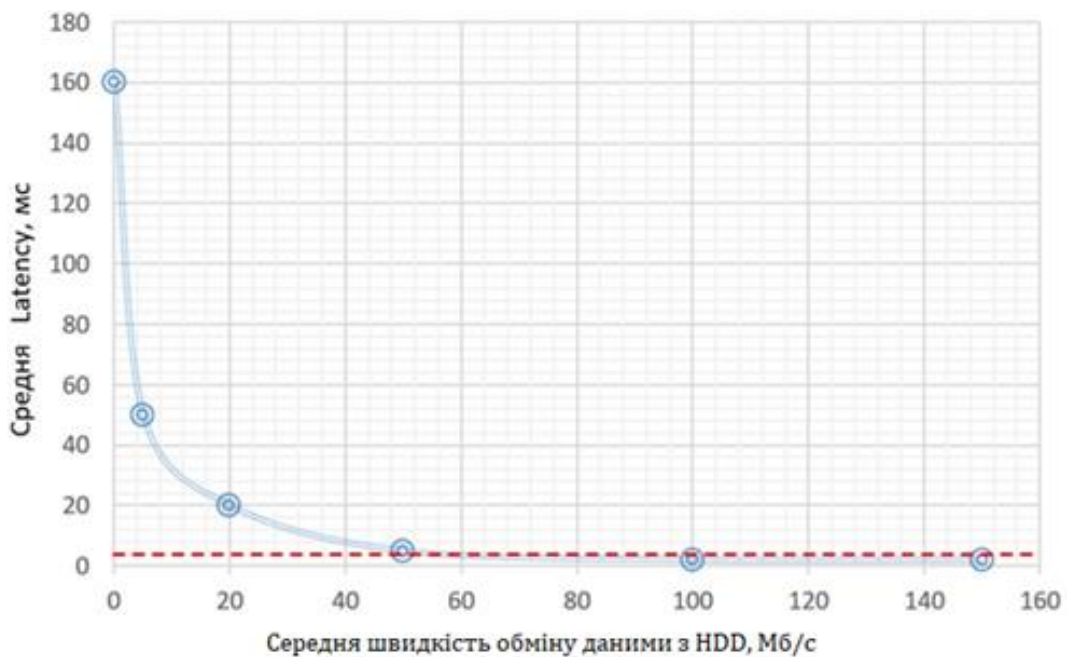


Рисунок 2.20 – Залежність часу очікування від середньої швидкості читання з жорсткого диска при 100 жорстких дисках

Крім часу, необхідного для читання з жорстких дисків блоків даних матриці кодування/декодування, на величину часу очікування також впливає

необхідність запуску процедури відновлення втрачених даних у матриці кодування/декодування. Однак запуск цієї процедури призводить до спонтанного збільшення часу очікування. Величина цього збільшення зазвичай враховується у профілі як максимально допустимий час очікування та накладає обмеження на показники продуктивності процедури відновлення втрачених даних, що враховуються у процесі розробки системи.

На основі аналізу графіка, представленого на рисунку 2.20, можна виділити мінімально необхідну середню швидкість обміну даними з масивом жорстких дисків, при якій досягається необхідний за профілем норматив часу очікування фіксованої кількості дисків в масиві. На базі математичної моделі можна побудувати криву залежності мінімально необхідної кількості жорстких дисків, необхідних для забезпечення середнього часу очікування профілю, від величини середньої швидкості читання блоків даних з жорсткого диска, яка наведена на рисунку 2.21.

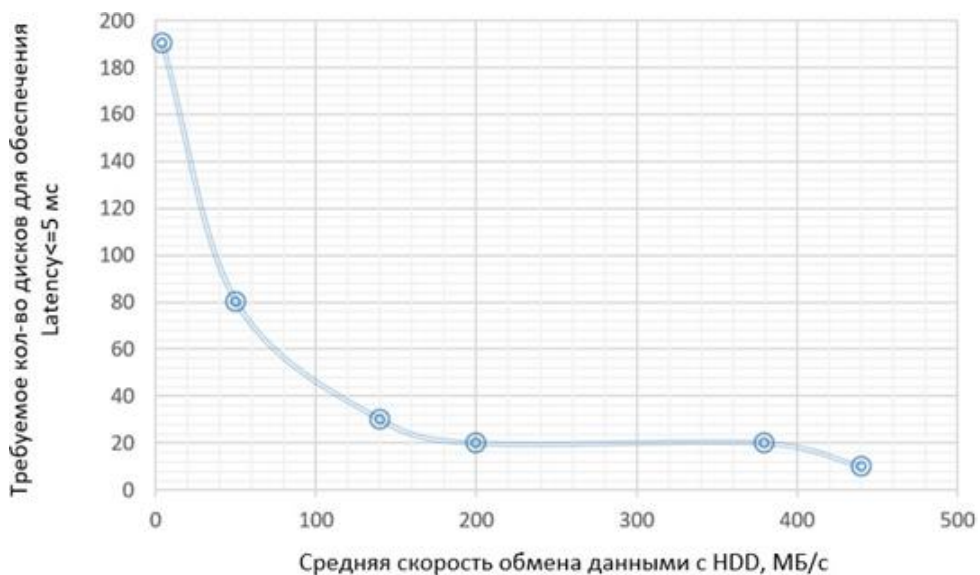


Рисунок 2.21 – Залежність мінімально необхідної кількості жорстких дисків від середньої швидкості читання блоків даних з жорсткого диска

Аналогічний підхід можна застосувати і параметр загальної продуктивності передачі між клієнтом і РСЗД з урахуванням фізичних

можливостей мережі передачі. Нагадаємо, що продуктивність процесу обміну даними залежить від швидкості проходження пакета даних, яка може змінюватись в деякому діапазоні. Графік залежності продуктивності обміну даними між клієнтом та РСЗД наведено на рисунку 2.22.

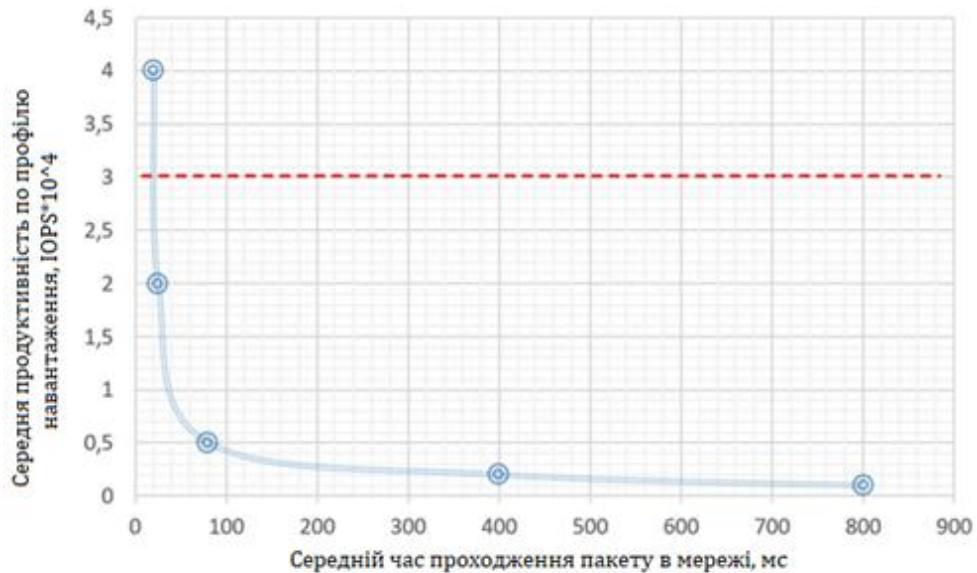


Рисунок 2.22 – Залежність продуктивності обміну даними між клієнтом та розподіленою системою зберігання даних зі 100 SSD від величини середнього часу проходження пакета

Збільшення продуктивності обміну даними при фіксованих показниках часу проходження можливе зі збільшенням кількості незалежних потоків передачі даних, у межах пропускної спроможності каналу зв'язку. Крива залежності мінімально необхідної кількості незалежних потоків передачі даних, необхідної для забезпечення загальної продуктивності обміну даними між клієнтом та системою від величини середнього часу проходження пакета даних, наведена на Рисунку 2.23.

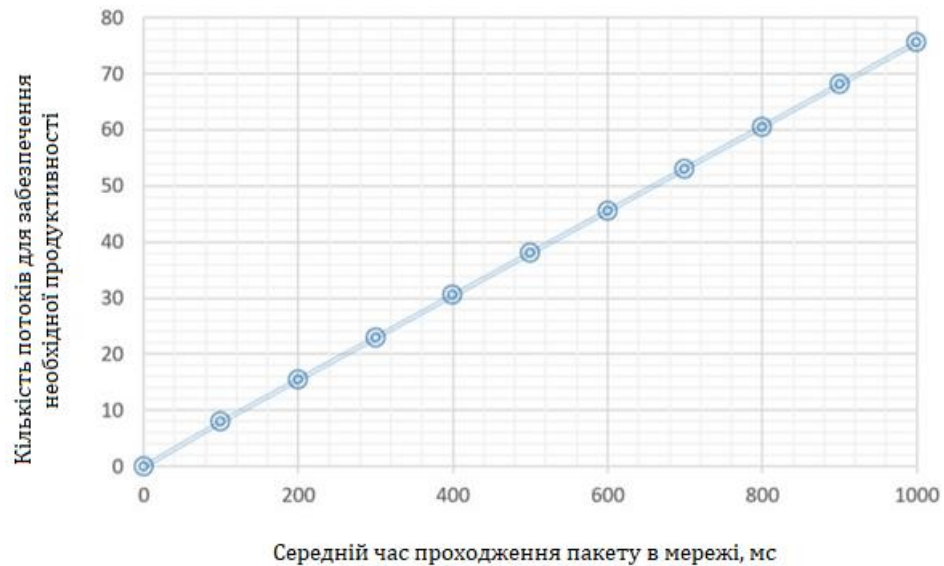


Рисунок 2.23– Залежність кількості незалежних потоків передачі даних від величини середнього часу проходження пакета даних при кількості чотирьох модулів зберігання

На підставі отриманих залежностей, наведених на рисунках 2.21 і 2.23, для відомого профілю навантаження цілком можливе проектування мінімально необхідної конфігурації РСЗД з однорідним масивом жорстких дисків фіксованого типу в мережі передачі даних з реально вимірними параметрами передачі пакета даних для забезпечення необхідних показників продуктивності часу очікування.

Важливо зазначити, що ефективність РСЗД визначається т.ч. використанням дискових носіїв з високою продуктивністю, наприклад, SSD-диски. Як метрика продуктивності процесів застосовують Мегабайт/секунду та/або іops (від англ. "in-out per second"). Прийmemo за одиницю виміру продуктивності інформаційного обміну даних – іops.

При проектуванні складу компонентів РСЗД одним з важливих параметрів є середня величина затримки в процесі читання блоків користувача з дисків.

ВИСНОВКИ

Розробка розподілених систем зберігання даних є перспективною для сучасних реалій, оскільки вона забезпечує високу надійність, ефективність доступу до даних і масштабованість. Впровадження механізмів реплікації з відповідною стратегією (синхронна, асинхронна, географічно розподілена) забезпечує високу доступність та надійність даних, що є ключовими факторами для досягнення стабільної та ефективної роботи розподілених систем зберігання даних у великих масштабах.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Білоконь А.С., Борисов С.О., Сальніков С.С., Федорченко В.М. Методи обробки та зберігання даних в розподілених системах // Системи управління, навігації та зв'язку. Полтава – 2024. № 1
2. Финн В.К. Об интеллектуальном анализе данных // Новости искусственного интеллекта, №3, 2004, С. 3-19.
3. W. Frawley, G. Piatetsky-Shapiro, C. Matheus Knowledge Discovery in Databases: An Overview. - AI Magazine. - 1992. pp. 213-228.
4. Kitchin Rob. The Data Revolution. United States: Sage. 2014, p. 6.
5. Huang L., Li J., Li X. InfoMall: A Large-Scale Storage System for Web Archiving
6. // In: Gao Y. et al. (eds) Web-Age Information Management. WAIM 2013. Springer, Berlin, Heidelberg. Lecture Notes in Computer Science. Vol. 7901. P. 89-98.
7. Kotoulas S., van Harmelen F., Weaver J. KR and Reasoning on the Semantic Web
8. // Web-Scale Reasoning. 2011. Handbook of Semantic Web Technologies. P. 441- 466.
9. Bouquet P., Palpanas T., Stoermer H., Vignolo M. A Conceptual Model for a Web- Scale Entity Name System // In: Gómez-Pérez A., Yu Y., Ding Y. (eds) The Semantic Web. ASWC 2009. Springer, Berlin, Heidelberg. Lecture Notes in Computer Science. Vol. 5926.
10. Качество обслуживания хранилища. URL: <https://docs.microsoft.com/ru-ru/windows-server/storage/storage-qos/storage-qos-overview> (дата обращения: 24.02.2022).
11. Sun J., Yu X., Wang R., Zhong N. A Model for Personalized Web-Scale Case Base Maintenance // In: Liu J., Wu J., Yao Y., Nishida T. (eds) Active Media Technology. AMT 2009. Springer, Berlin, Heidelberg. Lecture Notes in Computer

Science. Vol. 5820. P. 44.

12. Zhang B., Qu Y., Peng J., Fan J. An automatic image-text alignment method for large-scale web image retrieval // *Multimedia Tools and Applications*. 2017.