

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЗАПОРІЗЬКА ПОЛІТЕХНІКА»

Факультет комп'ютерних наук та технологій
Кафедра комп'ютерних систем та мереж

Пояснювальна записка

до дипломного проекту (роботи)

магістра

(ступінь вищої освіти)

на тему КОМП'ЮТЕРНА СИСТЕМА ДАТА ЦЕНТРУ З ДІАГНОСТУВАННЯМ
СТАНУ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ

Виконав: студент 2 курсу, групи КНТ-512м
спеціальності _____

123 Комп'ютерна інженерія

(код і найменування спеціальності)

Освітня програма (спеціалізація)

Комп'ютерні системи та мережі

ДЕМІЧЕВ А.В.

(ПРИЗВИЩЕ та ініціали)

Керівник СКРУПСЬКИЙ С.Ю.

(ПРИЗВИЩЕ та ініціали)

Рецензент САМОЙЛИК С.С.

(ПРИЗВИЩЕ та ініціали)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»
(повне найменування вищого навчального закладу)

Факультет комп'ютерних наук і технологій
Кафедра «Комп'ютерні системи та мережі»
Ступінь вищої освіти магістерський
Спеціальність 123 Комп'ютерна інженерія
(код і найменування)
Освітня програма (спеціалізація) Комп'ютерні системи та мережі
(назва освітньої програми (спеціалізації))

ЗАТВЕРДЖУЮ

Зав. кафедри Кудерметов Р.К.

“ 24 ” жовтня 2023 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ) СТУДЕНТА

ДЕМІЧЕВА Андрія Валерійовича

(прізвище, ім'я, по батькові)

1. Тема проекту (роботи) Комп'ютерна система даних центру з діагностуванням стану на основі нейронної мережі

Керівник проекту (роботи) СКРУПСЬКИЙ С.Ю., к. т. н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від 24 жовтня 2023 року № 400

2. Строк подання студентом проекту (роботи) 10 грудня 2023 року

3. Вихідні дані до проекту (роботи) 1000 вузлів. Процесори мають бути з'єднані високошвидкісною низько-латентною шиною. Структура Даних центру: 25% - обчислювальні вузли, 35% - вузли веб-хостингу, 35% - вузли віртуальних приватних серверів, 5% - резервні вузли. Забезпечити засоби діагностування стану вузлів системи

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

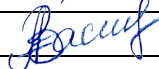
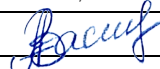
Аналіз технологій створення даних центрів

Проектування даних центру

Розробка діагностичної моделі та дослідження системи

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

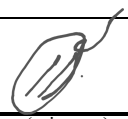
6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	прийняв виконане завдання
1-3	СКРУПСЬКИЙ С.Ю., к.т.н., доцент		
Нормоконтроль	ПОЛЬСЬКА О.В., ст.викладач		

7. Дата видачі завдання 29.09.2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Огляд та аналіз технологій створення дата центрів	09.10.2023	
2	Аналіз вимог до дата центру	11.10.2023	
3	Розробка інформаційної інфраструктури	25.11.2023	
4	Розробка телекомунікаційної інфраструктури	16.11.2023	
5	Розробка інженерної інфраструктури	21.11.2023	
6	Випробування системи	01.12.2023	
7	Аналіз засобів діагностування вузлів системи	04.12.2023	
8	Розробка діагностичної моделі	08.12.2023	
9	Оформлення отриманих результатів у ПЗ	09.12.2023	
10	Оформлення графічного матеріалу	10.12.2023	

Студент  Андрій ДЕМІЧЕВ
(підпис) (ім'я, ПРІЗВИЩЕ)

Керівник проекту (роботи)  Степан СКРУПСЬКИЙ
(підпис) (ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

ПЗ: 75 с., 13 рис., 11 таблиць, 18 джерел.

MIMD, MPP, INFINIBAND, XEON, ДАТА ЦЕНТР, ДІАГНОСТУВАННЯ, НЕЙРОННА МЕРЕЖА

Об'єкт дослідження – комп'ютерна система дата центру.

Предмет дослідження – нейронна мережа як засіб діагностування стану комп'ютерної системи.

Мета дослідження – підвищення ефективності використання дата центру за рахунок своєчасного діагностування стану його складових компонентів.

Для досягнення поставленої мети було виконано наступні задачі:

- проаналізовано технології побудови дата центрів;
- спроектовано інформаційну, телекомунікаційну та інженерну інфраструктури дата центрів;
- проаналізовано засоби побудови нейронних мереж;
- побудовано діагностичну модель, яка дозволяє діагностувати стан дата центру.

В результаті виконання роботи було розроблено комп'ютерну систему дата центру а також нейромережеву діагностичну модель.

ABSTRACT

Explanatory note : 75 pages, 13 figures, 11 tables, 18 references.

DATA CENTER, DIAGNOSIS, MIMD, MPP, NEURAL NETWORK,
INFINIBAND, XEON

The object of the research is the computer system of a data center.

The subject of the research is a neural network as a means of diagnosing the state of a computer system.

The aim of the research is to enhance the efficiency of data center utilization through timely diagnosis of the state of its constituent components.

To achieve the stated goal, the following tasks were undertaken:

- analyzed the technologies for constructing data centers;
- designed the information, telecommunication, and engineering infrastructures of data centers;
- analyzed the tools for building neural networks;
- developed a diagnostic model capable of diagnosing the state of the data center.

As a result of the research, a computer system for the data center and a neural network diagnostic model have been developed.

ЗМІСТ

Вступ.....	7
1 Аналіз технологій створення дата центрів	8
1.1 Опис предметної області	8
1.2 Аналіз технічного завдання.....	11
1.3 Огляд багатопроцесорних архітектур	13
1.4 Огляд технологій побудови комп'ютерних мереж в дата центрах.....	21
1.5 Огляд засобів захисту даних в датацентрах	22
1.6 Висновки за розділом.....	24
2 Проєктування дата центру.....	24
2.1 Розробка інформаційної інфраструктури	25
2.2 Розробка телекомунікаційної інфраструктури.....	31
2.3 Розробка інженерної інфраструктури	41
2.4 Налаштування операційних систем.....	48
2.5 Налаштування програмного забезпечення	52
2.6 Випробування кластера	59
2.7 Висновки за розділом.....	62
3 Розробка діагностичної моделі та дослідження системи	63
3.1 Вибір типу діагностичної моделі.....	63
3.2 Побудова моделі	65
3.3 Висновки за розділом.....	70
Висновки	72
Перелік джерел посилання	74

ВСТУП

У сучасному світі технологій, датацентри або центри обробки даних виступають як основа нашої цифрової ери, сприяючи зберігання, обробці та поширенню величезних обсягів інформації. З нескінченним зростанням цифрових даних попит на ефективну та надійну інфраструктуру центрів обробки даних стрімко зростає, що вимагає інноваційних підходів до проектування, управління та обслуговування. Паралельно з цим, галузь штучного інтелекту, зокрема нейронні мережі, зазнала вражаючих досягнень, пропонуючи нові можливості для покращення функціональності та діагностування систем.

В магістерській роботі ставиться мета розробити комп'ютерну систему дата центра і забезпечити можливість діагностування його стану за допомогою нейронної мережі. Для досягнення поставленої мети необхідно вирішити наступні завдання:

- проаналізувати технології побудови дата центрів;
- спроектувати інформаційну, телекомунікаційну та інженерну інфраструктуру дата центрів;
- проаналізувати засоби побудови нейронних мереж;
- побудувати діагностичну модель, яка дозволяє діагностувати стан дата центру.

Пояснювальна записка до магістерської роботи включає вступ, три розділи, висновки та перелік використаних джерел.

1 АНАЛІЗ ТЕХНОЛОГІЙ СТВОРЕННЯ ДАТА ЦЕНТРІВ

В цьому розділі аналізується технічне завдання, виконується огляд сучасних комп'ютерних архітектур, які можуть бути використані в датацентрах, обирається мережева технологія, розглядаються засоби забезпечення захисту даних в дата центрі.

1.1 Опис предметної області

У сучасному світі технологій, датацентри або центри обробки даних виступають як основа нашої цифрової ери, сприяючи зберіганню, обробці та поширенню величезних обсягів інформації. З нескінченним зростанням цифрових даних попит на ефективну та надійну інфраструктуру центрів обробки даних стрімко зростає, що вимагає інноваційних підходів до проєктування, управління та обслуговування. Паралельно з цим, галузь штучного інтелекту, зокрема нейронні мережі, зазнала вражаючих досягнень, пропонуючи нові можливості для покращення функціональності та діагностування систем.

Центри обробки даних, часто називані "цифровими нервовими центрами" сучасного світу, відіграють важливу роль у підтримці різних секторів, таких як фінанси, охорона здоров'я, зв'язок та електронна комерція [1]. Їхнє проєктування та управління впливають на глобальну економіку та суспільство в цілому. Однак стрімке зростання цифрових послуг призвело до збільшеного споживання енергії, операційних складностей та можливих обмежень продуктивності в центрах обробки даних. Для вирішення цих викликів все більше зростає потреба у нових стратегіях, які можуть покращити ефективність, надійність та сталість центрів обробки даних. Інтеграція нейронних мереж, підгрупи штучного інтелекту, має величезний потенціал у революціонізації операцій центрів обробки даних,

дозволяючи робити розумне прийняття рішень, передбачувальне обслуговування та адаптивний розподіл ресурсів.

Центри обробки даних відіграють фундаментальну роль у широкому спектрі завдань у різноманітних галузях, сприяючи цифровій трансформації, інноваціям та забезпеченню безперервності доставки послуг та додатків. Можливості датацентрів використовуються у широкому спектрі сфер інформаційних технологій [1].

У сфері хмарних обчислень та віртуалізації забезпечується створення та керування віртуальними машинами, розподіл ресурсів за запитами та ефективне використання апаратних ресурсів. Центри обробки даних дозволяють підприємствам розгортати застосунки, масштабувати ресурси та здійснювати зберігання даних у гнучкий та економічно ефективний спосіб.

У контексті обробки великих даних, датацентри сприяють зберіганню, доступу та обробці великих масивів даних, дозволяючи організаціям отримувати цінні висновки, приймати рішення на основі даних та виявляти закономірності, що впливають на стратегічне планування.

В області електронної комерції та онлайн-роздрібної торгівлі центри обробки даних підтримують функціонування електронних платформ, надаючи хостинг для онлайн-магазинів, управління транзакціями та обслуговування запасів. Вони забезпечують безпечну обробку платежів, ефективну взаємодію зі споживачами та оновлення інвентарю в реальному часі.

У сфері фінансових послуг та торгівлі, центри обробки даних мають важливе значення для високочастотної торгівлі та фінансової аналітики, забезпечуючи швидке виконання транзакцій та обробку даних ринку в реальному часі. Вони надають обчислювальну потужність для складних фінансових моделей та алгоритмічних торгових стратегій.

У контексті соціальних мереж, які генерують великі обсяги вмісту, взаємодій та даних, створених користувачами, центри обробки даних дозволяють зберігати, управляти та аналізувати ці дані, сприяючи персоналізації вмісту, цільовій рекламі та покращенню взаємодії з користувачами.

У сфері відеострімінгу та онлайн-ігор центри обробки даних є незамінними для служб відеострімінгу та платформ онлайн-ігор, які надають високоякісний відеоконтент та інтерактивні ігрові враження користувачам по всьому світу, вимагаючи низької затримки та високої пропускної здатності.

У контексті Інтернету Речей, розповсюдження пристроїв Інтернету Речей генерує значний обсяг даних сенсорів, які центри обробки даних обробляють та аналізують, дозволяючи реалізовувати ініціативи розумних міст, промислову автоматизацію та передбачувальне обслуговування у різних секторах.

Діагностика стану датацентру під час його експлуатації набуває важливого значення для забезпечення безперервної та ефективної роботи. У контексті стрімкого зростання обсягів та значення даних, а також збільшення складності обладнання та програмного забезпечення, діагностична діяльність виступає як ключовий інструмент забезпечення високої продуктивності та надійності. Систематичне та своєчасне діагностування сприяє ідентифікації та вирішенню потенційних проблем, які можуть призвести до збоїв або відмов, тим самим дозволяючи здійснити вчасні виправлення та запобігти серйозним наслідкам.

Крім того, діагностика сприяє плануванню обслуговування та ремонту, передбачаючи терміни заміни компонентів та потенційний ремонт, що зменшує час простою та позитивно впливає на продуктивність та доступність системи. Також, на основі діагностичних даних розробляються прогнози щодо технічного стану та витрат на обслуговування, що допомагає раціонально розподіляти ресурси та планувати бюджет.

У цьому контексті діагностування стану датацентру в процесі експлуатації стає вирішальним для забезпечення надійності, ефективності та безпеки його функціонування, сприяючи оптимальному використанню ресурсів у відповідності до вимог сучасного цифрового середовища.

Розробка комп'ютерної системи датацентру з діагностуванням його стану на основі нейронної мережі є основною метою дипломної роботи. Для досягнення цієї мети необхідно провести аналіз сучасних паралельних та розподілених комп'ютерних архітектур, на підставі якого обрати оптимальну основу для

датацентру, виконати проектування інформаційної, телекомунікаційної та інженерної інфраструктури датацентру, а також провести дослідження стану датацентру та розробити модель його діагностування з використанням нейронних мереж.

1.2 Аналіз технічного завдання

Планується створення комп'ютерної системи дата-центру, що складатиметься з тисячі вузлів, забезпечених відповідними ресурсами обчислювального потенціалу, зберігання та хостингу. Обчислювальний потенціал передбачає використання багатоядерних процесорів із значним об'ємом оперативної пам'яті (мінімум тридцять два гігабайти на кожен вузол), з'єднаних високошвидкісною низько-латентною шиною для ефективного обміну даними між вузлами. Ресурси зберігання мають забезпечувати можливість зберігання не менше одного петабайта даних із резервуванням аналогічного обсягу. Ресурси хостингу повинні дозволити клієнтам розгорнути веб-сервіси та забезпечити хостинг веб-сайтів.

Передбачається, що дев'яносто відсотків ресурсів дата-центру будуть надаватися в оренду в рамках програм віртуальних серверів, виділених серверів та виділених стійок. В дата-центрі планується використання тільки провідних сегментів мережі, поділених на інфраструктурні та інформаційні.

Структура датацентру передбачає розміщення 25% обчислювальних вузлів, 35% вузлів веб-хостингу, 35% віртуальних приватних серверів та 5% резервних вузлів. Всі вузли розміщуються у стійках, що займають весь виділений простір. Для зв'язку між вузлами планується створення інфраструктурної мережі на базі технології 10G Ethernet. Пограничні роутери забезпечуватимуть двосторонній доступ вузлів до зовнішньої мережі Інтернет з пропускнуою здатністю не менше 10 Гбіт/с. Управління базою даних облікових записів, політиками та

адмініструванням мережі планується здійснювати за допомогою контролера домена.

Обчислювальні ресурси повинні забезпечити надійне функціонування апаратно-програмних засобів на сучасному рівні, забезпечуючи тривалий термін експлуатації обладнання (не менше дев'яти років), після чого обладнання повинне бути готовим до модернізації.

Програмне забезпечення має відповідати сучасним вимогам надійності та безпеки, дозволяючи використання операційних систем як від Microsoft, так і розповсюджуваних за ліцензією GNU. На всіх вузлах повинні бути встановлені та системи комплексного захисту від шкідливих програм та хакерських атак.

Вимоги до обчислювальних вузлів комп'ютерів включають низькопрофільний форм-фактор, материнську плату з підтримкою двох сучасних процесорів та вбудованим відеоадаптером. Процесори мають бути 16-ядерні з 32 потоками, тактовою частотою не менше 2.8 ггц, виконувати не менше двох потоків на ядро та мати кеш не менше 20 мб. Оперативна пам'ять має становити не менше 32 гб, а накопичувач - не менше 10 тб.

Для мережі висуваються вимоги наявності інформаційної мережі з пропускною здатністю не менше 10 гбіт/с, перспективи модернізації, резервування критичних комутаторів та апаратної системи безпеки.

Система безперебійного живлення повинна активно забезпечувати захист від перепадів напруги, відмов та відключень, забезпечувати роботу від акумуляторів протягом 30 хвилин, бути придатною для монтажу в стійку 19" та мати розподільник живлення до всіх вузлів стійки.

Система контролю навколишнього середовища має забезпечувати підтримку заданої температури та вологості повітря у будь-яку пору року та в широкому діапазоні температур, бути прецизійною та мати резервування установок.

Електричні установки, кабелі та засоби заземлення повинні відповідати діючим нормативним актам, а електроживлення здійснюватися за допомогою

сертифікованих засобів знешкодження мережеских перешкод відповідно до вимог безпеки.

1.3 Огляд багатопроцесорних архітектур

Ключовою системою класифікації багатопроцесорних систем є класифікація Флінна, яка викликала появу нових класифікацій, які уточнили початкові класи цієї системи [2]. Флінн впровадив дві основні концепції: концепцію потоку команд та концепцію потоку даних. Термін "потік команд" зручно інтерпретувати як послідовність команд, які виконуються в рамках однієї програми. В свою чергу, "потік даних" можна вважати послідовністю даних, що обробляються у рамках одного потоку команд. Виходячи з цих визначень, Флінн запропонував класифікацію паралельних обчислювальних систем, що базується на принципі грубозернистої паралелізації.

Системи з одиночним потоком команд та одиночним потоком даних, або SISD, являють собою стандартні послідовні обчислювальні системи. В таких системах один потік інструкцій обробляє один потік даних, що зображено на рисунку 1.1.

Системи з одиночним потоком команд та багатозначним потоком даних, або SIMD, характеризуються виконанням однієї програми, де кожна інструкція обробляє багато даних одночасно. Це означає, що одиночний потік інструкцій опрацьовує багатозначний потік даних, що відповідає векторному типу паралелізації, що зображено на рисунку 1.2.

Найпопулярніша концепція у класі SIMD - це векторна обробка. Векторний процесор підтримує опрацювання не тільки окремих скалярних, але і векторних операндів. Ефективне розкодування інструкцій та зручні дані позитивно впливають на продуктивність. Тому векторну обробку використовують і у процесорах класів SISD та MIMD. Наприклад, розширення SIMD IA32 - це

технології MMX та SSE. В якості провідного процесора використовується 64-бітний двопотоковий RISC-процесор від IBM, який традиційно має векторне розширення.

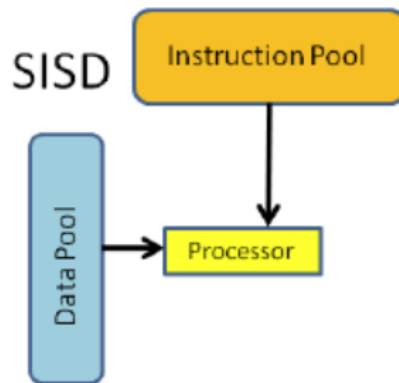


Рисунок 1.1 – Архітектура одиночний потік команд - одиночний потік даних

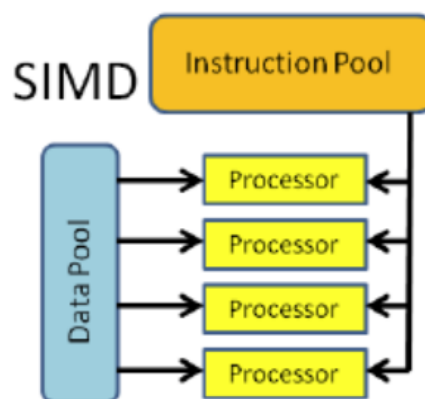


Рисунок 1.2 – Архітектура одиночний потік команд - багатозначний потік даних

Відомою представницею класу SIMD є архітектура VLIW (дуже довге слово команди). Одна інструкція в такій системі представляє собою набір декількох RISC-інструкцій, які не залежать одна від одної за даними. VLIW-процесору не потрібно перевіряти інструкції на наявність структурних залежностей, залежностей за даними чи управлінням [2]. Це завдання тепер лежить на плечах компілятора. Процесор може відразу переходити до фази виконання. Елементи, такі як динамічний планувальник, станції резервування тощо, в цій архітектурі

відсутні. Звільнені ресурси (транзистори) перерозподіляються для підвищення продуктивності системи.

Наступною архітектурою розглянемо багатозначний потік команд - одиночний потік даних або MISD. Тут кілька команд одночасно взаємодіють з одним елементом даних (рис. 1.3). Цей клас поки що не має практичного застосування, але це не означає, що схема є недосконалою.

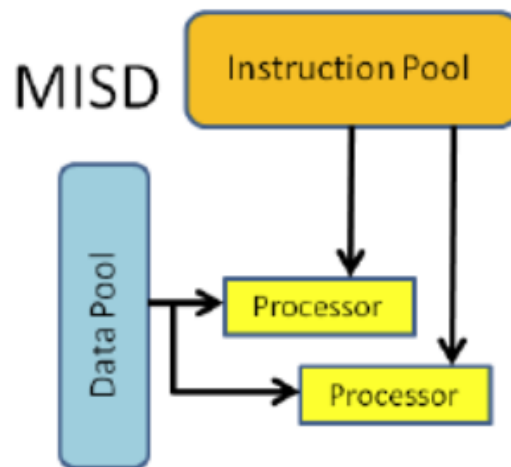


Рисунок 1.3 – Архітектура одиночний потік команд - багатозначний потік даних

Архітектура багатозначний потік команд - багатозначний потік даних або MIMD. В таких системах одночасно та незалежно від інших виконуються кілька програмних гілок, які в певні інтервали часу обмінюються даними (рис. 1.4). Зазвичай ці системи називаються багатопроцесорними [2].

Представники класу MIMD поділяються на системи з об'єднаною (тісно зв'язаною) та розподіленою (слабко зв'язаною) пам'яттю. Системи з об'єднаною пам'яттю (мультіпроцесори) конструюються шляхом збільшення кількості процесорів у системі. Така мультіпроцесорна система симетрична з точки зору процесорів. Вона дозволяє відновити завдання, яке було виключено раніше, на будь-якому з доступних процесорів. Операційна система розглядає кожен процесор як загальний ресурс, який розподіляється між різними завданнями або потоками виконання.

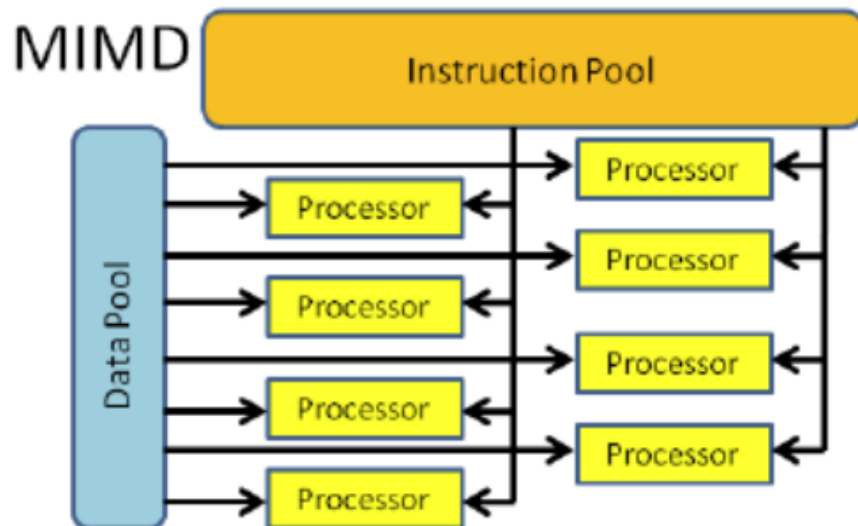


Рисунок 1.4 – Багатозначний потік команд - багатозначний потік даних

Мінімізація перемикань процесорів для виконуваних програм допомагає оптимально використовувати ресурси кожного процесора. Ця симетрична мультіпроцесорна архітектура (Symmetric Multiprocessor, SMP) підтримується сучасними операційними системами. Зміни в кількості, відповідно до закону Мура, дозволили розміщувати кілька процесорів на одному чипі. Багатоядерний (Multi-Core) процесор з точки зору операційної системи представляє собою набір одноядерних процесорів, які дозволяють симетричне виконання завдань. В інших випадках різні ядра можуть спільно використовувати кеш другого або третього рівнів. Мультикомп'ютерні системи передбачають розподілену пам'ять. Вони складаються з потужних серверів, об'єднаних у єдиний ресурс завдяки швидкій комунікаційній мережі (Massively Parallel Processing, MPP). Кластерні системи представляють спрощену версію MPP. Ця запропонована система класифікації залишається найбільш популярною для початкової характеристики комп'ютерів. Якщо ми заявляємо, що комп'ютер належить до класу SIMD або MIMD, це відразу надає базового розуміння принципів його роботи, і, в деяких випадках, цього вистачає. Отже, з точки зору дата-центру, найперспективнішою багатопроцесорною архітектурою є MIMD.

Симетричні мультіпроцесорні системи, або SMP системи, є складними структурами, що включають в себе низку однорідних процесорів та спільного

пула пам'яті. Кожен процесор має рівний доступ до будь-якої області пам'яті з однаковою швидкістю [3]. Крім того, апаратно забезпечується когерентність кешів, що дозволяє уникнути неузгодженості даних. Процесори можуть бути підключені до пам'яті через спільну шину або crossbar-комутатор. Останній, використовуючи принцип локальності програм, має за мету побудову паралельних з'єднань між процесорами та блоками пам'яті. Зазвичай, в кожен момент часу кожен процесор має доступ лише до певного фрагмента загальної пам'яті. Наявність спільної пам'яті спрощує взаємодію між процесорами, але одночасно на них накладаються обмеження щодо їхньої кількості. У реальних системах зазвичай встановлюється не більше 32 ядер (зазвичай багатопотокові).

Операційна система, що використовується в SMP системах, спрямована на розподіл процесів між процесорами. Програми можуть легко переноситись з завантаженого процесора на менш завантажений. Багатопотокова програма динамічно отримує доступ до певної кількості процесорів в залежності від завантаження системи в цілому. Планувальник операційної системи легко розподіляє високе обчислювальне навантаження. Спільною рисою цих систем є високі витрати на їхнє обладнання та обмежена масштабованість.

Масивно-паралельні системи, або MPP системи, будується на основі однорідних обчислювальних вузлів з кількома центральними процесорами, власною локальною пам'яттю та комунікатором. Між вузлами встановлюється комунікаційна мережа. Загальна кількість процесорів у реальних системах може сягати кількох сотень тисяч. Операційна система в MPP надає повну функціональність лише на управляючому вузлі, підтримуючи лише базовий функціонал для паралельних програм на кожному іншому вузлі. Часта передача обчислювального навантаження між вузлами в MPP не рекомендується, оскільки це може призвести до повільних процедур зовнішнього введення/виведення. Фактично, це не один могутній ресурс, а об'єднання ресурсів для розв'язання певних завдань, що вимагає спеціального програмування. Такі вимоги до спеціальної організації обчислень змушують змінювати наявне програмне забезпечення або створювати нове, що незаперечно ускладнює перехід від SMP до

MPP. Головною перевагою MPP є висока масштабованість, обмежена лише параметрами комунікаційної мережі.

Основна відмінність між системами SMP (Symmetric Multiprocessing) та MPP (Massively Parallel Processing) полягає в характері адресної системи. В системах зі спільною пам'яттю адресний простір всіх процесорів є єдиним, і коли декілька процесорів використовують одну змінну X , вони звертаються до тієї ж самої фізичної ячейки спільної пам'яті [3]. Така організація має як переваги, так і недоліки.

Наявність спільної пам'яті дозволяє уникнути фізичного переміщення даних між програмами, що паралельно виконуються на різних процесорах, спрощуючи програмування та знижуючи часові витрати на міжпроцесорний обмін.

Водночас, для виконання команд кожним процесором потрібно звертатися до спільної пам'яті, що ставить високі вимоги до пропускної здатності комутатора цієї пам'яті і обмежує максимальну кількість процесорів у системі зі спільною пам'яттю.

Окрім того, можливість одночасного звернення кількох процесорів до спільних даних може призвести до отримання некоректних результатів. Це вимагає реалізації механізмів синхронізації паралельних процесів, що ускладнює роботу операційних систем.

У системах з індивідуальною пам'яттю кожен процесор має незалежний адресний простір, і коли одна і та ж змінна X використовується в програмах різних процесорів, вони звертаються до фізично відмінних ячеек індивідуальної пам'яті цих процесорів. Це призводить до фізичного переміщення даних між взаємодіючими програмами на різних процесорах. Однак, оскільки більшість обращень до пам'яті здійснюється кожним процесором власної пам'яті, вимоги до комутатора послаблюються, і кількість процесорів у системах з розподіленою пам'яттю може досягати сотень або навіть тисяч.

Системи з неоднорідним доступом до пам'яті (NUMA). Така система складається з однорідних базових модулів, які включають декілька процесорів та блок пам'яті. Модулі з'єднані високошвидкісним комутатором. Вона підтримує

спільний адресний простір і має апаратну підтримку доступу до віддаленої пам'яті, тобто пам'яті інших модулів. Затримки при обробці процесором "своїї" пам'яті є невисокими, особливо у порівнянні з SMP-системами. Проте доступ до "чужої" пам'яті, належної іншому процесору, супроводжується вищими затримками.

Термін "неоднорідність" організації пам'яті має тут своє коріння. При правильній організації доступу до пам'яті, коли кожен процесор працює з даними лише у "своїй" пам'яті, така схема буде вигідною порівняно з класичним SMP-рішенням завдяки відсутності обмежень пропускну здатності загальної системної шини. У цьому випадку загальна пікова пропускну здатність підсистеми пам'яті буде дорівнювати подвійному значенню пропускну здатності використовуваних модулів пам'яті.

Справедлива організація доступу до пам'яті – це ключове та критично важливе поняття. Платформи з архітектурою NUMA повинні підтримуватися як зі сторони операційної системи, так і зі сторони додатків. У випадку, коли апаратно підтримується когерентність кешів у всій системі (зазвичай це так), говорять про архітектуру cc-NUMA (cache-coherent NUMA). Масштабованість NUMA-систем обмежується обсягом адресного простору, можливостями апаратної підтримки кешів та можливостями операційної системи управляти великою кількістю процесорів. На сьогоднішній день максимальна кількість процесорів у NUMA-системах становить 1024.

Паралельні векторні системи (PVP). Головною рисою систем PVP є наявність спеціальних векторно-конвеєрних процесорів, які мають команди однотипної обробки векторів незалежних даних, що ефективно виконуються на функціональних пристроях конвеєрів [3]. Зазвичай, кілька таких процесорів (1–16) одночасно працюють зі спільною пам'яттю (аналогічно SMP) в межах багатопроцесорних конфігурацій. Кілька таких вузлів можуть бути об'єднані за допомогою комутатора (аналогічно MPP). Паралельні векторні системи об'єднують в собі переваги SMP і MPP. Однак PVP – це все ж спеціалізована

архітектура, яку неможливо отримати шляхом масштабування звичайних комп'ютерів, як в SMP або MPP.

Кластерні системи. При симетричному мультипроцесорному підході процесори об'єднуються за допомогою спільної пам'яті, що суттєво спрощує програмування, але ставить високі вимоги до продуктивності самої пам'яті. Дешевшим і добре масштабованим підходом (порівняно з SMP) є побудова мультикомп'ютерної системи з розподіленою пам'яттю. Такий мультикомп'ютер є масивом незалежних спеціалізованих обчислювальних модулів, об'єднаних в єдиний обчислювальний ресурс за допомогою високопродуктивної комунікаційної мережі зі спеціалізованими каналами зв'язку. І обчислювальні модулі, і канали зв'язку створюються виключно під конкретний суперкомп'ютер.

Ідея створення так званого кластера робочих станцій стала розвитком підходу MPP. Концепція кластера робочих станцій сформувалася наприкінці 1990-х років завдяки розвитку шини PCI (PCI-X, PCI-E), появі гігабітного Ethernet тощо. Звичайні персональні комп'ютери, що були об'єднані локальною мережею, почали наздоганяти спеціалізовані системи MPP за комунікаційними можливостями. Таким чином, з'явилася можливість побудувати повноцінну систему MPP зі звичайних універсальних робочих станцій за допомогою стандартних комунікаційних технологій. Причому вартість такого кластера була в середньому на два порядки менше. На сьогоднішній день кластер складається з процесорних елементів (стандартних процесорів), з'єднаних високошвидкісною мережею передачі даних [2].

Отже, після аналізу сучасних багатопроцесорних архітектур для обчислювальних вузлів дата центру виявилось, що найкращим варіантом є архітектура MPP. Її переваги у масштабованості та можливості незалежної обробки даних кожним вузлом вписуються у концепцію Центру обробки даних. При цьому кожний обчислювальний вузол, згідно з класифікацією Флінна, представляє собою MIMD-систему.

1.4 Огляд технологій побудови комп'ютерних мереж в дата центрах

Сучасні дата центри, що здійснюють великі обчислювальні завдання та забезпечують обробку великих обсягів даних, вимагають потужних та швидкодіючих мережевих інфраструктур. Оптимізація продуктивності, масштабованості та надійності мереж визначає ефективність діяльності дата центру.

10G Ethernet та InfiniBand є ключовими технологіями, що забезпечують високошвидкісний обмін даними в дата центрах. 10G Ethernet заснований на стандарті IEEE 802.3ae та дозволяє передавати дані зі швидкістю 10 гігабіт на секунду. Ця технологія заснована на протоколі Ethernet, що робить її сумісною з існуючими мережами. 10G Ethernet зазвичай використовується для підключення серверів, сховищ даних та комутаторів у великих дата центрах. Переваги 10G Ethernet включають його сумісність з існуючими структурами, низьку вартість обладнання та підтримку стандартних мережевих протоколів [4].

InfiniBand, натомість, спеціально розроблений для високопродуктивних обчислювальних середовищ. Ця технологія забезпечує вражаючу пропускну здатність до 200 гігабіт на секунду та низьку латентність менше 1 мікросекунди. InfiniBand широко використовується для обчислювано-інтенсивних завдань, наукових досліджень та моделювання. Він підтримує технологію RDMA, яка дозволяє безпосередній доступ до пам'яті іншого вузла без зайвого навантаження на процесор.

10G Ethernet, використовуючи оптичні волоконні кабелі, може досягати пропускну здатності до 10 гігабіт на секунду з латентністю близько 2-5 мікросекунд. Вартість обладнання для 10G Ethernet може розпочинатися від \$500 за порт на комутаторі, залежно від виробника та можливостей пристрою. Що стосується вартості кабелів, кожен метр оптичного волоконного кабелю вартує від \$1 до \$10.

Обладнання InfiniBand може бути вартим значно більше порівняно з 10G Ethernet - ціна за порт на комутаторі може становити від \$1000 до \$2000 і більше, залежно від моделі та можливостей. Оптичні волоконні кабелі для InfiniBand також дорожчі, приблизно \$5 до \$15 за метр.

Основні відмінності між 10G Ethernet та InfiniBand полягають у їхніх характеристиках та застосуваннях. 10G Ethernet зазвичай вартує менше, зокрема вартість обладнання та кабелів, що робить його більш доступним для менших бюджетів. В той час як InfiniBand володіє вражаючою швидкістю та низькою латентністю, ідеально підходячи для високопродуктивних обчислювань.

Вибір між 10G Ethernet та InfiniBand залежить від конкретних потреб дата центру. 10G Ethernet підходить для завдань управління вузлами та обміну даними, зокрема завдяки його сумісності та низькій вартості. InfiniBand найкращий для завдань високопродуктивних обчислень, де важлива швидкість та низька латентність. Оптимальне рішення буде залежати від обсягу завдань, бюджету та технічних вимог дата центру.

Як висновок, використаємо обидві технології: 10G Ethernet – для керування вузлами та для копіювання даних, а Infiniband для обміну даними в процесі розрахунків вузлами дата центру.

1.5 Огляд засобів захисту даних в датацентрах

Засоби захисту даних в датацентрах є важливою складовою інфраструктури для забезпечення безпеки, конфіденційності та доступності інформації. У сучасному цифровому середовищі, де кількість кібератак зростає, використання ефективних засобів захисту є невід'ємною частиною успішної діяльності будь-якого дата центру [5].

Одним з ключових засобів захисту є брандмауер. Він фільтрує мережевий трафік і дозволяє забороняти або дозволяти певні типи комунікації з інтернетом.

Брандмауери можуть бути апаратними або програмними, і вони допомагають зменшити ризик доступу несанкціонованих користувачів до мережі датацентра.

Ще одним ефективним засобом є системи виявлення та запобігання вторгнень (IDS/IPS). Системи IDS виявляють аномальну активність або незвичні патерни в мережі, тоді як системи IPS можуть автоматично реагувати та блокувати потенційно шкідливий трафік.

Для забезпечення конфіденційності даних використовують шифрування. Все, починаючи від зберігання даних на дисках до передачі через мережу, може бути зашифровано для унеможливлення несанкціонованого доступу до інформації.

Засоби аутентифікації та авторизації також грають важливу роль. Системи двофакторної аутентифікації, біометричні методи (відбитки пальців, розпізнавання обличчя) та управління правами доступу допомагають обмежити доступ до даних лише авторизованим користувачам [5].

З метою забезпечення високої доступності даних, використовують резервне дублювання. Засоби для створення резервних копій даних дозволяють відновити інформацію в разі аварій або нежданої втрати даних.

Комплексне використання цих засобів є необхідною умовою для створення надійної системи захисту даних в датацентрах. Жоден засіб не є універсальним, і вони взаємодіють, доповнюючи один одного. Брандмауери разом із системами IDS/IPS можуть виявляти та блокувати загрози ще на ранньому етапі. Шифрування забезпечує конфіденційність даних, що передаються через мережу або зберігаються на серверах. Засоби аутентифікації та авторизації гарантують, що лише вповноважені особи матимуть доступ до даних. Резервне дублювання забезпечує можливість відновлення даних в разі критичних ситуацій.

Отже, комплексне використання засобів захисту даних в датацентрах є критично важливим для забезпечення високого рівня безпеки та надійності інформації. Кожен засіб вносить свій внесок у створення комплексної системи захисту, яка здатна відповідати сучасним викликам кібербезпеки.

1.6 Висновки за розділом

В ході проведеного аналізу сучасних багатопроцесорних архітектур для обчислювальних вузлів дата центру обрано архітектуру MPP. Її переваги у масштабованості та можливості незалежної обробки даних кожним вузлом вписуються у концепцію центру обробки даних. При цьому кожний обчислювальний вузол, згідно з класифікацією Флінна, представляє собою MIMD-систему. Тобто кожен вузол незалежно від інших може повноцінно виконувати будь яку задачу і забезпечувати паралелізм всередині вузла.

Для обміну даними в дата центрі використаємо оби дві сучасні технології: 10G Ethernet – для керування вузлами та для копіювання даних, а Infiniband для обміну даними в процесі розрахунків. Середовище передачі даних – оптичне волокно.

Захист даних в дата центрі має бути забезпечений комплексно на всіх рівнях мережевої моделі OSI: від фізичного рівня перешкоджання доступу неавторизованим користувачам до програмного рівня з шифруванням, аутентифікацією, авторизацією, захистом від людини посередині.

2 ПРОЄКТУВАННЯ ДАТА ЦЕНТРУ

Процес проєктування дата центру є комплексним завданням, яке включає розробку різних складових інфраструктури. Ключовим елементом є розробка апаратної частини, яка охоплює кілька основних напрямів.

Першим напрямком є створення інформаційної інфраструктури. Це передбачає визначення та встановлення обладнання для обробки, зберігання та передачі даних, включаючи сервери, системи зберігання даних та мережеве

обладнання. При цьому важливо забезпечити високу пропускну здатність, надійність та масштабованість системи для відповідності майбутнім потребам.

Другим елементом є розробка телекомунікаційної інфраструктури. Це включає планування та встановлення мережевого обладнання для забезпечення зв'язку між внутрішніми вузлами дата центру та зовнішнім світом. Особливу увагу слід приділити пропускній здатності мережі, механізмам захисту даних та високій доступності мережевих сервісів.

Третім важливим аспектом є інженерна інфраструктура. Цей напрямок охоплює системи електропостачання, охолодження, контролю навколишнього середовища та фізичної безпеки. Ефективні системи електропостачання та охолодження є критично важливими для забезпечення безперебійної роботи обладнання.

Також слід розробити стратегії для резервного електропостачання та запобігання перегріву обладнання. Системи контролю навколишнього середовища повинні забезпечувати оптимальні умови для роботи апаратури, а системи безпеки – захищати обладнання та дані від несанкціонованого доступу та інших загроз.

2.1 Розробка інформаційної інфраструктури

Інформаційна інфраструктура Дата-центру охоплює обладнання, яке виконує основні функції, не пов'язані з комунікаційною складовою. Це включає низькопрофільні корпуси, стійки 19 дюймів, материнські плати, процесори, пам'ять, дискові накопичувачі, монітори, багатофункціональні пристрої та системи KVM.

Згідно з технічним завданням, чверть вузлів Дата-центру планується відвести під обчислювальний кластер. У цих вузлах передбачається використання по два 16-ядерних 32-поточних процесорів INTEL Xeon Gold 6226R, що належать

до сучасного і відносно бюджетного покоління процесорів від Intel. Проведено аналіз технічних характеристик цих процесорів [6].

Вони виготовлені за 14-нанометровою технологією, мають базову частоту 2.90 GHz і максимальну частоту Turbo до 3.90 GHz, кеш-пам'ять 22 MB, два канали UPI з максимальною споживаною потужністю 150 Вт. Підтримують пам'ять DDR4-2933 з максимальною швидкістю 2933 MHz, шість каналів пам'яті, підтримку Intel® Optane™ Persistent Memory, ECC Memory, масштабованість до двох сокетів, PCI Express Revision 3.0 з максимальною кількістю 48 шин PCI Express. Також процесори оснащені технологіями Intel® Deep Learning Boost, Intel® Resource Director Technology, Intel® Speed Shift Technology, Intel® Hyper-Threading Technology, Intel® Transactional Synchronization Extensions та іншими сучасними технологіями.

Ці характеристики процесора повністю відповідають технічному завданню, поєднуючи сучасні досягнення Intel, і планується їх використання на обчислювальних вузлах кластера..

Вибір серверної оперативної пам'яті включає в себе розгляд спеціалізованої варіантів, придатних для серверних систем. Однією з таких варіантів є RDIMM / UDIMM RAM. RDIMM представляє собою вид оперативної пам'яті, де модулі містять додатковий регістр між мікросхемами пам'яті та системним контролером [7]. Ці регістри відповідають за зменшення навантаження на контролер та дозволяють використовувати більше пам'яті в одному каналі. Однак через наявність регістрів, вартість RDIMM вища через складніший процес виготовлення та додаткові компоненти. Використання регістрів також призводить до додаткової затримки при доступі до пам'яті. Кожен операція читання або запису спочатку буферизується в регістрі протягом одного такту, перед тим як потрапити до чіпа DRAM через шину пам'яті. Це робить регістрову пам'ять повільнішою за UDIMM, яка не використовує регістрів.

Для вузлів кластера передбачено використання оперативної пам'яті об'ємом 48 Гб типу DDR4-2933 UDIMM RAM, а для інших вузлів – по 24 Гб такої ж оперативної пам'яті. У якості серверних материнських плат вибрано

низькопрофільні моделі, які встановлюються у корпусах 1U та монтуються в стійки 19 дюймів. Обрані материнські плати мають підтримувати два процесори вибраного раніше типу та модулі оперативної пам'яті з об'ємом від 24 Гб DDR4-2933 UDIMM і більше. Важливою є також підтримка мінімум двох накопичувачів SATA зі швидкістю 6 Гбіт/с.

Відповідно до заданих критеріїв обрано материнську плату Gigabyte з сімейства S12-P04R [8] (рис. 2.1), оснащену наступними технічними характеристиками: два сокети P (LGA 3647) для підтримки процесорів 2-го покоління Intel® Xeon® Scalable, 8 слотів для DDR4 RDIMM/LRDIMM пам'яті з чотирьохканальним доступом і загальною ємністю до 1 ТБ, два порти GbE LAN (Intel i210AT) плюс один порт для дедикованого керування GbE, чотири знімні відсіки для жорстких дисків SATAIII 3.5 дюйма, два фіксованих відсіку для SSD 2.5 дюйма SATA, опційний слот PCIe x4 2280/2260 M.2 для завантажувального пристрою, один слот розширення PCIe x8 Gen3 FHHL для вимог мережі 10G SFP+, віддалений контролер управління Aspeed® AST2500 та блок живлення на 650W з функцією 1+1 резервування.

Для серверних систем необхідно використовувати високошвидкісні дискові накопичувачі типу HDD на 7200 обертів на хвилину на шині SATA 6 Гбіт/с об'ємом 10 Тбайт кожен і кешем 6 МБ. На кожному вузлі встановимо по 2 такі жорсткі диски. Таким чином, загальний обсяг дискової пам'яті Дата-центру складе 2 жорсткі диски на 10 терабайт на 1000 вузлів, що дорівнює 20000 Тбайт.

В якості накопичувачів використовуємо серверний WD Ultrastar DC HC330 10 TB SATA (WUS721010ALE6L4/0B42266). Його технічні характеристики [9]: 10TB capacity point supports both OEM & cloud deployments • Best in Class sustained transfer rate up to 273MB/s • Advanced Format 512e models • Self-Encrypting Drive options • 5-year limited warranty. Sector size (bytes)3, 5 512e: 512 4Kn: 4096. Data buffer (MB) 256 Rotational speed (RPM) 7200 Latency average (ms) 4.16. Interface transfer rate (MB/s, max) 600 1200 Sustained transfer rate (MiB/sec, typ.) (MB/sec, typ.) Up to 260 Up to 273 Seek time (read/write, ms, typ.) 8.0 / 8.6 Error Rate (non-

recoverable bits read) 1 in 1015 Load/Unload cycles (at 40°C) 600,000 MTBF4 2M hours Annual failure rate (AFR) 4 0.44% Availability (hrs/day x days/wk) 24x7.

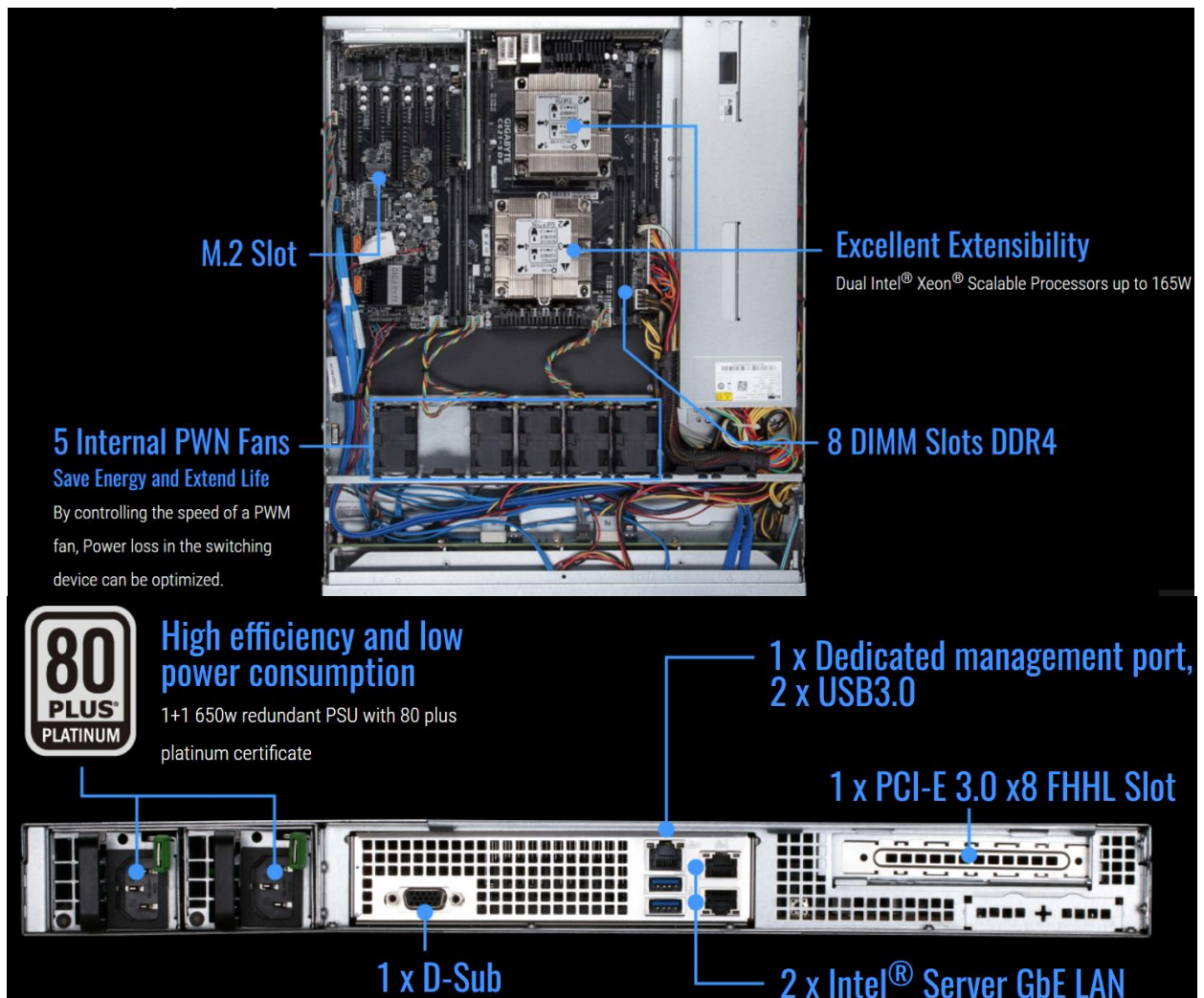


Рисунок 2.1 – Материнська плата Gigabyte S12-P04R

Відмінною особливістю даного жорсткого диска є його підвищений напрацювання на відмову, а саме 2 млн. годин проти 650 тис. годин у несерверного жорсткого диска.

У кожній серверній стійці планується встановлення одного монітора з модулем KVM, який надасть можливість користувачеві взаємодіяти з будь-яким вузлом цієї стійки. Вибір монітора здійснюється з урахуванням оптимального співвідношення якості та вартості, тому віддається перевага моніторам із матрицею типу IPS.

Сучасні монітори відрізняються наступними технічними параметрами: оснащені LED-підсвічуванням, матрицею типу IPS або VA, часом реакції матриці менше 8 мс, яскравістю дисплея понад 250 кд/м², контрастністю дисплея 1000:1 та наявністю HDMI інтерфейсів. В якості економічно ефективного варіанта, що відповідає цим характеристикам, обрано 21.5-дюймовий монітор Dell S2240L.

У кожній 19-дюймовій стійці серверного приміщення буде розміщено наступне обладнання: 20 комп'ютерних вузлів (всього 20U), 15U - модулів безперебійного живлення, 3U - KVM з монітором, 1U - комутатор Gigabit Ethernet, 1U - Infiniband, та 2 додаткові резервні одиниці. Всього для цього знадобиться 42U стійки. Для цієї мети ми плануємо використовувати стійку APC NetShelter SX 42U (рис. 2.2). Таким чином, загалом знадобиться 50 стійок, щоб розмістити всі 1000 комп'ютерних вузлів. Щодо можливого масштабування, також передбачається включення до проекту ще 10 порожніх стійок, які можуть бути заповнені обладнанням при розширенні дата-центру.



Рисунок 2.2 – Серверна стойка APC NetShelter SX 42U

Вибір комутатора KVM. KVM-перемикач призначений для з'єднання одного набору пристроїв введення-виведення між декількома комп'ютерами. У нашому випадку необхідний KVM для 20 інтерфейсів введення-виведення. Ми вирішили використовувати модель ATEN CS-1720, яка має такі технічні характеристики:

Забезпечує управління 20 USB-комп'ютерами за допомогою однієї USB-клавіатури, однієї USB-миші та одного VGA-монітора. При послідовному підключенні до 31 додаткового перемикача CS-1716 або CS-1216A здатен керувати роботою до 512 комп'ютерів. Підтримує мультимедійні USB-клавіатури.

Сумісний з усіма операційними платформами. Для полегшення та швидкості навігації по меню він володіє системою екранних меню (OSD), керованою за допомогою миші. Гарантує можливість вибору мови операційної системи та клавіатури з екранних меню OSD. Має систему PSP для контролю доступу за допомогою адміністраторського пароля. Включає в себе вбудовану спеціалізовану інтегральну схему USB ASIC від ATEN.

Таким чином спроектована інформаційна інфраструктура наведена на рис. 2.3.

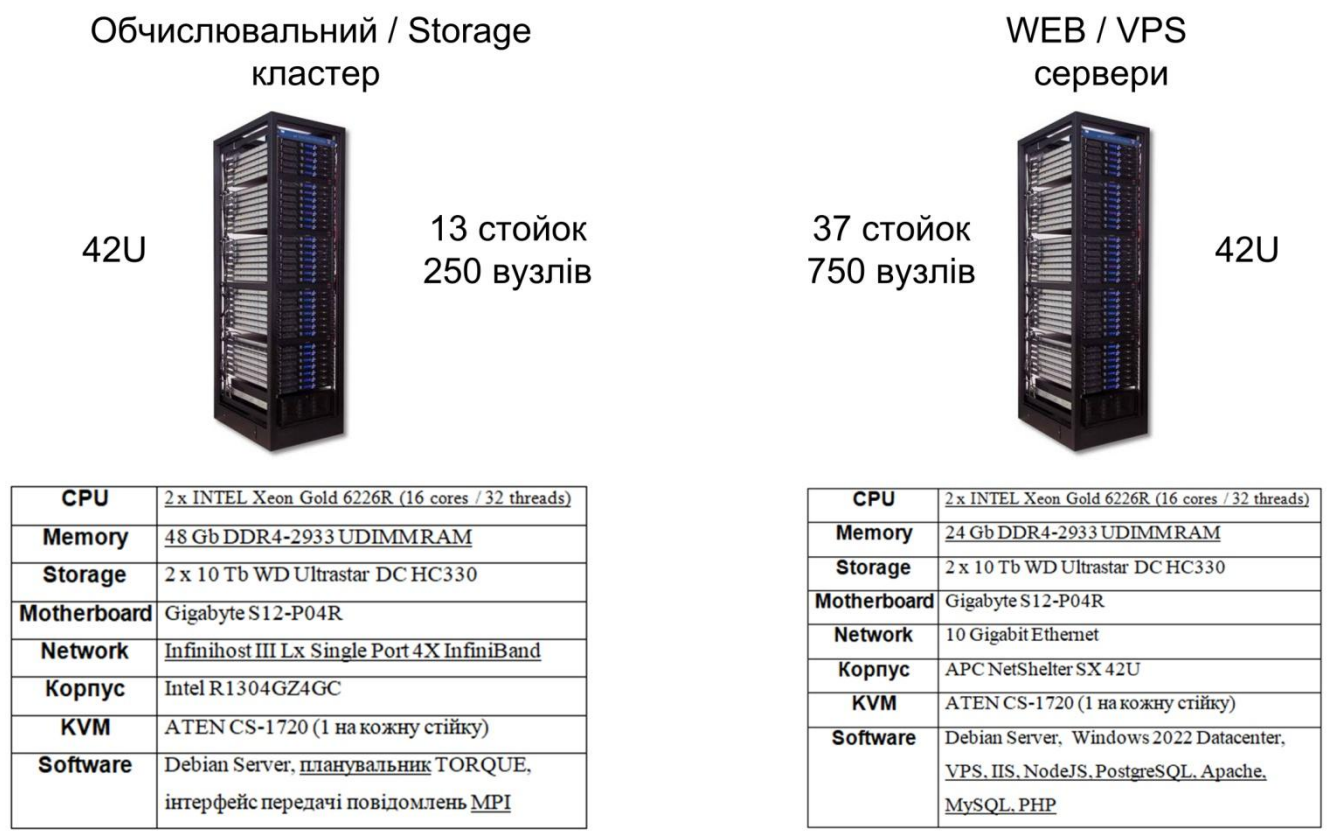


Рисунок 2.3 – Спроектowana інформаційна інфраструктура комп'ютерної системи

2.2 Розробка телекомунікаційної інфраструктури

Телекомунікаційна інфраструктура охоплює вибір засобів зв'язку, розробку монтажних та структурних схем мережі, розрахунок довжини кабелю, призначення IP-адрес, забезпечення захисту даних на всіх рівнях моделі OSI.

Для забезпечення зв'язку між вузлами необхідно використовувати 20 інтерфейсів 10 Gigabit Ethernet у кожному стійці. Кожен комп'ютер у кластері повинен бути підключений до оптичного кабелю з інтерфейсом Infiniband. Крім цього, до кожного вузла також може бути підведений зовнішній кабель для зв'язку з Інтернетом (для зовнішнього з'єднання з користувачами). Отже, на кожному комп'ютері має бути не менше 2-х інтерфейсів 10 Gigabit Ethernet (один з яких підключається до стійкового комутатора), а на вузлах кластера - по одному інтерфейсу Infiniband, який підключається до високошвидкісного стійкового комутатора. Загалом, необхідно 50 комутаторів 10 Gigabit Ethernet з 24 інтерфейсами кожний та 13 комутаторів Infiniband, також з 24 інтерфейсами кожний.

Серед інтелектуальних функцій комутаторів особливу роль відіграє підтримка алгоритму покриваючого дерева (STA), який відключає комунікаційні канали, що можуть спричинити утворення петель. Також важлива функція автоматичного визначення та зберігання MAC-адрес, що сприятиме зменшенню обсягу службового трафіку протоколу ARP під час відповідності IP-адрес до MAC-адрес.

Для реалізації 10G Ethernet комутації вибрано комутатор NETGEAR 24-Port 10G/Multi-Gigabit Plus Switch (XS724EM), який характеризується наступними властивостями [10]: керований, з двома портами 10G SFP+, можливість розміщення на столі або в стійці, обмежена пожиттєва гарантія, 24 порти Multi-gig Ethernet 1G/2.5G/5G/10G, два порти 1G/10G SFP+ зі спільною роботою, зручний інтерфейс ПЗ для базових налаштувань, захисту та моніторингу мережі, енергоефективний дизайн, відповідає стандарту IEEE802.3az.

Для Infiniband комутації застосовується Mellanox 24-портовий 4X SDR з підтримкою адаптера для зв'язку та однієї блокової живлення F-X430060. Він забезпечує одночасну передачу даних зі швидкістю 20 Гбіт/с з будь-якого інтерфейсу, підтримує 8 віртуальних Infiniband ліній, оптичний кабель та монтаж у стійку формату 1U 19 дюймів.

Для вузлів кластера необхідно встановлення низькопрофільних Infiniband карт у роз'єм PCI-Express. Вибір зроблено на користь InfiniHost III Lx, карти з одним портом 4X InfiniBand / PCI-Express x8, низькопрофільної, без пам'яті, з підтримкою оптичних волокон, відповідної стандарту RoHS (R5). Усього використовується 250 таких карт.

Виходячи з технічного завдання, маючи план будівлі (рис. 2.4) і розробивши проект конфігурації обладнання, створено схему монтажу мережі (рис. 2.5).

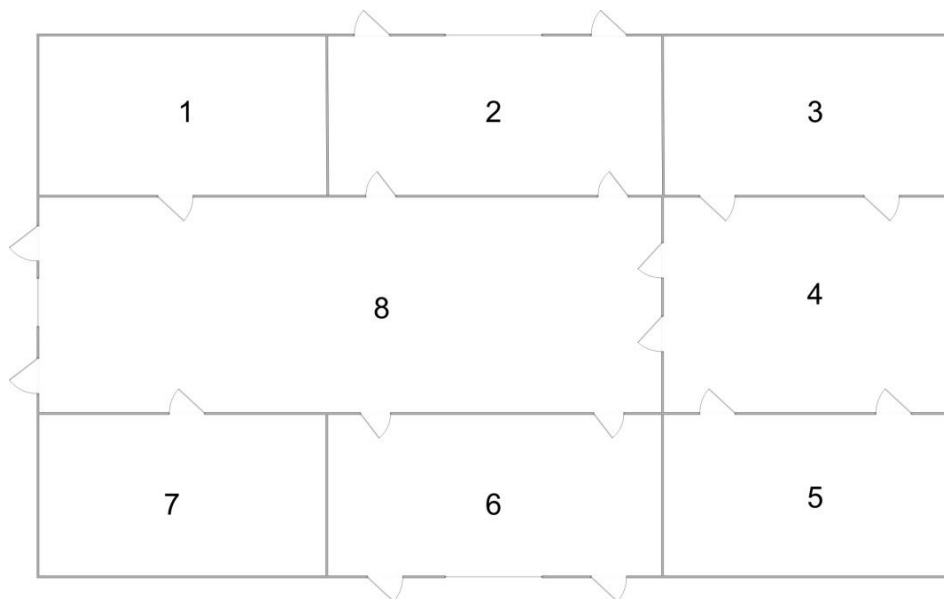


Рисунок 2.4 – План будівлі, де розміщується обладнання

Розмістимо обладнання в кімнатах (таблиця 2.1). Для прокладання кабельних трас застосовуються короби, розташовані на відстані 40 см від підлоги. Враховуючи відстані, на які буде прокладено кабель, комутатори будуть з'єднані наступним чином:

- комутатор 4 буде з'єднаний з комутатором 1;
- комутатор 1 буде з'єднаний з комутатором 2;

- комутатор 2 буде з'єднаний з комутатором 3;
- комутатор 3 в майбутньому буде з'єднаний з комутатором 5, який буде встановлений у разі розширення центру обробки даних.

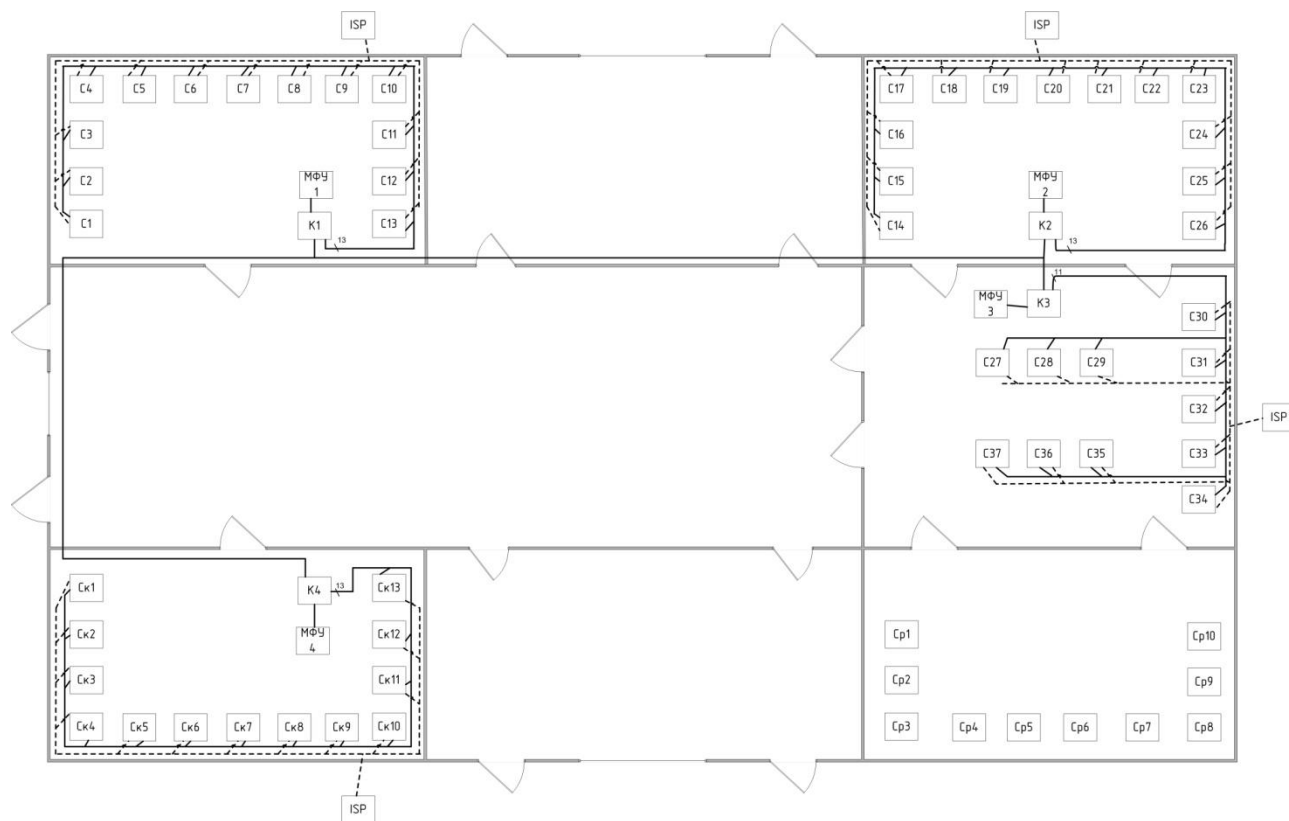


Рисунок 2.5 – Монтажна схема мережі

Таблиця 2.1 – Розміщення обладнання по кімнатах

Номер кімнати	Обладнання в кімнаті
1	13 серверних стійок, комутатор 1, МФУ1
2	вхід в Дата-центр
3	13 серверних стійок, комутатор 2, МФУ2
4	11 серверних стійок, комутатор 3, МФУ3
5	10 резервних серверних стійок
6	вхід в Дата-центр
7	13 кластерних стійок, комутатор 4, МФУ4
8	вхід в Дата-центр

Кожен вузол у кожній стійці має окремий канал від постачальника послуг Інтернет-провайдера. Цей канал підключається до окремого інтерфейсу Gigabit Ethernet, який має видиму зовні статичну IP-адресу. Це виконується з метою того,

щоб користувачі Дата-центру, які будуть орендовувати сервери, використовувати обчислювальні ресурси або розміщувати веб-сайти, могли зареєструвати зовнішні IP-адреси в DNS-базах. У коробах буде прокладено оптичне волокно для забезпечення безпеки персоналу та уникнення перешкод.

Для визначення потрібної середньої довжини кабеля (L_{cp}) використовується відома емпірична формула [11]. У центрі обробки даних обчислювальні вузли, які підключаються до мережі, рівномірно розподілені по обслуговуваній площі (з винятком маршрутів між комутаторами). Довжина маршрутів між комутаторами значно менше, ніж довжина маршрутів між вузлами та комутаторами, тому при розрахунку довжини кабеля ми візьмемо довжину маршруту між комутаторами у співвідношенні 1 до 10 відносно маршрутів між вузлами та комутаторами:

$$L_{cp} = \frac{(L_{\max}/10 + L_{\min})}{2}, \quad (2.1)$$

де $L_{\min}$ і $L_{\max}$ – довжини кабельного маршруту від точки розташування перехресного обладнання до інформаційного з'єднувача (інтерфейса комутатора) найближчого та найвіддаленішого комп'ютера відповідно, обчислені з урахуванням технології прокладання кабелю, усіх спусків, підйомів, поворотів і особливостей будівлі.

Виходячи з технології прокладання кабелю та плану будівлі, маємо $L_{\min} = 0.5$ м, $L_{\max} = 60$ м:

$$L_{cp} = \frac{(60/10 + 0.5)}{2} = 3.25 \text{ м} \quad (2.2)$$

При визначенні довжини маршрутів необхідно додати технологічний запас розміром 10% від L_{cp} та запас X для процедур розводки кабелю у розподільчому вузлі та інформаційному роз'ємі. Таким чином, довжина маршруту L буде:

$$L = (L_{cp} \cdot 1.1 + X) \cdot N, \quad (2.3)$$

де N – кількість активних мережевих пристроїв у дротових сегментах мережі, до яких відносяться комп'ютерні вузли, мережеві МФУ та комутатори.

Проведемо розрахунок приблизної довжини кабельних трас згідно з емпіричною методикою, де ми враховуємо кількість активного мережевого обладнання наступним чином: комп'ютерні вузли - 1000, мережеві МФУ - 4, комутатори - 54 (50 у стійках та 4 зовнішні). Отже, $N = 1058$.

Приблизне значення необхідної довжини оптоволоконного кабеля для 10G Ethernet буде:

$$L = (3.25 \cdot 1.1 + 0.5) \cdot 1058 = 4311,35 \text{ м} \quad (2.4)$$

Тепер розрахуємо довжину оптичного кабелю, який потрібний для з'єднання сегментів кластера (сегменти Infiniband). Кількість комп'ютерних вузлів - 250, кількість комутаторів - 13, отже, $N = 263$. Приблизне значення необхідної довжини оптичного кабелю буде:

$$L = (1 \cdot 1.1 + 0.2) \cdot 263 = 341.9 \text{ м} \quad (2.5)$$

Структурна схема відображає компоненти мережі та зв'язки між ними. Відповідно до монтажною схемою, розроблено структурну схему мережі (рис. 2.6). На схемі пунктиром показані з'єднання, які будуть виконані у разі розширення Центру обробки даних. Не використовувані 10 серверних стійок будуть з'єднані за допомогою комутатора 5, який буде підключений до комутатора 3.

Для призначення IP-адрес можна використовувати одну з двох політик: статичне призначення або динамічне. Також можливо використовувати поєднання цих двох підходів.

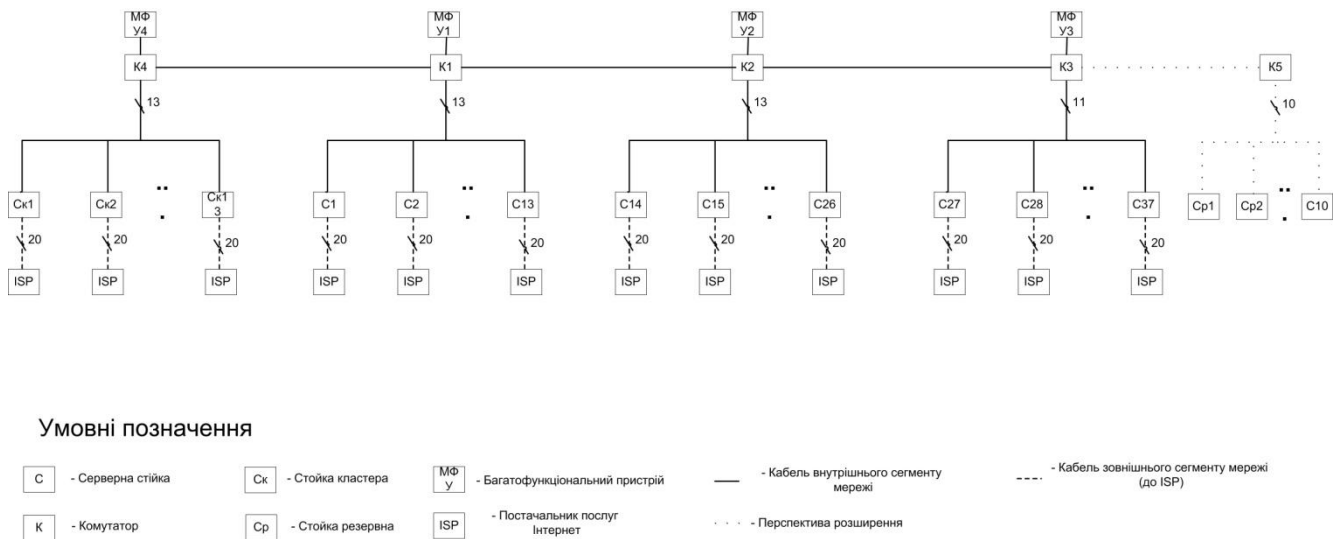


Рисунок 2.6 – Структурна схема мережі

Оскільки сервери Центру обробки даних повинні мати зовнішні адреси для постійного доступу ззовні, використано політику статичних адрес для мережевих адаптерів, які підключені до постачальника послуг Інтернет.

Для карт Infiniband також використовуємо політику статичних адрес, оскільки вузли кластера повинні мати постійні IP-адреси, що відповідають доменним іменам, зареєстрованим у таблиці хостів.

Мережевим адаптерам 10 Gigabit Ethernet, які відносяться до внутрішньої магістралі, також призначимо статичні IP-адреси.

Статичні адреси будуть також призначені мережевим багатофункціональним пристроям, для того щоб з консолей KVM було можливо доступатися до постійних адрес, належних цим пристроям.

Усі вузли Центру обробки даних повинні бути рівноправними, але можливість доступу окремих вузлів до інших вузлів повинна бути надана тільки в виняткових випадках, наприклад, при спільних розрахунках на кластерах.

З цією метою всі вузли будуть входити до робочої групи "WORKGROUP", а не до домену. Список IP-адрес та доменних імен для вузлів кластера однієї стійки наведений у таблицях 2.2–2.4; список для серверних вузлів однієї стійки - у таблицях 2.5 та 2.6.

Таблиця 2.2 – IP-адреси інтерфейсів Infiniband кластера стойки "СК1"

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
Inf1	10.11.0.1	255.255.0.0	—	локальна таблиця /etc/hosts
Inf2	10.11.0.2	255.255.0.0		
...				
Inf19	10.11.0.19	255.255.0.0		
Inf20	10.11.0.20	255.255.0.0		

Таблиця 2.3 – IP-адреси "зовнішніх" інтерфейсів кластера стойки "СК1"

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
DC_zp1	77.120.0.1	255.255.0.0	видає ISP	видає ISP
DC_zp2	77.120.0.2	255.255.0.0		
...				
DC_zp19	77.120.0.19	255.255.0.0		
DC_zp20	77.120.0.20	255.255.0.0		

Таблиця 2.4 – IP-адреса "внутрішніх" інтерфейсів кластера

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
Ge1	177.16.0.1	255.255.0.0	—	локальна таблиця /etc/hosts
Ge2	177.16.0.2	255.255.0.0		
...				
Ge19	177.16.0.19	255.255.0.0		
Ge20	177.16.0.20	255.255.0.0		

Таблиця 2.5 – IP-адреси "зовнішніх" інтерфейсів серверів стойки "С1"

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
DC_zp251	77.125.1.1	255.255.0.0	видає ISP	видає ISP
DC_zp252	77.125.1.2	255.255.0.0		
...				
DC_zp269	77.125.1.19	255.255.0.0		
DC_zp270	77.125.1.20	255.255.0.0		

Для серверів, розташованих у стійці "С1", присвоєно адреси, починаючи з інтерфейсу № 251, що відповідає першому серверу в загальному просторі імен.

Цей сервер вважається початковим у нумерації, не враховуючи ті сервери, що входять до складу кластера.

Для інших серверів, що не входять до стійки "C1", використовується аналогічний підхід у призначенні IP-адрес та доменних імен. Це забезпечує послідовність та систематичність у розподілі мережевих ресурсів. Деталі щодо IP-адрес мережевих мультифункціональних пристроїв (МФУ) представлені у таблиці 2.7.

Таблиця 2.6 – IP-адреси "внутрішніх" інтерфейсів серверів стійки "C1"

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
Ge251	177.16.1.1	255.255.0.0	—	локальна таблиця /etc/hosts
Ge252	177.16.1.2	255.255.0.0		
...				
Ge269	177.16.1.19	255.255.0.0		
Ge270	177.16.1.20	255.255.0.0		

Таблиця 2.7 – IP-адреса мережевих МФУ дата-центра

Доменне ім'я	IP-адреса	Маска підмережі	Шлюз	DNS-сервер
MFU1	177.16.250.1	255.255.0.0	—	—
MFU2	177.16.250.2	255.255.0.0		
MFU3	177.16.250.3	255.255.0.0		
MFU4	177.16.250.4	255.255.0.0		

Безпека та надійність функціонування комп'ютерної мережі є обов'язковими вимогами, які ставляться до мережі Центру обробки даних. Забезпечення безпеки функціонування мережі передбачає захист комп'ютерів та іншого активного мережевого обладнання від хакерських атак, атак типу DDOS, шкідливих програм, включаючи віруси, троянів та шпигунських програм.

Для забезпечення безпеки мережі в цілому необхідно застосувати заходи на всіх рівнях моделі OSI [12]. На прикладному рівні буде використана політика безпеки Active Directory для серверних операційних систем, що працюють під

управлінням MS Windows. Додатково, будуть застосовані антивірусні програми та шифрування трафіку. При роботі з обліковими записами будуть встановлені такі вимоги до паролів: мінімальна довжина пароля - 12 символів, обов'язкова присутність цифр, символів верхнього та нижнього регістрів, а також регулярна зміна пароля кожні 3 місяці.

За допомогою контролю окремих подій у системі буде забезпечена можливість перевірки легітимності дій користувачів взагалі, наприклад, доступу користувача до ресурсів. Для цього система визначить події, які потрібно контролювати як окремі, так і в сукупності, що забезпечить загальний контроль дій. При великій кількості подій вони будуть групуватися за функціональними або іншими спільними ознаками. З цього набору буде створена еталонна копія, що представляє собою список дозволених подій. Кожен з таких списків утворює окремий рівень контролю. Якщо спостерігається відхилення між еталонною копією та реальною подією, то вважається, що властивості системи порушені. В процесі роботи системи будуть виявлятися відхилення путем порівняння списків рівнів контролю та їх еталонних значень. В разі виявлення відхилень буде створено стан порушення безпеки системи. Як реакція на це може бути відновлення цілісності контрольованих системних файлових об'єктів чи об'єктів реєстру ОС, зупинення події (наприклад, завершення несанкціонованого процесу, призупинення несанкціонованого драйвера), відновлення події (наприклад, запуск обов'язкового процесу) або інший сценарій, передбачений системою.

Антивірусний захист робочих станцій, які функціонують під управлінням операційних систем GNU/Linux та Windows, здійснюється за допомогою відомого антивірусного програмного забезпечення, наприклад Norton. Серед переваг використання Norton можна відзначити його ефективність у виявленні та блокуванні різноманітних загроз, включаючи віруси, шкідливі програми, троянські коні, черв'яки та інші типи шкідливих кодів. Norton регулярно оновлює свою базу даних відомих загроз, що дозволяє своєчасно реагувати на нові віруси та шкідливі програми [12]. Програма забезпечує захист в реальному часі, активно перевіряючи та блокуючи шкідливі дії, перш ніж вони можуть завдати шкоди

системі. Наявність центру управління дозволяє адміністраторам контролювати захист, налаштовувати параметри сканування, переглядати журнали подій та виконувати інші управлінські дії. Крім того, багато антивірусних програм, включаючи Norton, надають технічну підтримку для вирішення будь-яких питань або проблем, пов'язаних із програмним забезпеченням. Norton сумісний з різними операційними системами та іншими програмами, що дозволяє ефективно використовувати його на різних типах серверів.

Для забезпечення безпеки мережі на транспортному рівні використовуються атрибути, які містяться у заголовках сегментів транспортних протоколів TCP і UDP. Це пов'язано з тим, що будь-яке мережеве додаток використовує транспортний протокол для передачі даних, а кожен клас додатків має свій унікальний номер транспортного порту. Ця особливість може бути використана зловмисниками для атаки на певну мережеву послугу чи службу або ж адміністратором для її захисту..

Використання політики захисту на рівні транспорту досягається за допомогою міжмережових екранів. На рівні мережі з'являється можливість встановлення захищеного з'єднання між двома комп'ютерами, що розташовані в різних точках мережі. Тут ми реалізуємо фільтрацію трафіку між внутрішньою та зовнішньою мережами, захист від несанкціонованого доступу з зовнішньої мережі до внутрішньої, маскування топології внутрішньої мережі за допомогою NAT.

Алгоритм визначення доступності середовища ґрунтується на постійному прослуховуванні середовища передачі всіма підключеними до нього вузлами [12]. Ця особливість використовується зловмисниками для організації різних видів атак на комп'ютерні мережі. Завдання протоколу канального рівня вважається виконаним, якщо обмінюються даними вузли знають апаратні адреси один одного та можуть інкапсулювати мережеві пакети в кадри канального рівня, ідентифіковані комутатором за MAC-адресами вузлів.

Вразливість системи розрішення мережових адрес стоїть у тому, що вузол (А) довіряє змісту кадра з відповіддю. Іншими словами, дані, передані у відповідь

на запит щодо розрішення мережевого адресу, ніяк не перевіряються та нічим не підтверджуються. Цю вразливість може використовувати зловмисник, який бажає підмінити вузол (Б) або відслідковувати потік кадрів, передаваних між будь-якими двома вузлами мережі.

Для захисту мережі на рівні каналу необхідно заборонити зв'язки типу "один до багатьох". Також необхідно визначити робочі місця, з яких дозволено налаштування засобів комутації каналного рівня. Адміністратор мережі повинен регулярно контролювати відповідність налаштувань комутаторів розробленій політиці захисту та вести моніторинг мережевої активності з метою виявлення джерел аномально високої кількості широкомовних запитів.

На фізичному рівні необхідно забезпечити технічні заходи щодо запобігання можливості витоку інформації за допомогою спеціальних технічних засобів через фізичні канали (провідні лінії зв'язку, міські комунікації, обладнання з виходячим побічним випромінюванням, радіопередавачі тощо) [12]. Для цього необхідно проводити прокладання мережевого кабелю в схованій проводці або в закритих кабельних каналах з можливістю опечатування за допомогою не знімних ярликів – "стікерів". Кабельні сегменти, які використовуються для підключення всіх вузлів комп'ютерної мережі, повинні бути сконцентровані на одній комутаційній панелі. Керування конфігурацією фізичних зв'язків між вузлами проводиться лише на комутаційній панелі. Комутаційну панель слід розміщувати у запереченому комутаційному шафі. Доступ до комутаційного обладнання має бути обмеженим.

2.3 Розробка інженерної інфраструктури

Інженерна інфраструктура включає в себе системи безперебійного живлення, підтримки мікроклімату, охорони та пожежогасіння.

Розрахунок системи безперебійного живлення. В кожному стійку серверної необхідно встановити систему безперебійного живлення. Відповідно до технічного завдання, система повинна забезпечувати активний захист від перепадів напруги, відключень та вимикань, працюючи при цьому протягом години від акумуляторів.

Розрахунок споживаної потужності системи у кожній стійці передбачає врахування потужності, споживаної 20 комп'ютерними вузлами. Пікова потужність, яку споживає один вузол, становить 750 Вт. У звичайному режимі роботи споживання потужності одним вузлом не перевищить 500 Вт, а для вузлів, які не входять до складу кластера, ця цифра ще менша. Додатково монітор та KVM споживають не більше 50 Вт. В результаті, у стандартному режимі роботи система в стійці споживає не більше ніж 500 Вт на 20 вузлів плюс 50 Вт, що в сумі становить 10050 Вт або 10,05 кВт. З урахуванням запасу потужності для покриття пікових навантажень необхідно систему безперебійного живлення потужністю 12 кВт.

Розглядаємо систему безперебійного живлення APC Smart-UPS RT 6000VA 230V – Marine Rack 5U, зображення фронтальної частини якої подано на рисунку 2.7, для використання у розроблюваній системі, аналізуючи її характеристики: вхідна напруга 230 вольт, частота 50/60 герц з автоматичним визначенням, тип вхідного з'єднання – жорсткий провід 3-жильний, діапазон вхідної напруги при роботі від мережі 180 - 270 вольт, максимальна вихідна потужність 4200 ватт / 6000 вольт-ампер, номінальна вихідна напруга 230 вольт, ефективність при повному навантаженні 92,0%, скаження форми вихідної напруги менше 3%, вихідна частота синхронізована з електромережею 50/60 герц з можливістю регулювання користувачем, пік-фактор 3 : 1, тип форми напруги - синусоїдальний сигнал, тип батареї - необслуговувана герметична свинцево-кислотна батарея з загущеним електролітом, захист від витоків, дві встановлені батареї, типовий час перезарядки 2,50 години, автоматичний внутрішній обхід, підтримка підключених навантажень від електромережі у випадку перевантаження або несправності, масштабована автономність за часом роботи

від акумуляторів, інтелектуальне керування батареями, замінні батареї, автоматичний перезапуск підключених навантажень після вимкнення, автоматичний запуск підключеного обладнання після відновлення електроживлення, зарядка батарей з температурною компенсацією, підвищення терміну служби акумуляторів, керування через мережу, автоматичний вбудований тест, регулярне виконання процедури самодіагностики батарей, сповіщення про передбачувані відмови, регулювання частоти та напруги, корекція відхилень частоти та напруги від норми без використання батарей, умови живлення, захист від імпульсів та викидів напруги, ударів блискавки та інших відхилень від нормальних параметрів роботи електромережі, корекція коефіцієнта потужності на вході, мінімізація витрат на інсталяцію завдяки можливості використання генераторів та кабелів, розрахованих на більш низький рівень потужності, сумісність з резервними генераторами.



Рисунок 2.7 – Система безперебійного живлення APC Smart-UPS RT 6000VA

Отже, розглянута система безперебійного живлення повністю задовольняє технічному завданню. Для масштабування потужності необхідно використовувати три такі системи в кожній стійці (разом 15U). Оскільки в Центрі обробки даних 50 стоек, то такі системи необхідно встановити у кількості $50 \times 3 = 150$ штук.

Аналіз системи контролю мікроклімату. Для забезпечення ефективної роботи центру обробки даних слід розглянути спеціалізовану систему охолодження, яка зможе виводити викидаєму потужність. Згідно з будівельними нормами, "Системи кондиціонування, а також припливні загальні системи, призначені для цілодобового та цілорічного забезпечення вимог параметрів

повітря у приміщеннях", слід передбачати щонайменше два обладнання, які забезпечать охолодження приміщень центру обробки даних. При виході з ладу одного з обладнань, слід забезпечити принаймні 50% необхідної вентиляції та задану температуру (але не менше 12 °C) в холодний період року. З урахуванням технологічних вимог або замовлення на проектування, можна передбачити установку резервних кондиціонерів, вентиляторів, насосів тощо для підтримки необхідних параметрів повітря. Практично для охолодження серверних приміщень застосовуються прецизійні кондиціонери [14].

Прецизійні кондиціонери представляють собою найбільш професійний і вартісний варіант систем охолодження, який використовується в умовах, де вирішальним є надійність та стабільність роботи, а також в ситуаціях, де вимагаються чітко задані температурні та вологісні умови експлуатації, або при створенні систем гарантованого безперебійного живлення. В технічній документації обладнання зазвичай вказуються оптимальні температурно-вологісні умови для його нормальної роботи, як правило, це температура $+22 \pm 2$ °C та вологість $50 \pm 10\%$. Важливо враховувати, що акумуляторні батареї, використовувані в блоках безперебійного живлення, є чутливими до температури. Зокрема, з підвищенням температури на 7-10°C термін їх служби може скоротитися вдвічі.

Окремі види обладнання чутливі до чистоти навколишнього повітря. Накопичений пил може уповільнити процес виведення тепла та спричинити порушення повітряного обміну всередині обладнання, зокрема через забивання вентиляторів систем охолодження та вентиляційних решіток.

Для забезпечення необхідного мікроклімату в Центрі обробки даних застосовуються прецизійні кондиціонери, які мають такі основні функції: альтернативне включення для рівномірного використання ресурсів та забезпечення максимальної надійності системи, автоматичний перезапуск після відновлення електропостачання, реагування на перевищення встановлених температурних порогів у приміщенні, відправлення повідомлень про аварійні ситуації на віддалений термінал системи моніторингу та зовнішній індикаційний

модуль, збір, зберігання та аналіз даних про стан кондиціонерів системи, температуру в приміщенні та реєстрація аварійних ситуацій, запис та збереження (не менше одного року) усіх параметрів системи та історії основних подій, можливість читання "журналу подій" за допомогою власного рідинно-кристалічного дисплея або комп'ютера. При виборі прецизійної системи кондиціонування необхідно здійснити розрахунок споживаної потужності в кожному приміщенні, а потім визначити відповідні системи кондиціонування за допомогою емпіричної формули для визначення потужності системи кондиціонування:

$$P_{\text{конд}} = P_{\text{споживана}} \cdot 0.8 \text{ кВт} \quad (2.6)$$

Розрахуємо потужність системи кондиціонування для кожної кімнати.

Результати розрахунків представлені у таблиці 2.8.

Таблиця 2.8 – Розрахунок потужності системи кондиціонування

Номер кімнати	Сумарна споживана потужність в кімнаті, кВт	Потужність системи кондиціонування, кВт
1	10.05 кВт * 13 = 130.65 кВт	104.52 кВт
2	—	—
3	10.05 кВт * 13 = 130.65 кВт	104.52 кВт
4	10.05 кВт * 11 = 110.55 кВт	88.44 кВт
5	у перспективі 10.05 кВт * 10 = 100.5 кВт	у перспективі 80.4 кВт
6	—	—
7	10.05 кВт * 13 = 130.65 кВт	104.52 кВт
8	—	—

Оскільки обладнання розміщено у кімнатах майже рівномірно, використовується одна модель системи підтримки мікроклімату, яка буде встановлена в кімнатах № 1, 3, 4, 7, а з перспективою можливості установки також і в кімнаті № 5. Як систему прецизійного кондиціонування обрано добре зарекомендовану систему STULZ CyberAir 3 PRO DX, dual-circuit, оснащену

охолоджувальною потужністю 118 кВт, споживаною потужністю 35 кВт, розмірами 2550 x 1980 x 890 мм, можливістю резервування в разі відмови одного блоку, цілодобовою роботою 365 днів на рік, точністю задаваної температури 1 градус Цельсія, а також аккумуляторною батареєю на 20 хвилин автономної роботи.

У розробці системи безпеки та пожежогасіння використовується система uSystem CAP з керуванням комплексом через GSM, включаючи сигнальні елементи та функції контролю температури, виявлення пожежі та затоплення. Система пожежогасіння забезпечує виявлення пожежі, передачу сигналів на пожежну станцію, керування інженерними системами будівлі, видачу попереджувальних сигналів людям та подачу засобу гасіння в приміщення, що підлягають захисту. У технологічних приміщеннях Дата-центру встановлено автоматичні газові системи пожежогасіння з використанням гасячого агента – інергена, який має важливі властивості, такі як безпека для здоров'я людини та природи, "сухий газ", що не утворює конденсату, збереження концентрації у повітрі, не змінює температуру приміщення, безпечний для електроніки, прозорий та не залишає слідів, допомагає в евакуації.

Система uSystem CAP дозволяє створити єдину систему охоронної та тривожної сигналізації, контролюючи зовнішній периметр Дата-центру. Для визначення загальної потужності, яку споживає Дата-центр, потрібно підсумувати потужність всього обладнання, включаючи сервери у десяти стійках, призначених для масштабування, з урахуванням технологічного запасу 10%. Цей розрахунок наведено у формулі 2.7 та 2.8.

$$P_{total} = (P_{servers} + P_{other}) \cdot 1.1 \text{ кВт}, \quad (2.7)$$

де P_{total} – загальна споживана потужність;

P_{other} – потужність, споживана системами охолодження, безпеки та пожежогасіння;

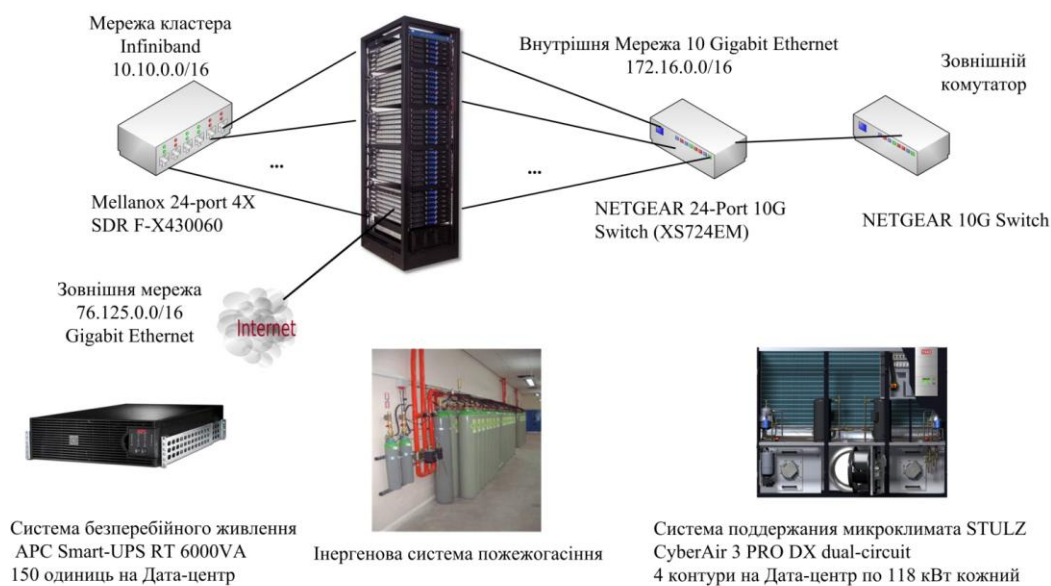
$P_{servers}$ – потужність, споживана серверами Дата-центру.

$$P_{total} = ([130.65 + 130.65 + 110.55 + 100.5 + 130.65] + [35 \cdot 5]) \cdot 1.1 = 855.8 \text{ кВт} \quad (2.8)$$

Постачання електроенергії до Дата-центру забезпечимо з власної трансформаторної підстанції з двома незалежними трансформаторами. Дата-центр підключатиметься до міських електричних мереж розподілу за другою категорією живлення.

У всіх серверних приміщеннях буде встановлено підлогу з піднятими плитами (фальшпідлога). Замість того, щоб прокладати кабелі під стелею, монтажники розміщують їх під фальшпідлогою. Існуючі рішення в цій області забезпечують ефективне охолодження, зменшення кількості та довжини схованих кабелів, консолідацію фізичних портів і зниження кількості проводів для підключення обладнання. Фальшпідлога представляє собою підлогове покриття, що складається з знімних плиток, які встановлюються на опорну конструкцію. Завдяки цьому, між покриттям і підлогою-основою утворюється простір, куди розкладаються кабелі різного призначення.

Розроблені телекомунікаційна та інженерна інфраструктури наведені на рис. 2.8.



Споживана Дата-центром потужність = 855.8 кВт/г Від власної підстанції

Рисунок 2.8 – Телекомунікаційна та інженерна інфраструктури дата центра

2.4 Налаштування операційних систем

В залежності від потреб і задач необхідно використовувати і Windows і Linux системи. Windows для платформи .NET і всієї відповідної екосистеми. Linux краще стикується з NodeJS та взагалі з javascript за стосунками.

У контексті датацентрів Windows Server 2022 пропонує кілька значущих переваг, які роблять його привабливим вибором для організацій, які шукають надійні та ефективні рішення для серверів. Ці переваги охоплюють різні аспекти, включаючи покращені можливості забезпечення безпеки, покращену продуктивність і масштабованість, розширені можливості управління та підтримку віртуалізації.

Windows Server 2022 включає в себе низку передових засобів безпеки, які спрямовані на захист даних та систем у межах дата центрів. Ці засоби включають Secured-core Server, що поєднує апаратні та програмні компоненти безпеки для захисту від атак на рівні прошивки, а також Windows Defender Antivirus, Windows Defender Exploit Guard та Windows Defender Credential Guard, що забезпечують комплексний захист від загроз.

Windows Server 2022 також приніс покращення у продуктивності та масштабованості, відповідаючи вимогам сучасних датацентрів. Служба міграції даних дозволяє безперешкодно мігрувати дані зі старих серверів на нові системи, оптимізуючи ресурси зберігання. Функція SMB Direct покращує продуктивність файлового сервера, використовуючи технологію віддаленого прямого доступу до пам'яті (RDMA) для передачі даних з високою швидкістю.

Windows Server 2022 пропонує розширені можливості управління через Windows Admin Center, сучасний засіб управління, який надає єдиний інтерфейс для управління серверами, кластерами, гіперзгуртованою інфраструктурою (HCI) та службами Azure. Цей оптимізований інтерфейс управління дозволяє адміністраторам датацентрів ефективно контролювати та керувати їх інфраструктурою.

Для підтримки передової віртуалізації Windows Server 2022 покращує свої можливості віртуалізації за допомогою інтеграції Windows Admin Center з Azure Stack HCI, що дозволяє безперешкодно розгортати гібридні хмарні рішення. Функція Azure Automanage допомагає автоматизувати та оптимізувати завдання з управління віртуальними машинами.

Windows Server 2022 також продовжує підтримувати технології контейнеризації, дозволяючи датацентрам використовувати Docker та Kubernetes для ефективного розгортання та управління додатками. Це дозволяє створювати архітектури на основі мікросервісів, що покращують масштабованість та обслуговування додатків.

Windows Server 2022 вводить інноваційні рішення для зберігання, такі як Storage Spaces Direct, яке дозволяє створювати високодоступні та масштабовані пули зберігання за допомогою невеликих апаратних засобів. Ця функція особливо корисна для датацентрів, які потребують гнучких та ефективних рішень для зберігання.

З Windows Server 2022 датацентри можуть легко інтегруватися з хмарною платформою Azure, дозволяючи реалізувати сценарії гібридного хмарного рішення. Ця інтеграція забезпечує гнучкість та агільність, дозволяючи організаціям використовувати як внутрішні, так і хмарні ресурси.

Windows Server 2022 пропонує розширений цикл підтримки, що забезпечує довгострокову стабільність та надійність. Це гарантує, що датацентри зможуть продовжувати отримувати оновлення, патчі безпеки та технічну допомогу протягом тривалого періоду, зменшуючи потребу у частих оновленнях.

Таким чином, Windows Server 2022 пропонує комплексні переваги для датацентрів, охоплюючи вдосконалені можливості забезпечення безпеки, покращену продуктивність, масштабованість, розширені можливості управління, підтримку віртуалізації, контейнеризації, інноваційні рішення для зберігання та інтеграцію гібридного хмарного середовища, включаючи розширений цикл підтримки [7].

Інструкція з автоматизованого встановлення Windows Server 2022 в датацентрі за допомогою Kubernetes передбачає кілька кроків.

На першому етапі необхідно забезпечити підготовку, яка включає встановлення кластера Kubernetes, такого як AKS (Azure Kubernetes Service) або Minikube для тестування в локальному середовищі, та використання інструментів автоматизації, наприклад Kubernetes Helm чи Ansible.

На другому етапі застосовується Kubernetes Helm для автоматизованого встановлення, що включає встановлення Kubernetes Helm на кластері, створення Helm Chart для встановлення Windows Server 2022 із заданими параметрами, а також використання команди Helm для встановлення цього Chart у кластері.

Третій крок включає використання Ansible для автоматизованого встановлення, зокрема встановлення Ansible на локальному комп'ютері або сервері у кластері Kubernetes, створення Ansible playbook для встановлення Windows Server 2022 та запуск Ansible playbook за допомогою команди.

Четвертий крок полягає в моніторингу та управлінні, який включає використання Kubernetes API для моніторингу та управління серверами та встановлення додаткових інструментів для моніторингу, таких як Prometheus і Grafana.

Останній крок полягає в оптимізації та налаштуванні Windows Server 2022 згідно з вимогами користувачів, включаючи встановлення ролей та функцій, налаштування групових політик та активацію параметрів безпеки, а також перевірку сумісності додатків з Windows Server 2022.

Для Linux операційної системи використовується Debian, який відзначається своєю стабільністю та надійністю, довгостроковою підтримкою, безкоштовністю та відкритим кодом, великою спільнотою та документацією, широким вибором пакунків програмного забезпечення, високою увагою до безпеки, гнучкістю налаштування, підтримкою різних архітектур, ефективним пакетним управлінням та високою продуктивністю. Debian є потужним і надійним варіантом для датацентрів завдяки своїм перевагам у стабільності, безпеці та гнучкості налаштування.

Алгоритм встановлення Debian в датацентрі за допомогою Kubernetes включає кілька етапів.

Спочатку потрібно підготувати інфраструктуру Kubernetes, встановивши кластер Kubernetes на серверах, використовуючи рішення, такі як Kubernetes kubectl, kops, або Kubernetes на хмарних платформах, наприклад, AKS або GKE.

Далі потрібно підготувати файли конфігурації, створивши файл конфігурації Kubernetes для ресурсів, таких як Pod, Deployment та Service, та визначивши образ Docker для контейнера Debian з Docker Hub.

Після цього необхідно створити YAML-файли (маніфести) для ресурсів, які планується створити, визначивши образ контейнера, мережеві налаштування, ресурси та інші параметри.

Четвертий крок включає встановлення Debian за допомогою Kubernetes, використовуючи kubectl для створення ресурсів з маніфестів, застосовуючи маніфести до кластеру за допомогою команди `kubectl apply -f`. Приклад такого маніфесту приведено у лістингу 2.1.

Лістинг 2.1 – Приклад маніфесту для ресурсів

```
yaml
apiVersion: v1
kind: Pod
metadata:
  name: debian-pod
spec:
  containers:
  - name: debian-container
    image: debian:latest
```

Для моніторингу та управління використовуються інструменти Kubernetes, такі як kubectl для перегляду стану ресурсів.

Нарешті, після встановлення Debian у контейнері на Kubernetes, потрібно налаштувати систему за потреби, включаючи мережеві налаштування та забезпечення доступу до Debian з інших серверів у кластері.

Налаштування мережевих адаптерів включає вказівки на статичну IP-адресу, маску підмережі, шлюз за замовчуванням та автоматичне включення

інтерфейсу eth0 під час завантаження системи. Приклад такого налаштування приведено у лістингу 2.2.

Лістинг 2.2 - Налаштування мережевих інтерфейсів сервера стойки "Ск1"

```
# Зовнішня мережа
iface eth0 inet static
address 77.125.0.1
netmask 255.255.0.0
gateway # значення, отримане від постачальника послуг Інтернет
auto eth0
# Внутрішня мережа
iface eth1 inet static
address 177.16.0.1
netmask 255.255.0.0
auto eth1
# Мережа Infiniband
iface ib0 inet static
address 10.11.0.1
netmask 255.255.0.0
auto ib0
```

Для вузлів, які не входять до кластера, налаштування мережевих адаптерів виконується подібним чином, але без налаштувань Infiniband.

2.5 Налаштування програмного забезпечення

З метою подолання недоліків, характерних для класичної архітектури клієнт-сервер, розробники програм реалізують спеціальний середній шар програмного забезпечення, який виступає як посередник (middleware) [15].

Цей шар міжпрограмного зв'язку може оптимізувати обробку даних, використовуючи високорівневі механізми, а не прямий обмін між клієнтом і сервером.

В умовах великих обчислювальних кластерів використовуються особливі компоненти проміжного програмного забезпечення, такі як планувальники та інтерфейси передачі повідомлень, для ефективно організації та координації

роботи різних вузлів кластера. Один із найбільш поширених та популярних планувальників (менеджерів ресурсів) - це Torque (Terascale Open-Source Resource and QUEUE Manager).

Після встановлення програми на головний вузол ("Infl"), де буде використовуватись демон `pbs_server` для контролю всіх вузлів, доцільно розпакувати та скомпілювати дистрибутив на цьому вузлі. Це досягається через послідовне виконання кроків, які приведено у лістингу 2.3.

Після встановлення необхідно змінити значення змінної середовища `PATH`, додавши шлях до встановленого менеджера ресурсів, оскільки за замовчуванням виконувані файли встановлюються в `/usr/local/bin`.

Лістинг 2.3 – Приклад команд для інсталювання дистрибутиву

```
tar -xzvf torqueX.tar.gz;
cd torqueX;
./configure;
make;
make install.
```

У контексті встановлення Torque на обчислювальні вузли можна використовувати вбудовану систему `trackage`, що генерує саморозпаковувальні архіви, які потім можна розповсюдити та запустити на обчислювальних вузлах. Для створення цих архівів необхідно після встановлення `pbs_mom` виконати команду `make packages`. Після цього отримані архіви можна копіювати на інші машини та запускати їх із флагом `-install`.

Налаштування демона `pbs_server` виконується через використання команди `qmgr`. Для нової системи перед налаштуванням слід виконати команду `pbs_server -t create`. Потім потрібно налаштувати бажану структуру черг та підключити планувальник.

Спершу активується планувальник (`scheduling=true`). Далі створюється черга та вказується її тип. В разі PBS чергу слід оголосити як `execution`, щоб у ній можна було виконувати завдання. Додаткові налаштування (наприклад,

встановлення параметрів черги `started` та `enabled`) дозволяють черзі приймати заявки на виконання завдань та запускати задачі, що були поставлені у чергу.

Після цього встановлюються значення за замовчуванням для параметрів `node` та `walltime` для задач, які потрапляють до черги. Ці значення використовуються при постановці завдання у чергу, якщо користувач не вказав інші значення. Команда `default_queue=short` вказує на те, що завдання буде розміщено в чергу `batch`, якщо для нього не вказана інша черга.

Після налаштування сервера потрібно визначити обчислювальні вузли. Для того, щоб `pbs_server` міг взаємодіяти з кожним з `tom`-демонів, необхідно вказати йому, які вузли враховувати. Кожен обчислювальний вузол, що повинен стати частиною обчислювальної системи, вказується у файлі кластерних вузлів. Цей файл розташований у `$TORQUE_HOME/server_priv/nodes`.

У більшості випадків досить просто зазначити ім'я вузла. Якщо обчислювальні вузли мають багато процесорів, слід вказувати кількість процесорів за допомогою `pr=кількість процесорів`. Далі необхідно налаштувати Torque на обчислювальних вузлах. Для цього на кожному обчислювальному вузлі `tom`-сервер повинен бути налаштований на політику довіри до `pbs_server` демона. Це вимагає створення конфігураційного файлу та встановлення в ньому параметра `$pbsserver`. Після налаштування необхідно запустити демона `pbs_tom` на всіх вузлах.

Для перевірки правильності конфігурації та функціональності кластера, здійснюється виклик команди `"pbsnodes"`. Після виконання цієї команди, очікується виведення переліку обчислювальних вузлів зі статусом "вільний". У контексті нашої ситуації, цей перелік матиме наступний вигляд як у лістингу 2.4.

Лістинг 2.4 – Стан кластера після налаштування проміжного програмного забезпечення Torque

```
node=Inf2
state=free
np=12
ntype=cluster
...
```

```
node=Inf250
state=free
np=12
ntype=cluster
```

Властивість "node" вказує на доменне ім'я вузла у Infiniband мережі кластера, "np" - це кількість потоків, які можуть бути виконані одночасно на вузлі, "ntype" - це тип вузла (кластер), "state" - це стан вузла (вільний, зайнятий, відключений). Таким чином, у переліку мають бути відображені всі обчислювальні вузли від другого до 250-го.

Далі варто встановити та змоделювати інтерфейс передачі повідомлень (MPI). MPI є набором функцій, які дозволяють обмінюватися даними між процесами, а також визначають стандарти передачі повідомлень. Для того, щоб така бібліотека функціонувала в середовищі виконання, необхідний шар, який є проміжним між описом бібліотеки та середовищем виконання, і його називають реалізацією MPI для даного середовища. Реалізація це набір програмних інструментів, які забезпечують виконання всіх можливостей MPI в конкретному середовищі. Ці інструменти включають парні та колективні обміни, реалізацію комунікаторів, топології, способи спілкування з середовищем виконання, інструменти профілювання та інші операції, пов'язані з запуском обчислень, оцінкою продуктивності паралельних обчислень та іншими важливими аспектами. У даному проекті використовувалася бібліотека OMPI 1.2.7.

Перший крок встановлення MPI полягає у завантаженні вихідних кодів пакету з офіційного веб-сайту OMPI. Отримавши архів `openmpi-1.2.7.tar.bz2`, його слід розпакувати і запустити конфігураційний скрипт `configure` за допомогою команди: `./configure --prefix=/usr`.

У параметрах скрипта `configure` визначається тип архітектури машини, на яку планується встановлення пакету MPI (в даному випадку Ubuntu Linux), а також шлях до каталогу, в який буде встановлений пакет. Для всіх вузлів кластера ("Inf1"–"Inf250") слід використовувати однаковий каталог для встановлення MPI. Скрипт `configure`, запущений у результаті, аналізує операційну систему та враховує її особливості для підготовки пакету MPI до компіляції.

Стандартним для MPI є використання rsh для міжвузлового зв'язку. Однак для забезпечення більш комфортного адміністрування, замінено rsh на ssh та налаштуємо безпарольний доступ до вузлів кластера з консольної машини. Це досягається завдяки використанню параметра `-rsh=/usr/bin/ssh` під час запуску скрипта `configure`.

Наступним етапом є компіляція та встановлення MPI. Це досягається послідовним викликом двох команд, таких як команда `make` та команда `make install`.

Перша з них компілює пакет MPI, а друга - встановлює його в систему, переміщуючи бінарні файли та інші необхідні компоненти в файлову систему.

Налаштування OMPI полягає у внесенні змін до відповідного файлу, який описує вузли кластера. Ім'я та розташування цього файлу можуть бути вибрані на розсуд користувача. Файл містить простий перелік комп'ютерів, які входять до кластерів `Inf1-Inf250`.

Для забезпечення ефективної взаємодії між вузлами кластера важливо налаштувати безпарольний доступ через протокол SSH.

Цей процес починається з підключення до консолі кластера за допомогою команди `ssh -p 222 root@Inf1`, що дозволяє адміністратору управляти кластером.

Після входу в систему необхідно перейти до каталогу `ssh`, використовуючи команду `cd ~/.ssh`, щоб зайти в потрібний директорій для роботи з ключами безпеки.

Далі слід згенерувати rsa-ключі, використовуючи команду `ssh-keygen -t rsa`. Це важливий крок для створення надійного ключа, що використовується для безпечного з'єднання. При запиті про ім'я файлу необхідно натиснути клавішу `Enter`, залишаючи ім'я файлу за замовчуванням `id_rsa`.

Після цього, на прохання ввести пароль, потрібно натиснути клавішу `Enter` двічі, щоб створити безпарольний ключ.

На наступному кроці виконується копіювання публічного ключа на вузол кластера за допомогою команди `scp id_rsa.pub root@Inf2:~/.ssh`.

Після цього здійснюється підключення до вузла Inf2 через `ssh root@Inf2`, щоб виконати наступні дії безпосередньо на цьому вузлі. Тут необхідно перейти до каталогу `ssh` і скопіювати публічний ключ у файл `authorized_keys2`, використовуючи команду `cat id_rsa.pub >> authorized_keys2`.

Це дозволить забезпечити безпарольний доступ до цього вузла. Після цього відключаються від вузла Inf2 та повторюють цю послідовність дій для інших вузлів кластера, щоб налаштувати безпарольний доступ до них.

У результаті, застосування цього алгоритму дозволяє налагодити взаємодію між вузлами кластера, забезпечуючи безпечне і ефективне управління ресурсами. Це становить важливу частину налаштування інфраструктури кластера і сприяє оптимальному використанню його потенціалу. Успішне встановлення та налаштування планувальника Torque та інтерфейсу передачі повідомлень OMPI на вузлах кластера є ключовими компонентами для забезпечення високої продуктивності та надійності роботи всього кластера. Віртуальний приватний сервер (VPS) принесе безліч переваг, що схожі на виділений сервер, одночасно діючи в межах спільного апаратного середовища разом з іншими віртуальними серверами. Це досягається завдяки сучасному програмному забезпеченню для віртуалізації, яке створює та керує численними "віртуальними серверами" в межах загальних апаратних ресурсів.

Кожен окремий віртуальний сервер надається замовникам для їхніх програм, баз даних та іншого, ніби це справжній виділений сервер з попередньо визначеним обсягом пам'яті, ємністю диска та пропускнуою здатністю інтернету. Ця можливість дає замовникам можливість повністю налаштовувати середовище для їхньої конкретної програми та забезпечення безпеки.

Програмне забезпечення для віртуалізації гарантує, що віртуальний сервер залишається ізольованим від дій всіх інших віртуальних серверів, які працюють на тому ж самому фізичному обладнанні. Якщо програма одного замовника стане працювати неправильно через помилки у програмному коді, інші віртуальні сервери продовжать працювати як зазвичай. Важливо, навіть серйозні відмови в межах одного віртуального сервера не вплинуть на роботу всіх інших.

Для створення екземплярів VPS ми використовуємо OpenVZ, технологію, яка дозволяє виконувати кілька ізольованих копій операційної системи на одному фізичному сервері [16]. Ці екземпляри називаються "віртуальними приватними серверами" або "віртуальними середовищами". OpenVZ базується на ядрі Linux і відрізняється від віртуальних машин, таких як VMware або Parallels Desktop, тим, що він виключно підтримує дистрибутиви Linux як "гостьові" системи.

Варто відзначити, що віртуалізація на рівні операційної системи в OpenVZ надає кращу продуктивність, масштабованість, щільність розміщення ресурсів, динамічне управління ресурсами та легкість у адмініструванні порівняно з альтернативними рішеннями [4]. Буде здійснено встановлення та налаштування OpenVZ на серверах у Центрі обробки даних, призначених для хостингу екземплярів VPS.

Для початку роботи з OpenVZ ми встановлюємо ядро з підтримкою OpenVZ: `apt-get install linux-image-openvz-686`. Після оновлення ядра ми налаштовуємо необхідні параметри ядра в файлі `/etc/sysctl.conf`, як показано на лістингу 2.5.

Лістинг 2.5 – Приклад налаштування необхідних параметрів ядра в файлі `/etc/sysctl.conf`

```
# Параметри, специфічні для вузла апаратної платформи
# Активувати пересилання пакетів і вимкнути проху
arp
net.ipv4.conf.default.forwarding=1
net.ipv4.conf.default.proxy_arp = 0
net.ipv4.ip_forward=1
# Активувати перевірку джерела маршруту
net.ipv4.conf.all.rp_filter = 1
# Активувати ключ magic-sysrq
kernel.sysrq = 1
# TCP Explicit Congestion Notification
#net.ipv4.tcp_ecn = 0
# Вимкнути відправку всіх інтерфейсів перенаправлень
net.ipv4.conf.default.send_redirects = 1
net.ipv4.conf.all.send_redirects = 0
```

Для застосування змін після оновлення конфігурації необхідно виконати команду `sysctl -p`.

Далі потрібно створити віртуальний екземпляр, базуючись на обраному образі. Оскільки на VPS хостингу ця задача повторюється кілька разів, зручно використовувати скрипт, наведений у лістингу 2.6.

Лістинг 2.6 – Скрипт налаштування VPS

```
root_pw="root_password"
# Пароль користувача root в гостьовій системі
os_template="debian-x86_64"
# Ім'я образу ОС для гостьової системи
vm_id="1"

# Ідентифікатор гостьової системи
/usr/sbin/vzctl create ${vm_id} --ostemplate ${os_template} --
config vps.basic

# Створення гостьової системи на основі обраного образу
/usr/sbin/vzctl set ${vm_id} --onboot yes --save
# Запускати гостьову систему при запуску сервера
/usr/sbin/vzctl set ${vm_id} --hostname vps${vm_id}.domain.com
# Встановити hostname гостьової ОС
/usr/sbin/vzctl set ${vm_id} --ipadd 77.125.10.1 --save
/usr/sbin/vzctl set ${vm_id} --nameserver 8.8.8.8 --save
/usr/sbin/vzctl set ${vm_id} --userpasswd root:${root_pw}
/usr/sbin/vzctl set ${vm_id} --cpulimit 25 --save
/usr/sbin/vzctl set ${vm_id} --diskspace 9000000 --save
/usr/sbin/vzctl set ${vm_id} --quotatime 600 --save
/usr/sbin/vzctl set ${vm_id} --privvmpages 1024M --save
/usr/sbin/vzctl set ${vm_id} --vmguarpages 1024M --save
/usr/sbin/vzctl start ${vm_id}
```

Після виконання цього скрипта налаштування VPS можна здавати в оренду з виділенням квот ресурсів кожному користувачеві. Доступ до гостьової системи можна отримати за допомогою протоколу SSH.

2.6 Випробування кластера

Для виконання перевірки функціонування кластера ми вирішили використовувати стандартну задачу, пов'язану з обчисленням числа Пі методом

прямокутників, і виконати це на N-процесорних системах. Код, що реалізує цей підхід за допомогою платформи OMPI [2], можна знайти у лістингу 2.7.

Лістинг 2.7 – Код застосунку "task" обчислення числа Π на кластері

```
#include <mpi.h>
#include <stdio.h>
#include <math.h>
double f(double);
double f(double a)
{   return (4.0 / (1.0 + a*a));}
int main(int argc, char *argv[])
{   int n, myid, numprocs, i;
    double PI25DT = 3.141592653589793238462643;
    double localPI, pi, h, sum, x;
    double start_wtime = 0.0, end_wtime;
    int namelen;
    char processor_name[MPI_MAX_PROCESSOR_NAME];
    MPI_Init(&argc, &argv);
    MPI_Comm_size(MPI_COMM_WORLD, &numprocs);
    MPI_Comm_rank(MPI_COMM_WORLD, &myid);
    MPI_Get_processor_name(processor_name, &namelen);
    if (myid == 0) {
        fprintf(stdout, "Enter the number of intervals: ");
        fflush(stdout);
        if (scanf("%d", &n) != 1)
        {
            fprintf(stdout, "No number entered; quitting\n");
            n = 0;
        }
        start_wtime = MPI_Wtime();
    }
    MPI_Bcast(&n, 1, MPI_INT, 0, MPI_COMM_WORLD);
    h = 1.0 / (double) n;
    sum = 0.0;
    for (i = myid + 1; i <= n; i += numprocs)
        {x = h * ((double)i - 0.5);
         sum += f(x); }
    localPI = h * sum;
    MPI_Reduce(&localPI, &pi, 1, MPI_DOUBLE, MPI_SUM, 0,
MPI_COMM_WORLD);
    if (myid == 0)
    {   printf("pi = %.16f, Error is %.16f\n",
        pi, fabs(pi - PI25DT));
        end_wtime = MPI_Wtime();
        printf("wall clock time = %f\n", end_wtime-start_wtime);
        fflush(stdout);
    }
    MPI_Finalize();
    return 0;}
```

Щоб запустити цю програму на кластері, необхідно спочатку здійснити компіляцію, підключивши необхідні заголовкові файли та бібліотеки MPI. Далі слід розробити скрипт керування для системи планування завдань TORQUE. Приклад такого скрипту зображено на рисунку 2.8.

Лістинг 2.8 – Приклад скрипту керування для системи планування завдань TORQUE

```
#!/bin/bash
#PBS -l nodes=20:ppn=24, mem=320 gb
#PBS -S /bin/bash
#PBS -q fast
./task
exit 0
```

Перший рядок вказує на стандартну директорію для виконання завдань. У наступних рядках наведені директиви планувальної системи PBS, в яких передбачені такі параметри:

l – затребувані ресурси (20 вузлів, усі сервери відносяться до одного шасі, 24 потоки та 320 ГБ пам'яті);

S – оболонка для виконання завдання;

q – черга, в яку слід поставити завдання.

передостанній рядок передбачає виконання виконавчого файлу, а останній рядок відповідає за виходження з черги зі статусом "нуль".

Для запуску програми слід виконати команди, які приведено у лістингу 2.9.

Лістинг 2.9 – Команди для запуску програми

```
cd /opt/ompi/bin
./mpirun -machinefile node.LINUX -np 20 ./task
```

Параметр "machinefile" вказує на файл, який містить список вузлів у кластері, параметр "np" визначає кількість вузлів, на яких програма буде виконуватися в середовищі MPI.

Ми також вирішили виконати цю програму на серверах кожного шасі в кластері. Зі збільшенням кількості вузлів, відповідно до кількості інтервалів для

обчислення числа Π , очікуємо пропорційного збільшення продуктивності. Для досягнення цього ми встановили кількість інтервалів на рівні 10^{15} .

2.7 Висновки за розділом

У цьому розділі було розглянуто аспекти проектування інформаційної, телекомунікаційної та інженерної інфраструктури дата центру. Розглянуто дата-центр, який містить 1000 комп'ютерних вузлів, розташованих у 50 стійках, з яких 250 вузлів призначені для обчислювального кластера, а решта 750 вузлів відведені для хостингу, включаючи VPS- та веб-хостинг. Розглянуто також можливість масштабування, передбачивши 10 стійок у окремому приміщенні.

Проаналізовано сучасні багатопроцесорні архітектури комп'ютерів та розроблено технічне обґрунтування апаратної частини Дата-центру, яке включає стійки на 19 дюймів з 20 вузлами в кожній та архітектуру Intel з двома процесорами на кожному вузлі, кожен з яких має 16 ядер і 32 потоки. Обґрунтовано вибір технології та топології побудови комп'ютерної мережі, зокрема обчислювальні вузли з'єднані мережею InfiniBand 10 Гбіт/с, а всі вузли підключені внутрішньою мережею 10 Gigabit Ethernet та зовнішнім окремим каналом від постачальника Інтернету на 1 Гбіт/с.

Розроблено монтажну та структурну схеми мережі, надано статичні IP-адреси всім кінцевим вузлам мережі. Пропрацьовано питання захисту даних в датацентрі, розроблено інженерну інфраструктуру, що охоплює живлення, безперебійне живлення та пожежогасіння. Встановлено та налаштовано операційні системи Windows Server 2022 та Debian, встановлено проміжне програмне забезпечення TORQUE, налаштовано VPS OpenVZ та виконано випробування кластера на завданні обчислення числа Π методом прямокутників.

3 РОЗРОБКА ДІАГНОСТИЧНОЇ МОДЕЛІ ТА ДОСЛІДЖЕННЯ СИСТЕМИ

3.1 Вибір типу діагностичної моделі

У контексті вирішення завдань діагностики, порівняння різних підходів до моделювання, зокрема нейронних мереж та поліноміальних регресійних моделей, є ключовим для визначення стратегії аналізу даних. Проведемо аналіз цих двох підходів у контексті діагностичних завдань [17].

Нейронна мережа прямого розповсюдження є потужним інструментом для моделювання складних та нелінійних залежностей в даних. Вона складається зі шарів нейронів, які можуть вчити складні функції завдяки використанню активаційних функцій та глибокому вивченню характеристик.

В діагностичних задачах, де залежності можуть бути дуже складними та нелінійними, нейромережі прямого розповсюдження можуть ефективно виявляти та моделювати різноманітні взаємозв'язки між даними.

Згорткова нейронна мережа підходить для завдань, пов'язаних із зображеннями та виявленням просторових закономірностей. Згорткові та пулінгові шари допомагають виявляти ознаки на різних рівнях деталізації, що може бути корисним у діагностичних задачах, де важливі деталі та структура.

Поліноміальні регресійні моделі є спрощеними інструментами, які розширюють лінійну регресію шляхом включення в них степеневих членів вхідних змінних. Вони добре підходять для даних, де залежності можуть бути апроксимовані поліноміальними функціями.

Однак ускладненіші та нелінійніші залежності можуть бути важко апроксимовані поліномами, що може обмежити їхню ефективність у діагностичних задачах зі складними взаємозв'язками.

Порівнявши ці підходи, можна зробити висновок, що для вирішення завдань діагностики нейронна мережа прямого розповсюдження виявляється найбільш підходящою альтернативою. Її глибока структура, здатність до вивчення нелінійних та складних залежностей, а також автоматичне виявлення

репрезентацій характеристик роблять її потужним інструментом для аналізу діагностичних даних.

У цьому розділі детально розглянуто процес створення моделі прямого розповсюдження нейромережі, відомої як багатошаровий перцептрон.

Початковий етап полягає у підготовці даних, який включає збір та передобробку набору даних із ознаками та цільовими значеннями, а також розподіл даних на навчальний, валідаційний та тестовий набори.

Далі слід ініціалізація, де визначається архітектура нейромережі, включаючи кількість шарів, нейронів у кожному шарі та функції активації. Потрібно також ініціалізувати ваги та зміщення для кожного нейрона.

На етапі прямого розповсюдження, для кожного навчального прикладу вхідні ознаки подаються через мережу, обчислюється зважена сума вхідних даних та застосовується функція активації.

Потім на етапі обчислення втрат визначаються втрати між передбаченим виводом та фактичною цільовою величиною за допомогою відповідної функції втрат, такої як середньоквадратична помилка.

Зворотнє розповсюдження помилок включає обчислення градієнту втрат відносно ваг та зміщень мережі, а потім оновлення цих параметрів за допомогою алгоритмів оптимізації.

На етапі тренування модель проходить через навчальні дані протягом певної кількості епох, де на кожній епосі відбувається пряме розповсюдження, обчислення втрат та зворотнє розповсюдження.

Процес валідації та раннього зупинення включає періодичну оцінку продуктивності моделі на валідаційному наборі, що допомагає уникнути перенавчання.

Після тренування модель оцінюється на тестовому наборі, щоб оцінити її загальну здатність до узагальнення.

На етапі тюнінгу гіперпараметрів проводиться налаштування параметрів, таких як швидкість навчання, розмір пакета, кількість нейронів та кількість прихованих шарів.

Нарешті, навчену модель розгортають для здійснення передбачень на нових, невидимих даних.

Таким чином, створення моделі прямого розповсюдження нейромережі є багатоетапним процесом, що охоплює збір та обробку даних, визначення архітектури, тренування, валідацію та тестування моделі, а також тюнінг гіперпараметрів та розгортання для кінцевого використання [18].

3.2 Побудова моделі

Партнерська ІТ компанія, яка веде активну співпрацю в рамках магістерської роботи, надала цінний набір історичних даних, що стосуються їхнього власного комп'ютерного кластера. Цей кластер має структурні схожості з дата центром, проектування якого є предметом дослідження у магістерській роботі.

Надані дані включають інформацію про робочі характеристики, використання ресурсів, ефективність системи та історію проблемних ситуацій, що виникали в кластері. Ця інформація стане корисним ресурсом для аналізу та може допомогти у визначенні оптимальних параметрів та потенційних поліпшень для планованого дата центру, що дозволить забезпечити більш високу ефективність та надійність у його роботі.

Серед усіх даних ми виділили три параметри, які зручно моніторити і вони впливають на стан вузла, що діагностується. Мова йде про температуру центрального процесора (T_{cpu}) – в моделі позначимо як X_1 , температуру накопичувача ($T_{storage}$) – в моделі позначимо як X_2 , а також про час відклику вузла ($T_{response}$) в мілісекундах – в моделі позначимо як X_3 .

Вихідним параметром є стан вузла, що може приймати одне з двох значень (працює = 1 та не працює = 0). Отримані від ІТ компанії дані використаємо як

навчальну вибірку для моделі. Всього 169 результатів впливу вхідних параметрів на стан вузла. Фрагмент навчальної вибірки наведено в таблиці 3.1.

Таким чином навчальна вибірка представляла собою таблицю чисел, яка складалася з 169ти рядків і 4х стовпців. Ці стовпці містили значення трьох вхідних ознак та одну вихідну змінну для кожного випадку діагностування вузла системи:

$$V = \langle X, Y \rangle, \quad (3.1)$$

де $X = \{x_1, x_2, x_3\}$, $x_i = \{x_{i1}, x_{i2}, \dots, x_{iN}\}$, $N=169$;

$Y = \{y_1, y_2, \dots, y_N\}$.

Таблиця 3.1 – Фрагмент навчальної вибірки

T_{cpu} , градусів	$T_{storage}$, градусів	$T_{response}$, мс	Стан вузла
64	59	2326	1
68	49	1496	1
70	45	3200	1
72	53	3901	1
76	40	2216	1
64	51	2728	1
69	59	4084	1
75	39	3808	1
68	46	1683	1
79	43	2701	1
80	56	2032	1
82	42	4206	1
84	43	3760	1
87	56	0	0
73	53	488	1
79	47	2546	1
73	46	3991	1
75	48	4011	1
69	49	972	1
78	52	3629	1
90	60	6237	0
65	43	3230	1
62	50	825	1
60	56	4112	1
78	47	2104	1

З метою зниження впливу варіації у порядку значень ознак на синтезовану модель, було застосовано процес нормалізації ознак. Це означає, що значення ознак були приведені до єдиного інтервалу $x \in [0;1]$ за допомогою формули (3.2):

$$x_{ijn} = \frac{x_{ij} - x_{i\min}}{x_{i\max} - x_{i\min}}, \quad (3.2)$$

де $i=\overline{1;3}, j=\overline{1;169}$.

Фрагмент отриманої в результаті нормалізації виборки наведено в таблиці 3.2.

Таблиця 3.2 – Фрагмент нормованої вибірки

$X1$	$X2$	$X3$	Y
0,462	0,474	0,121	1
0,538	0,298	0,078	1
0,577	0,228	0,166	1
0,615	0,368	0,203	1
0,692	0,140	0,115	1
0,462	0,333	0,142	1
0,558	0,474	0,212	1
0,673	0,123	0,198	1
0,538	0,246	0,087	1
0,750	0,193	0,140	1
0,769	0,421	0,105	1
0,808	0,175	0,218	1
0,846	0,193	0,195	1
0,904	0,421	0,000	0
0,635	0,368	0,025	1
0,750	0,263	0,132	1
0,635	0,246	0,207	1
0,673	0,281	0,208	1
0,558	0,298	0,050	1
0,731	0,351	0,188	1
0,962	0,491	0,324	0
0,481	0,193	0,168	1
0,423	0,316	0,043	1
0,385	0,421	0,213	1
0,731	0,263	0,109	1

Для діагностування стану вузла системи було використано паралельну систему та побудовано трьохшаровий персептрон. Перший шар містив три нейрони, другий - також три, а третій - один нейрон. Усі нейрони функціонували з використанням сигмоїдної функції активації (3.3):

$$\psi(\varphi) = \frac{1}{1 + e^{-\varphi}}, \quad (3.3)$$

де $\varphi = \varphi(w; x)$.

Під час синтезу нейромоделі, в якості дискримінантної функції використовувалася зважена сума (3.4):

$$\varphi(w; x) = w_0 + \sum_{i=1}^{|x|} w_i x_i, \quad (3.4)$$

де w_i визначає вагу i -го вхідного параметру x_i в функції $\varphi(w; x)$.

Для створення нейромоделі та визначення значень її параметрів (вагових коефіцієнтів та зсувів кожного нейрона) на вході нейромоделі використовувалися нормалізовані значення характеристик, а на виході – стан вузла системи (працює чи не працює).

Під час навчання нейромоделі як цільову функцію використовували мінімізацію середньоквадратичної помилки.

Отже, структура синтезованої трьохшарової нейромоделі Y_{NN} може бути подана так (3.5):

$$\begin{cases} Y_{NN} = \psi_{(3,1)}(\varphi_{(3,1)}(w_{(3,1)}; \psi_{(2)})); \\ \psi_{(2)} = \{\psi_{(2,1)}; \psi_{(2,2)}; \psi_{(2,3)}\} \\ \psi_{(2,k)} = \psi_{(2,k)}(\varphi_{(2,k)}(w_{(2,k)}; \psi_{(1)})), \quad k = 1, 2, 3; \\ \psi_{(1)} = \{\psi_{(1,1)}; \psi_{(1,2)}; \psi_{(1,3)}\} \\ \psi_{(1,l)} = \psi_{(1,l)}(\varphi_{(1,l)}(w_{(1,l)}; X)), \quad l = 1, 2, 3. \end{cases} \quad (3.5)$$

Процес навчання нейронної мережі виконувався методом оберненого розповсюдження помилки.

Відповідно до стандартів, вважалось прийнятним досягнення середньоквадратичної помилки, яка не перевищує 10^{-4} . У таблиці 3.3 наведено фрагмент отриманої матриці вагових коефіцієнтів w .

Таблиця 3.3 – Фрагмент отриманої матриці вагових коефіцієнтів

Номер шару	Номер нейрону в шарі	Значення зсуву	Зв'язки (з'єднання)	
			Вузол, з якого іде з'єднання	Значення вагового коефіцієнту
1	1	-1,42	ознака x_1	4,2
			ознака x_2	7,2
			ознака x_3	-2,6
			ознака x_4	-6,85
	2	-17,13	ознака x_1	-2,17
			ознака x_2	-37,46
			ознака x_3	27,36
	3	-14,46	ознака x_1	1,18
			ознака x_2	23,11
			ознака x_3	1,46
...	...	...	...	...
3	1	-3,76	нейрон (2, 1)	4,13
			нейрон (2, 2)	21,34
			нейрон (2, 3)	-32,76

Графічна ілюстрація цієї синтезованої нейромережевої моделі подана на рис. 3.1.

Після встановлення значень вагових коефіцієнтів та зсувів у вираз (3.5), враховуючи функцію активації (3.3) та дискримінантну функцію (3.4), отримуємо математичний опис синтезованої нейромережевої моделі (3.6). Цей опис відображає взаємозв'язки між характеристиками вузла, що діагностується та його станом.

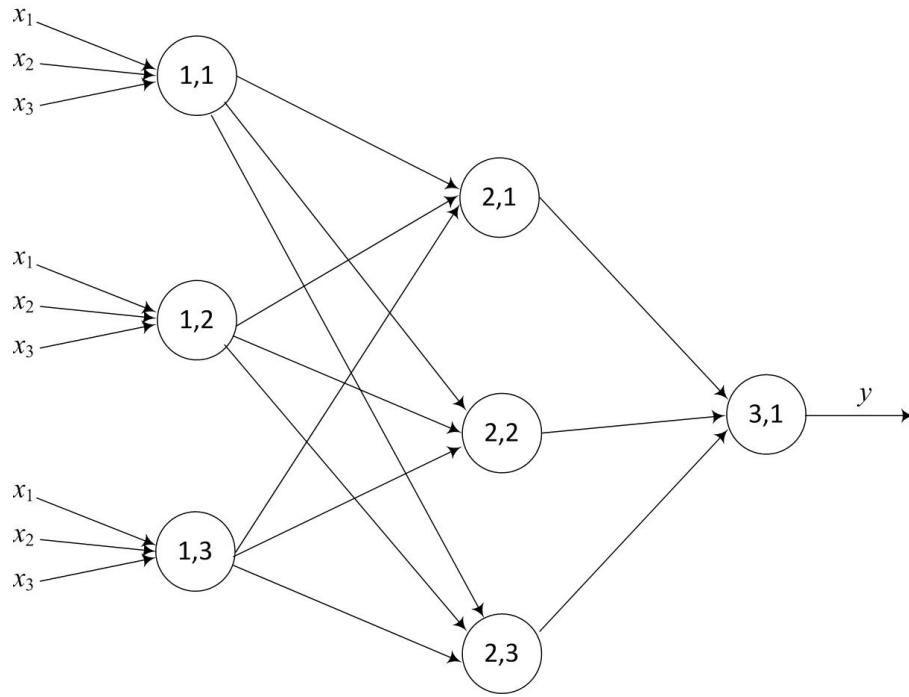


Рисунок 3.1 – Графічна ілюстрація запропонованої нейромережевої моделі

$$\begin{cases} Y_{NN} = \psi_{(3,1)} = \left(1 + e^{-(-3,76 + 4,13\psi_{(2,1)} + 21,34\psi_{(2,2)} - 32,76\psi_{(2,3)})}\right)^{-1}; \\ \psi_{(2,1)} = \left(1 + e^{-(0,21 - 4,31\psi_{(1,1)} - 3,29\psi_{(1,2)} + 7,65\psi_{(1,3)})}\right)^{-1}; \\ \psi_{(2,2)} = \left(1 + e^{-(0,37 + 14,62\psi_{(1,1)} + 10,24\psi_{(1,2)} - 7,89\psi_{(1,3)})}\right)^{-1}; \\ \psi_{(2,3)} = \left(1 + e^{-(4,23 - 19,59\psi_{(1,1)} + 22,7\psi_{(1,2)} - 4,59\psi_{(1,3)})}\right)^{-1}; \\ \psi_{(1,1)} = \left(1 + e^{-(1,42 + 4,2x_1 + 7,2x_2 - 2,6x_3)}\right)^{-1}; \\ \psi_{(1,2)} = \left(1 + e^{-(17,13 - 2,17x_1 - 37,46x_2 + 27,36x_3)}\right)^{-1}; \\ \psi_{(1,3)} = \left(1 + e^{-(14,46 + 1,18x_1 + 23,11x_2 + 1,46x_3)}\right)^{-1}. \end{cases} \quad (3.6)$$

3.3 Висновки за розділом

Таким чином, побудована нейромережева модель (3.6) має ієрархічну структуру, що включає нейрони, та дозволяє спрогнозувати стан вузла дата центру при заданих вхідних параметрах. Модель можна використовувати і як діагностичну і як прогностичну. Тобто можна не тільки діагностувати стан по

факту, но і прогнозувати стан, який буде з вузлом, якщо його параметри стануть такими як ми задамо в моделі. При цьому значення середньоквадратичної помилки моделі склало $4,39 \cdot 10^{-3}$, що вважається прийнятним для аналогічних завдань, які вирішуються з використанням синтезованої моделі.

ВИСНОВКИ

В результаті виконання роботи була розроблена комп'ютерна система дата центру з діагностуванням його стану на основі нейронної мережі. Дата-центр включає 1000 комп'ютерних вузлів, розташованих на 50 стійках: 250 вузлів об'єднані в обчислювальний кластер, 750 вузлів відведено під хостинг, включаючи VPS та веб-хостинг. Передбачено можливість масштабування до 10 стійок в окремому приміщенні.

Проведено аналіз сучасних багатопроцесорних комп'ютерних архітектур, на основі якого розроблено технічне обґрунтування апаратної частини Дата-центру: 19" стійки, в кожній з яких міститься 20 вузлів, архітектура Intel - з 2 процесорами у кожному вузлі. Обрано технологію побудови комп'ютерної мережі на підставі обґрунтування: обчислювальні вузли під'єднані до мережі Infiniband 10 Гбіт/с, всі вузли з'єднані внутрішньою мережею 10 Gigabit Ethernet, до кожного вузла підведено зовнішній відведений канал від постачальника Інтернет-послуг зі швидкістю 1 Гбіт/с. Розроблено монтажну та структурну схеми мережі. Призначено статичні IP-адреси для всіх кінцевих вузлів мережі.

Розроблено інженерну інфраструктуру: 150 систем безперебійного живлення потужністю 6000 ВА кожна, система гасіння пожежі з інергеновими реагентами, система підтримання мікроклімату з 4 контурами по 105 кВт кожний, власна підстанція на 1 МВт/год.

Обрано програмні компоненти та виконано їх установку та налаштування: Debian Linux, Windows 2022 Datacenter edition, OpenVZ VPS. Виконано встановлення та налаштування проміжного програмного забезпечення TORQUE і OMPI, перевірена функціональність кластера.

Наукова новизна роботи полягає у тому, що розроблено модель діагностування стану вузлів дата центру на основі нейронної мережі прямого розповсюдження, яка апроксимує залежність стану вузлів дата центру від температури центрального процесора, температури накопичувача і часу відклику

вузла. Модель дозволяє діагностувати стан та навіть прогнозувати стан вузла дата центру (працює чи не працює) в залежності від значень параметрів, що спостерігаються.

Практична цінність отриманих результатів полягає у тому, що спроектовано дата центр для вирішення широкого кола задач. Окрім цього запропонована діагностична модель дозволяє на практиці проводити своєчасні зупини вузлів з подальшим ремонтом, що запобігає виходу вузлів з ладу і продовжує термін служби дата центру.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1 Інститут фахівців центрів обробки даних. URL: <http://idcp.marist.edu> (date of access: 14.09.2023)

2 Голотенко О. Архітектура комп'ютерних систем: конспект лекцій для студентів усіх форм навчання з курсу «Архітектура комп'ютерних систем». Тернопіль: ТНТУ імені Івана Пулюя, 2016. 120 с.

3 Мельник А. Архітектура комп'ютера. Луцьк : Вол. обласна друк., 2008. 470 с.

4 ANSI/TIA-942 STANDARD. URL: <https://tiaonline.org/products-and-services/tia942certification/ansi-tia-942-standard/> (date of access: 18.09.2023)

5 Дрововозов В., Дидар М. Розвиток корпоративної мережі центру високопродуктивної обробки даних. Проблеми інформатизації та управління: зб. наук. праць. 2014. № 1(45). С. 42–46.

6 Intel Xeon Gold 6226R Processor. URL: <https://www.intel.com/content/www/us/en/products/sku/199347/intel-xeon-gold-6226r-processor-22m-cache-2-90-ghz/specifications.html> (date of access: 20.09.2023)

7 ExplainingComputers youtube channel. URL: <https://www.youtube.com/@ExplainingComputers> (date of access: 01.10.2023)

8 S12-P04R (rev. 1.0) Key Features | Motherboard - GIGABYTE Ukraine. *GIGABYTE Global*. URL: <https://www.gigabyte.com/ua/Motherboard/S12-P04R-rev-10> (date of access: 05.10.2023)

9 Ultrastar DC HC330 from Western Digital. URL: <https://www.westerndigital.com/products/internal-drives/data-center-drives/ultrastar-dc-hc330-hdd> (date of access: 11.10.2023)

10 Gigabit/Multi-Gigabit Plus (XS724EM). URL: <https://www.netgear.com/business/wired/switches/plus/xs724em/> (date of access: 15.10.2023)

- 11 Валецька Т. Комп'ютерні мережі. Апаратні засоби: навчальний посібник. Київ : Центр навчальної літератури, 2007. 208 с.
- 12 Network direction youtube channel. URL: www.youtube.com/channel/UCtuXekfqj-paqsxtqVNCC2A (date of access: 18.10.2023)
- 13 ДБЖ Smart-UPS Ultra. URL: <https://www.apc.com/ua/uk/campaign/smart-ups-ultra.jsp> (date of access: 01.11.2023)
- 14 Technology Connections youtube channel. URL: <https://www.youtube.com/@TechnologyConnections> (date of access: 03.11.2023)
- 15 Misra S., Bera S. Smart Grid Technology: A Cloud Computing and Data Management Approach. Cambridge University Press, 2018. 270 p.
- 16 Open source container-based virtualization for Linux. URL: <https://openvz.org/> (date of access: 12.11.2023)
- 17 Dopico J. R., de la Calle J. D., Sierra A. P. Encyclopedia of artificial intelligence. New York. 2009. Vol. 1–3.
- 18 Субботін С., Олійник А., Федорченко Є. Математичні та програмні засоби для прийняття рішень, розпізнавання образів й інтелектуального діагностування. Запоріжжя : НУ "Запорізька політехніка", 2020. 271 с.