МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

**кафедра “Інформаційні системи та мережі”**

**ПОЯСНЮВАЛЬНА ЗАПИСКА**

до дипломної роботи на тему:

Інформаційна система розподіленого зберігання файлів

**Студента групи** КН-411 Нанівського О.І.

**Керівник роботи ……….** Рішняк І.В.

**Завідувач кафедри ІСМ …………..**

**“….” ………….. 2021р.**

**ЛЬВІВ-2021**

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

Інститут комп’ютерних наук та інформаційних технологій

Кафедра “Інформаційні системи та мережі”

Напрям 6.050101 “Комп’ютерні науки”

|  |
| --- |
| “ЗАТВЕРДЖУЮ”  Завідувач кафедри ІСМ …………..  “….” ………….. 2021р. |

**ЗАВДАННЯ**

**на кваліфікаційну роботу (проект) студента групи КН-411 ОКР бакалавр**

Нанівського Олега Ігоровича

1. Тема роботи “Інформаційна система розподіленого зберігання файлів” затверджена наказом по НУ “ЛП” від “….” …………… 2021р. №…….

2. Термін здачі студентом закінченої роботи ………………..

3. Вихідні дані для роботи ………………….

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

5. Перелік графічного матеріалу ………………………………….

6. Перелік програмних продуктів, які належить використати в процесі розроблення роботи (проекту) …………………………………………………..

7. Консультування роботи, із зазначенням розділів роботи

|  |  |  |  |
| --- | --- | --- | --- |
| Розділ | Консультанти | Підпис, дата | |
| Завдання видав | Завдання отримав |
|  |  |  |  |
|  |  |  |  |

8. Дата, коли видано завдання …………………………….

Керівник …………………………… *(підпис)*

Завдання отримав до виконання ……………………….. *(підпис)*

КАЛЕНДАРНИЙ ПЛАН

|  |  |  |  |
| --- | --- | --- | --- |
| № з/п | Етапи дипломної роботи | Термін виконання етапів роботи | Примітки |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Студент …………………………… *(підпис)*

Керівник роботи …………………………… *(підпис)*

АНОТАЦІЯ

В даній роботі аналізуються проблеми зберігання файлів виликих обсягів та розроблена розподілена файлова система за допомогою мови Python. Система складається з трьох складових: клієнтської частини, сервера і серверів даних. Обмін інформацією між ними здійснюється за допомогою RPC.

Архітектура системи розроблена за зразком GFS [[1]](#Посилання1).

Ключові слова: розподілена файлова система, RPC, сервер, клієнт.

2000 символів

ANNOTATION

This paper analyzes problems of stroring large files and describes developing of disributed file system using Python programming language. The system consists of three components: client part, server and data servers. The components communicate with each other using Remote Procedure Call (RPC) protocol.

The system architecture is based on GFS princiles [[1]](#Посилання1).

Keywords: distributed file sytem, RPC, server, client.

ЗМІСТ

[ВСТУП 7](#_Toc64989890)

[РОЗДІЛ 1 (10-12%) Аналіз розподілених файлових систем 8](#_Toc64989891)

[1.1. Що таке розподілена файлова система. 8](#_Toc64989892)

[1.2. Відомі підходи побудови файлової системи. 8](#_Toc64989893)

[1.3. Актуальність теми (висновок). 8](#_Toc64989894)

[РОЗДІЛ 2 (10-15%) Системний аналіз 9](#_Toc64989895)

[2.1. Дерево цілей (варіанти побудови системи). 9](#_Toc64989896)

[2.2. Діаграми DFD і UML для деталізації структури. 9](#_Toc64989897)

[2.3. Діаграми ієрархії процесів DFD. 9](#_Toc64989898)

[РОЗДІЛ 3 (8-10%) Опис програмних засобів 10](#_Toc64989899)

[3.1. Обгрунтування вибору засобів розв’язання задачі. 10](#_Toc64989900)

[3.2. Технічні характеристики обраних програмних засобів. 10](#_Toc64989901)

[РОЗДІЛ 4 (15-20%) Практична реалізація 11](#_Toc64989902)

[4.1. Опис програми. 11](#_Toc64989903)

[4.2. Інструкція користувача (IEEE STD 1063-2001). 11](#_Toc64989904)

[4.3. Аналіз контрольного прикладу. 11](#_Toc64989905)

[РОЗДІЛ 5 Економічна 12](#_Toc64989906)

[ВИСНОВКИ 13](#_Toc64989907)

[СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 14](#_Toc64989908)

[ДОДАТКИ 15](#_Toc64989909)

ВСТУП

Найочевиднішим способом зберегти файл, це записати його на жорсткий диск, це і є найоптимальнішим варіантом в більшості ситуацій, проте, що якщо цих файлів стає занадто багато? Рішенням цієї проблеми може стати збільшення обсягу диску, або додавання ще одного, але цей процес може бути не зручним і не завжди можливим. Іншою проблемою може стати доступ до файлів, адже вони розміщені локально на диску.

Рішенням двох вищезгаданих проблем є розподілене зберігання файлів, що є зберіганням файлів на багатьох різних комп’ютерах, що дає можливість зберігати майже необмежену кількість файлів, зручну систему збільшення обсягу пам’яті в кластері. Також дана система дозволяє отримати доступ до файлів через інтернет, що дає клієнтам доступ до їхніх файлів з будь-якої точки Землі.

Google File System (GFS) – класичний приклад розподіленої файлової системи. Google як один з лідерів в сфері хмарних обчислень фолодіє значним практичним досвідом створення та підтримки розподілених файлових систем, що використовуються і сьогодні.

* В даній роботі буде описана і розроблена розподілена файлова система основана на принципах GFS [[1]](#Посилання1).

РОЗДІЛ 1 (10-12%)

Аналіз розподілених файлових систем

1.1. Що таке розподілена файлова система.

1.1.1. Короткий огляд

Розподілена файлова система – це система, яка розподілена на багатьох серверах. Це дозволяє програмам отримувати або зберігати файли як локальні з будь-якого комп’ютера або мережі.

Головна мета розподіленої файлової системи – це дозволити фізично віддаленим користувачам ділитися їхніми даними і ресурсами використовуючи звичайну файлову систему.

Розподілена файлова система має два компоненти:

* Прозорість місцезнаходження, яке здійснюється через простір імен.
* Надмірність використовуючи реплікацію файлів [[2]](#Посилання2).

Прозорість приносить зручність для кінцевого користувача даючи йому змогу маніпулювати, читати і зберігати файли на локальній машині в той час, як сам процес відбувається на серверах.

В загальному розподілені файлові системи використовуються в LAN мережах, але вони також можуть бути використаними в WAN.

Окрім розподіленої файлової системи іншим варіантом віддаленого доступу до файлів є ділитися диском. Ця система дає клієнту контроль доступу, що призводить до недоступності файлів коли клієнт офлайн. Розподілена файлова система в свою чергу є толерантною до помилок і клієнт має доступ до даних навіть якщо, декілька вузлів системи є недоступними [[3]](#Посилання3).

1.1.2. Переваги та властивості

Перевагами розподіленої файлової системи можна назвати обмеження в доступі базуючись на списку доступу або можливостей сервера і клієнта спираючись на то, як розроблено протокол. Оскільки сервер також надає одну центральну точку доступу припускається, що він буде справно працювати попри поломку окремих вузлів. Проте єдина точка доступу також і є його слабкістю. Наприклад сервер може перестати працювати через DDoS атаку. Тому зазвичай використовують запасний сервер для резервного збереження даних [[3]](#Посилання3).

Розглянемо короткий огляд основних властивостей розподіленої файлової системи:

* Прозорість
  + Прозорість структури: клієнт не знає про кількість або розміщення фалових серверів. Велика кількість файлових серверів використовується для збільшення продуктивності, адаптивності та надійності.
  + Прозорість доступу: отримання файлу на локальній і віддалених машинах не має відрізнятися, система автоматично знаходить відповідний файл і надсилає його клієнту.
  + Прозорість імене: ім’я файлу не має змінюватися, або вказувати на його місцезнаходження.
  + Прозорість реплікації: якщо файл скопійований на декількох вузлах, його копії і місцезнаходження повинні бути сховані від інших вузлів.
* Мобільність користувача: система автоматично приносить домашню папку користувача до вузла з якого користувач здійснює вхід.
* Продуктивність: продуктивність визначається середнім часом для обробки запиту клієнта. Рекомендовано, щоб цей час був подібним до централізованих систем.
* Простота: інтерфейс має бути простим з невеликою кількістю команд.
* Висока доступність: система повинна працювати в разі частоких поломок, як поломка з’єдання або поломка вузла. Високоадаптивні розподілені файлові системи повинні мати різні і незалежні файлові сервера для контролю різних і незалежних пристроїв зберігання [[2]](#Посилання2).

1.1.3. Історія

1.2. Відомі підходи побудови файлової системи.

1.3. Актуальність теми (висновок).

РОЗДІЛ 2 (10-15%)  
Системний аналіз

2.1. Дерево цілей (варіанти побудови системи).

2.2. Діаграми DFD і UML для деталізації структури.

2.3. Діаграми ієрархії процесів DFD.

РОЗДІЛ 3 (8-10%)  
Опис програмних засобів

3.1. Обгрунтування вибору засобів розв’язання задачі.

3.2. Технічні характеристики обраних програмних засобів.

РОЗДІЛ 4 (15-20%)  
Практична реалізація

4.1. Опис програми.

4.2. Інструкція користувача (IEEE STD 1063-2001).

4.3. Аналіз контрольного прикладу.

РОЗДІЛ 5  
Економічна

ВИСНОВКИ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. The Google File System [Електронний ресурс] / Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung Google – Режим доступу: <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/035fc972c796d33122033a0614bc94cff1527999.pdf>

2. What is DFS(Distributed File System)? [Електронний ресурс] / Snigdha Yambadwar - Режим доступу: <https://www.geeksforgeeks.org/what-is-dfsdistributed-file-system/>

3. Distributed File System (DFS) [Електронний ресурс] / techopedia - - Режим доступу: <https://www.techopedia.com/definition/1825/distributed-file-system-dfs>

4.

ДОДАТКИ