

Національний університет «Львівська політехніка»

Кафедра Інформаційні системи та мережі

Результат

перевірки на зовнішній плагіат бакалаврської кваліфікаційної роботи

студента групи КН-411 Нанівського О.І.

на тему: «Інформаційна система розподіленого зберігання файлів»

Проверка уникальности
Уникальность: **100.00%**
[Получить ссылку на проверку](#)
[Зафиксировать уникальность](#)
[Получить кнопку уникальности](#)
[Подробнее](#)

Проверка орфографии
В тексте найдено 39 ошибок:
* в
* i
* повтор пробела
[Подробнее](#)

SEO-анализ текста
Всего символов: **14185** Заспамленность: **63%**
Без пробелов: **12303** Воде: **4%**
Количество слов: **1902** Замена символов: **665**
[Подробнее](#)

Подсвечено: ■ Неуникальные фрагменты

ВСТУП
Розподілені системи для зберігання файлів надають можливості зручного та безпечного зберігання використовуючи віддалені машини. Починаючи від їх першої імплементації NFS, яка комерційно використовувалася, розподілені файлові системи постійно вдосконалювалися і набували ще більшої популярності. Сьогодні вони є незамінними в сфері великих даних і хмарних обчислень. Великий попит на розподілені файлові системи призвів до зростання пропозицій на ринку. Проте всього кілька компаній таких, як Google, AWS, Oracle разом займають більшу частину ринку, що не дозволяє іншим реалізаціям набути популярності, окрім локального ринку. Також враховуючи масштаб компаній, фокус їхньої уваги може легко обійти локальний ринок. Так розробка вітчизняного аналога зможе надавати кращі послуги орієнтовані на українські компанії.
Дана робота має на меті продемонструвати принципи роботи розподілених файлових систем. Вона може стати основою для майбутнього комерційного проекту, або використана для навчальних цілей.
Для її розробки необхідно виконати ряд завдань:
Провести аналіз предметної області. Дослідити відомі реалізації подібних систем.
Побудувати архітектуру системи, забезпечивши прозорість системи, клієнт не знає про структуру файлової системи, а бачить лише один простір імен, а також надмірність системи, що забезпечить її толерантність до відмов. Провести системний аналіз.
Проаналізувати технічні засоби за допомогою, яких можна побудувати систему.
Реалізувати систему та провести її тестування.
Об'єктом дослідження є зберігання файлів.
Предметом дослідження є збереження файлів на віддаленій машині з можливістю зручного доступу до них з будь-якого іншого пристрою.
Інтерфейс отримання та збереження файлів, як на звичайній файловій системі, а також запобігання втрати файлів.
Результат даної роботи може бути використаний для демонстрації роботи розподіленої файлової системи і вивчення її поведінки. Також система може бути практично використана для зберігання файлів одним, або кількома користувачами, проте усі користувачі будуть мати доступ до усього простору імен файлової системи.
В системі реалізовані базові можливості зберігання та читання файлів, а також маніпуляції з простором імен.

РОЗДІЛ 1
Аналіз джерел
1.1. Архитектурні принципи розподілених файлових систем
Розподілена файлова система — це система, яка розподілена на багатьох серверах. Це дозволяє програмам отримувати або зберігати файли як


Текст сохранен


☐ Проверка по Яндекс.Дзен

[Проверить уникальность](#)

Вы можете повысить уникальность текста на нашей Бирже рейтинга.
[Повысить уникальность](#)

Версии текста:
3 минуты назад (UTC +03:00)
Уникальность **100%** Орфография **39**
Всего символов **14185** Заспамленность **63%**
Без пробелов **12303** Воде **4%**
Количество слов **1902** Замена символ... **665**

**Доступность проверки**
Доступ к тексту закрыт. Если вы хотите открыть публичный доступ к тексту, вам необходимо [зарегистрироваться](#)

**Ссылка на проверку**
Сохраните в своем аккаунте данный текст

Уникальность: 100%

Проверка уникальности

Уникальность: 100.00%

[Получить ссылку на проверку](#)
[Зафиксировать уникальность](#)
[Получить кнопку уникальности](#)

Подробнее

Проверка орфографии

В тексте найдено 36 ошибок:

- повтор пробела
- основного серверу
- i

Подробнее

SEO-анализ текста

Всего символов: 12024

Заспамленность: 62%

Без пробелов: 10406

Вода: 3%

Количество слов: 1644

Замена символов: 568

Подробнее

Подсвечено: ☒ Смешанные слова

Рис. 1.5 Архитектура HDFS

Hadoop Distributed File System має master-slave архітектуру NameNode-DataNode, подібну до архітектури GFS. Кластер складається з одного NameNode, основного серверу, який керує простір імен файлової системи і регулює доступ до файлів, який відбуваються клієнтом. Також в кластері є кілька DataNode, зазвичай один на вузол.

Отже, HDFS показує високу масштабовану і толерантну до помилок файловою системою, яка дозволяє ефективно зберігати дані в файлах. Внутрішню, файл розподілений на багато блоків, які зберігаються на кількох DataNodes, щоб забезпечити живучість даних в разі помилок. NameNode відповідальний за виконання операції простору імен файлової системи, таких як, відкриття, перейменування, закриття файлів чи директорій. Розподілення блоків до серверів даних також виконується головним сервером. Операції клієнта читання або запису виконуються серверами даних. Сервери даних також допомагають в створенні, перейменуванні, або видаленні блоків, як інструкції головного серверу [4].

Таблиця 1.2

Порівняльний аналіз GFS та HDFS

Критерій оцінки GFS HDFS

1 2 3

Апаратне забезпечення Дешеві компоненти з Linux системою, які мають високий шанс поломки Дешеві компоненти з великою місткістю

Комунікація TCP комунікація

Періодичні командні повідомлення для синхронізації RPC поверх TCP/IP

Періодичні командні повідомлення для синхронізації

Ієрархія MasterNode і ChunkNodes NameNode і DataNodes

Безпека Google має багато серверів даних в невідомих місцях, щоб збільшити надійність Безпека основана на POSIX моделі користування і груп

Розмір блоку 64 мегабайти 128 мегабайт

Продовження Таблиці 1.2

1 2 3

Метадані для блоку Кожен чанк має 64КБ даних і 32 біти чек суми Для кожного чанку створено два файли:

1. Файл даних

2. Файл метаданих

Фактор реплікації 3 рази, але може бути встановлена будь-яка кількість 2 або 3 рази

Хто використовує Google Inc. Yahoo! Facebook IBM

Текст сохранен

Проверить SEO-данные

Вы можете повысить уникальность текста на нашей Бирже рейтинга.

Повысить уникальность

Версии текста:

2 минуты назад (UTC +03:00)

Уникальность 100%

Орфография 36

Всего символов 12024

Заспамленность 62%

Без пробелов 10406

Вода 3%

Количество слов 1644

Замена символов... 568

Доступность проверки

Доступ к тексту закрыт. Если вы хотите открыть публичный доступ к тексту, вам необходимо [зарегистрироваться](#)

Проверка уникальности

Уникальность: 92.02%

[dsas.to.ck.ua/ahp-help](#) 5%

[dspace.nau.edu.ua/bitstream/NAU/47215/1/...](#) 1%

Подробнее

Проверка орфографии

В тексте найдена 31 ошибка:

- розроблені
- -
- i

Подробнее

SEO-анализ текста

Всего символов: 14759

Заспамленность: 69%

Без пробелов: 12707

Вода: 3%

Количество слов: 2321

Замена символов: 737

Подробнее

Подсвечено: ☒ Заспамленность текста

Необхідними умовами для виконання фактору "Розроблення архітектури системи" є:

Вибір технологій для реалізації. Саме за допомогою них буде розроблятися дана система. Потрібно вибрати мову програмування, за допомогою якої можна буде розробити дану систему, а також бібліотеки, які знадобляться для автоматизації роботи.

Аналіз літературних джерел. Метою якого є дослідження теми побудови розподілених систем, дослідницьких статей, книг і курсів, що допоможе в розумінні роботи архітектури, поширених проблем і помилок, а також майбутній розвиток системи. Це допоможе приймати правильні рішення при проблемах створення системи.

Дослідження подібних систем допоможе зрозуміти популярні архітектурні рішення і їх вплив на поведінку системи. Використання досвіду при розробленні подібних систем допоможе значно скоротити час розробки. В особливості будуть досліджені такі популярні розподілені файлові системи, як GFS та HDFS.

Створення DFD діаграм допоможе візуально продемонструвати архітектуру системи, як на абстрактному рівні, так і на більш деталізованому та продемонструвати роботу процесів.

Побудова дерева цілей використовується для візуального представлення критичних факторів необхідних для успішної побудови системи.

Перший критичний фактор - це дослідження і аналіз теми. Для успішної реалізації системи необхідно володіти інформацією про актуальність даної теми та розуміти сучасні тренди. Також знання архітектур подібних систем значно спростить розробку.

Необхідними умовами для виконання фактору "Створення системи" є:

CLI клієнта - це інтерфейс командного рядка за допомогою якого клієнт може викликати функції системи, а саме записати дані до системи, прочитати його та видалити. Також клієнт може маніпулювати простором імен, тобто створювати директорії, видаляти їх, виводити існуючі та змінювати робочу директорію.

Методи основного серверу. Вони знаходяться на віддаленому сервері і викликаються RPC методами клієнта. Крім вищезазначених методів, які клієнт може викликати, він також проводить періодичний аналіз і відновлення від помилок.

Методи чанк серверів використовуються для виконання інструкцій основного сервера та запису даних. Також для розгортання системи потрібна інструкція для розгортання цілого кластеру чанк серверів.

Другий критичний фактор описує методи необхідні для функціонування системи, як зі сторони клієнта, так і основного серверу та чанк серверів.

Необхідними умовами для виконання фактору "Перевірка системи" є:

Написання юніт тестів допоможе переконатися в правильності роботи системи. Перевірка методів маніпулювання з метаданими, з файлами та перевірка відновлення від несправностей.

Написання документації. В цей пункт входить інструкція користувача та загальний опис системи. Він необхідний користувачам для розуміння роботи системи та зручного початку роботи з нею. Також він необхідний розробникам для розуміння внутрішньої структури системи.

Графічне представлення роботи відбувається за допомогою просто сайту, де зображено стан системи

Текст сохранен

Проверить SEO-данные

Вы можете повысить уникальность текста на нашей Бирже рейтинга.

Повысить уникальность

Версии текста:

2 минуты назад (UTC +03:00)

Уникальность 92%

Орфография 31

Всего символов 14759

Заспамленность 69%

Без пробелов 12707

Вода 3%

Количество слов 2321

Замена символов... 737

Доступность проверки

Доступ к тексту закрыт. Если вы хотите открыть публичный доступ к тексту, вам необходимо [зарегистрироваться](#)

Проверка уникальности

Уникальность: 99.18%

0%

Подробнее

Проверка орфографии

В тексте найдены 32 ошибки:

- i
- рівня
- Россумом

Подробнее

SEO-анализ текста

Всего символов: 13621

Заспамленность: 62%

Без пробелов: 11815

Вода: 4%

Количество слов: 1852

Замена символов: 649

Подробнее

Подсвечено: Заспамленность текста

Третім процесом для збереження файлу в розподілену файловою систему є запис блоків (Рис.2.6). В загальному це процес запису частини файлу на віддаленому сервері даних.

Розглянемо його підпроцеси:

- 1.Отримати адресу серверу даних. На вхід процес отримує локації репліки блоків даних за допомогою елементу контролю RPC, як результат попереднього процесу, а також ім'я файлу в просторі імен для ідентифікації. Механізмом є клієнт.
- 2.Передавати блок файлу. Отримавши на вхід блоку файлу та адрес серверу даних процес за допомогою елементу контролю RPC - віддалених процедурних викликів записує частину файлу в розподілений файловою системою. Як вихід ми отримуємо записаний блок в системі. Механізмами є клієнт який хоче записати блок та сервер даних на якому він записується.

Розглянувши деталізацію процесів для запису файлу в розподілену файловою систему можна спостерігати важливість і велику кількість комунікацій за допомогою віддалених процедурних викликів між клієнтом та основним сервером і в меншій мірі між клієнтом і серверами даних, проте саме комунікація між клієнтом та серверами даних займає більшу частину трафіку, оскільки основний сервер маніпулює лише метаданими, а сервери даних записують великі частини файлів. Дане рішення дозволило робити швидкі, але легкі запити до системи і окремо важкі, але стабільні передачі файлів.

Серед найосновніших процесів в системі можна виділити оптимальне розподілення реплік блоків серед серверів даних та правильна реєстрація ідентифікатора на основному сервері, помилка в якій може призвести до незворотних змін в розподілений файловою системою.

Також варто звернути на відсутність комунікації між основним сервером і серверами даних. В даному випадку це не необхідно, адже основний сервер володіє інформацією про сервери даних, як результат попередніх перевірок.

Текст сохранен

Проверить SEO-данные

Вы можете повысить уникальность текста на нашей Бирже рерайтинга.

Повысить уникальность

Версии текста:

3 минуты назад (UTC +03:00)

Уникальность	99%	Орфография	32
Всего символов	13621	Заспамленность	62%
Без пробелов	11815	Вода	4%
Количество слов	1852	Замена символов...	649

Доступность проверки

Доступ к тексту закрыт. Если вы хотите открыть публичный доступ к тексту, вам необходимо зарегистрироваться

Проверка уникальности

Уникальность: 100.00%

Подробнее

Проверка орфографии

В тексте найдено 49 ошибок:

- в значній мірі
- i
- інтерпритованою

Подробнее

SEO-анализ текста

Всего символов: 14917

Заспамленность: 63%

Без пробелов: 12950

Вода: 4%

Количество слов: 1998

Замена символов: 716

Подробнее

Подсвечено: Заспамленность текста

3.2. Технічні характеристики обраних програмних засобів

Технічні характеристики обраних програмних засобів в значній мірі впливають на процес розробки системи та її роботу, що вимагає детальнішого аналізу. Розглянемо технічні характеристики мови Python. Python є динамічною, високорівневою, безплатною і інтерпритованою мовою програмування. Вона підтримує об'єктно-орієнтоване програмування та процедурно орієнтоване [29].

Особливості Python:

- 1.Легке написання коду. Python є високорівневою мовою програмування. Python є досить легким для вивчення порівняно з мовами, такими, як C, C#, JavaScript, Java. Досить легко писати код використовуючи Python і будь-хто може вивчити основи за кілька годин, або днів.
- 2.Безплатна і з відкритим кодом. Мова Python вільнодоступна на офіційному сайті. Оскільки вона з відкритим кодом, це означає, що код доступний для публіки. Отже, є можливість завантажити її, а також поділитися.
- 3.Об'єктно-орієнтована. Однією з особливостей мови є її об'єктно-орієнтоване програмування. Python підтримує концепції класів, інкапсуляції і інші.
- 4.Підтримка розробки інтерфейсів. Графічні інтерфейси можуть бути розроблені використовуючи модулі такі, як: PyQt5, wxPython, Tk.
- 5.Високорівнева мова. Python є високорівневою мовою. При написанні програм на python нам не потрібно знати архітектуру системи, або взаємодіяти із пам'яттю.
- 6.Можливість розширення. Python дає можливість записати деякий python-код або компілювати його в C/C++.
- 7.Переносна мова. Python програми можна запускати на різних платформах, таких як, Linux, Windows, Mac, Unix.
- 8.Інтергована мова. Python надає можливість інтеграції із іншими мовами програмування.
- 9.Інтерпритована мова. Програми виконуються один рядок за раз, що дозволяє простіше відлагодити код. В внутрішній структурі Python-код перетворюється в bytecode.
- 10.Велика стандартна бібліотека. Існує велика стандартна бібліотека, яка надає модулі для великої кількості загальних задач, що дозволяє не писати код один і той самий код кожен раз.
- 11.Динамічно типізована мова. Це означає, що тип змінних визначається в процесі виконання програми.

Розглянувши особливості мови програмування Python варто також розглянути і її обмеження. Розглянемо найбільш важливі обмеження:

- 1.Продуктивність та швидкість. Численні дослідження показали, що Python в загальному повільніша, ніж інші сучасні мови програмування такі, як Java, C++. Проте існує декілька можливостей для програм працювати швидше. Однією з них є написання власного часу виконання програми і використовувати його замість стандартного.
- 2.Несумісність двох версій. Офіційно друга версія Python є застарілою, але обидві версії, друга і третя, оновлюються на регулярній основі. Також досі велика кількість програмістів надають перевагу другій версії враховуючи велику кількість популярних фреймворків та бібліотек, які підтримують тільки нову версію.

Текст сохранен

Проверить SEO-данные

Вы можете повысить уникальность текста на нашей Бирже рерайтинга.

Повысить уникальность

Версии текста:

2 минуты назад (UTC +03:00)

Уникальность	100%	Орфография	49
Всего символов	14917	Заспамленность	63%
Без пробелов	12950	Вода	4%
Количество слов	1998	Замена символов...	716

Доступность проверки

Доступ к тексту закрыт. Если вы хотите открыть публичный доступ к тексту, вам необходимо зарегистрироваться

Виконав студент групи КН-411

підпис

Нанівський О.І.

прізвище та ініціали

Керівник

підпис

Рішняк І.В.

прізвище та ініціали