

МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ЗБЕРІГАННЯ ТА ОБРОБКИ ВЕЛИКИХ ОБСЯГІВ ДАНИХ У ПРИВАТНИХ ХМАРНИХ СЕРВІСАХ

¹Вінницький національний технічний університет

Анотація

У тезах розглядаються сучасні методи підвищення ефективності зберігання та обробки великих обсягів даних у приватних хмарних сервісах. Проаналізовано підходи до оптимізації використання дискового простору, зменшення затримок доступу до даних та підвищення продуктивності обчислень. Особливу увагу приділено таким методам, як дедуплікація даних, стиснення, партиціювання, кешування та розподілена обробка даних. Наведено порівняльний аналіз основних технологій зберігання даних у приватних хмарних середовищах. Отримані результати можуть бути використані при проектуванні та оптимізації приватних хмарних інфраструктур.

Ключові слова: приватна хмара, великі дані, зберігання даних, обробка даних, оптимізація.

Abstract

The paper considers modern methods for improving the efficiency of storage and processing of large volumes of data in private cloud services. Approaches to optimizing disk space usage, reducing data access latency, and increasing computational performance are analyzed. Special attention is paid to data deduplication, compression, partitioning, caching, and distributed data processing. A comparative analysis of key data storage technologies used in private cloud environments is presented. The obtained results can be applied in the design and optimization of private cloud infrastructures.

Keywords: private cloud, big data, data storage, data processing, optimization.

Вступ

Зі зростанням обсягів цифрової інформації проблема ефективного зберігання та обробки великих даних набуває особливої актуальності. Приватні хмарні сервіси дедалі частіше використовуються організаціями для забезпечення контролю над даними, підвищення рівня безпеки та відповідності нормативним вимогам. Проте збільшення обсягів даних призводить до зростання навантаження на обчислювальні ресурси та системи зберігання.

Традиційні підходи до організації сховищ даних не завжди забезпечують необхідну продуктивність і масштабованість у приватних хмарних середовищах. Тому виникає необхідність застосування сучасних методів оптимізації, що дозволяють зменшити витрати ресурсів та підвищити швидкість систем. У роботі розглядаються основні методи підвищення ефективності зберігання та обробки великих обсягів даних у приватних хмарних сервісах.

Результати дослідження

Одним із ключових методів оптимізації зберігання даних у приватних хмарних сервісах є дедуплікація. Вона полягає у виявленні та усуненні дубльованих фрагментів даних, що дозволяє суттєво зменшити обсяг використовуваного дискового простору. За даними досліджень, використання дедуплікації може знизити обсяг сховища на 30–60 % залежно від типу даних [1].

Ще одним важливим методом є стиснення даних, яке зменшує фізичний розмір інформації без втрати її змісту. Стиснення дозволяє знизити навантаження на мережу та прискорити передачу даних між вузлами приватної хмари. Однак надмірне використання цього методу може збільшити обчислювальні витрати, що необхідно враховувати при проектуванні систем [2].

Для підвищення продуктивності доступу до даних широко застосовується партиціювання, тобто поділ великих масивів даних на логічні частини. Це дозволяє обробляти лише необхідні фрагменти інформації, зменшуючи час виконання запитів та навантаження на сервери баз даних [3].

Ефективним засобом прискорення обробки даних є кешування, яке полягає у збереженні часто використовуваних даних у швидкій пам'яті. Використання кешування в приватних хмарних сервісах може зменшити затримки доступу до даних у кілька разів, особливо при роботі з аналітичними запитами [4].

Окрему роль відіграють розподілені системи обробки даних, такі як Hadoop та Apache Spark. Вони дозволяють паралельно обробляти великі обсяги інформації, що суттєво підвищує масштабованість і продуктивність приватних хмарних рішень [5].

Табл. 1. Порівняльна характеристика методів оптимізації

Метод	Переваги	Недоліки	Сфера застосування
Дедуплікація	Значне зменшення обсягу даних	Додаткові витрати на аналіз	Резервне копіювання, архіви
Стиснення	Менше навантаження на сховище і мережу	Збільшення обчислень	Передача та зберігання даних
Партиціювання	Швидший доступ до даних	Ускладнення структури БД	Великі бази даних
Кешування	Зменшення затримок доступу	Обмежений обсяг кешу	Часті запити
Розподілена обробка	Висока масштабованість	Складність налаштування	Аналітика великих даних

Аналіз показує, що найбільшої ефективності можна досягти шляхом комбінування декількох методів оптимізації залежно від типу навантаження та архітектури приватної хмари.

Висновки

У результаті дослідження встановлено, що підвищення ефективності зберігання та обробки великих обсягів даних у приватних хмарних сервісах можливе завдяки застосуванню сучасних методів оптимізації. Дедуплікація та стиснення дозволяють суттєво зменшити обсяг даних, партиціювання та кешування — підвищити швидкість доступу, а розподілені системи обробки забезпечують масштабованість і високу продуктивність. Комплексне використання розглянутих методів є доцільним при проектуванні ефективних приватних хмарних інфраструктур.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Таненбаум Е. С., ван Стеен М. Розподілені системи. Принципи та парадигми. 2-ге вид. К.: Видавнича група BHV, 2013. 704 с.
2. Sayood K. Introduction to Data Compression. 5th Edition. Morgan Kaufmann, 2018. 744 p.
3. Garcia-Molina H., Ullman J. D., Widom J. Database Systems: The Complete Book. 2nd Edition. Pearson, 2008. 1119 p.
4. Pacheco P. Parallel Programming with MPI. Morgan Kaufmann, 2017. 384 p.
5. White T. Hadoop: The Definitive Guide. 4th Edition. O'Reilly Media, 2015. 756 p.

Науковий керівник: **Кисюк Дмитро Васильович** – старший викладач, кафедра обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця, e-mail: kneimad@gmail.com

Сірак Владислав Олегович – студент групи 1KI-25мс, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: vladsirak9@gmail.com

Supervisor: **Kysyuk Dmytro V.** – Senior Lecturer, Department of Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: kneimad@gmail.com

Sirak Vladislav O. – student of group 1KI-25ms, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: vladsirak9@gmail.com