

ОПТИМІЗАЦІЯ У СФЕРІ ЗБЕРІГАННЯ ДАНИХ В ХМАРНИХ СЕРЕДОВИЩАХ: ПОШУК КОМПЛЕКСНОГО НАБОРУ КРИТЕРІЇВ

Карташов А.Д., Глоба Л.С.

*Навчально-науковий інститут телекомунікаційних
систем КПІ ім. Ігоря Сікорського, Україна
E-mail: kartashov.anton.ukr@gmail.com*

OPTIMIZATION IN CLOUD COMPUTING: SEARCHING FOR A COMPREHENSIVE SET OF CRITERIA

Finding a comprehensive set of criteria for optimizing distributed storage and data access in a multi-cloud environment.

Не зважаючи на те, що хмарні обчислення все ще знаходяться в стадії постійного розвитку, їх широке поширення дало змогу багатьом дослідникам та компаніям почати оцінювати реальні та подальші потенційні виклики, з якими стикається ця технологія [1].

Дані більше не створюються лише в локальних центрах обробки даних. Обсяг даних, які створюються у хмарі та за допомогою нових технологій - таких як Інтернет речей (IoT) та розподілені обчислення (edge computing) - продовжує зростати. Успішне управління даними потребує всебічного погляду на зберігання даних, як в локальних, так і в хмарних архітектурах.

Одним з ключових викликів, з яким стикаються користувачі хмарних обчислень, є вибір найефективніших та найбільш економічних методів зберігання та доступу до даних, уникнення залежності від постачальника. Існуючі критерії для оптимізації розподілу даних в багатьох хмарних середовищах охоплюють лише підмножину вимог користувачів. Незважаючи на попередні спроби дослідити ці критерії, постійний подальший розвиток технології хмарних обчислень вимагає всебічного переоцінювання таких підходів. [2]

На основі проведеного огляду літератури та сучасних стандартів хмарних обчислень для зберігання та доступу ми визначаємо складний набір критеріїв для зберігання даних у багатьох хмарах, який включає в себе розгляд широкого спектру факторів, які можуть впливати на розташування, управління та отримання даних у різних постачальників хмарних послуг. [3] Таблиця 1 показує комплексний набір критеріїв, які потрібно враховувати у майбутніх дослідженнях.

Розглянемо для прикладу визначення та метод розрахунку Індекса Переносу Даних (Data Portability Index).

Визначення: Індекс Переносу Даних (Data Portability Index) вимірює ступінь складності, з якою дані можуть бути безперешкодно переміщені або передані між різними хмарними середовищами або системами. Це охоплює можливість міграції, передачі та доступу до даних на різних платформах та провайдерах хмарних послуг.

Одиниці вимірювання: Зазвичай індекс представляється у числовій шкалі, де вищі значення вказують на більшу переносимість даних. Наприклад, масштаб може становити від 0 до 100.

Таблиця 1. Набір критеріїв для розподілу зберігання даних у багатьох хмарах.

№	Criteria Category	Specific Criteria	Possible Measurement Metric
1	Data Accessibility Criteria	Latency Requirements	Milliseconds (ms)
2		Redundancy and Availability	Availability Percentage (%)
3		Data Consistency	Data Consistency Index
4		Data Encryption	Encryption Strength (e.g., AES-256)
5	Cost and Resource Utilization Criteria	Cost Efficiency	Cost per GB/month (\$)
6		Resource Allocation	Resource Utilization (%)
7		Data Lifecycle Management	Percentage of Archived Data (%)
8	Data Type and Format Criteria	Data Classification	Data Classification Score
9		Data Format	Data Format Compatibility
10	Compliance and Security Criteria	Regulatory Compliance	Compliance Audit Score
11		Data Ownership	Data Ownership Policy Adherence
12		Security Protocols	Security Protocol Strength
13	Scalability and Performance	Scalability	Scalability Factor
14		Performance Metrics	Throughput (requests/second)
15	Data Migration and Interoperability Criteria	Data Portability	Data Portability Index
16		Interoperability	Interoperability Score
17	Vendor Lock-In and Vendor Criteria	Vendor Lock-In Mitigation	Lock-In Reduction Score
18		Vendor Reputation	Vendor Reputation Rating
19	Disaster Recovery and Backup	Recovery Time Objective (RTO)	Recovery Time Objective (RTO, hours)
20		Recovery Point Objective (RPO)	Recovery Point Objective (RPO, hours)
21		Data Backup Frequency	Frequency (e.g., per day, per week)
22		Backup Storage Redundancy	Redundancy Level (e.g., dual-site)
23	Monitoring and Reporting	Monitoring Tools	Tool Effectiveness (e.g., Score)
24		Reporting	Reporting Accuracy (e.g., Percentage)
25	Sustainability	Environmental Impact	Carbon Emission Reduction (%)
26		Energy Efficiency	Energy Usage (kWh)
27		Resource Sustainability	Resource Conservation Index

Вимірювання Індексу Переносу Даних включає оцінку різних факторів, які сприяють складності переміщення та доступу до даних на різних середовищах.

Запропоновані можливі способи вимірювання Індексу Переносу Даних:

1. Швидкість Передачі Даних:

а. Вимірювання: Швидкість, з якою дані можуть бути передані між різними платформами або провайдерами хмар.

б. Метрики: Mbps або GBps (Мегабіти або Гігабіти в секунду).

2. Сумісність Форматів Даних:

- a. Вимірювання: Ступінь сумісності форматів даних між вихідною та цільовою системами.
 - b. Метрики: Оцінка сумісності на основі стандартизованих форматів даних.
3. Підтримка API:
- a. Вимірювання: Наявність та сумісність інтерфейсів програмування додатків (API) для передачі даних.
 - b. Метрики: Оцінка сумісності API.
4. Інструменти Міграції Даних:
- a. Вимірювання: Наявність та ефективність інструментів, призначених для міграції даних.
 - b. Метрики: Оцінка ефективності інструментів на основі функціональних можливостей, зручності використання та частки успішних операцій.
5. Обробка Метаданих:
- a. Вимірювання: Здатність обробляти та зберігати метадані під час передачі даних.
 - b. Метрики: Оцінка збереження метаданих.
6. Вплив Відмови сервісу:
- a. Вимірювання: Вплив перерв у роботі системи під час передачі даних.
 - b. Метрики: Тривалість та частота відмов.
7. Вартість Передачі Даних:
- a. Вимірювання: Фінансові наслідки переміщення даних між різними середовищами.
 - b. Метрики: Вартість за ГБ/місяць за передачу даних.

Результатом Індексу Переносу Даних є загальний погляд на те, наскільки легко та ефективно дані можуть бути переміщені та доступні на різних платформах або провайдерах хмарних послуг.

Створення комплексної формули для Індексу Переносу Даних передбачає поєднання кількох факторів і надання відповідних ваг для відображення їх відносної важливості. Наведемо узагальнену формулу, яка включає ключові метрики:

$$\text{Data Portability Index (DPI)} = \sum_{i=1}^n W_i \times \left(\frac{M}{\text{Max}_i} \right),$$

де: n - кількість розглянутих метрик; W_i - вага, надана i -й метриці; M_i - виміряне значення i -ї метрики; Max_i - максимально досяжене значення для i -ї метрики; Таким чином, в дослідженні будуть проаналізовані і інші критерії оптимізації та дано їх визначення і методи розрахунку.

Література

1. Tomarchio, O., Calcaterra, D. & Modica, G.D. Cloud resource orchestration in the multi-cloud landscape: a systematic review of existing frameworks. J Cloud Comp 9, 49 (2020). <https://doi.org/10.1186/s13677-020-00194-7>.
2. Salil Bharany & Sandeep Sharma & Osamah Ibrahim Khalaf & Ghaida Muttashar Abdulsahib & Abeer S. Al Humaimeedy & Theyazn H. H. Aldhyani & Mashaal Maashi & Hasan Alkahtani, 2022. "A Systematic Survey on Energy-Efficient Techniques in Sustainable Cloud Computing," Sustainability, MDPI, vol. 14(10), pages 1-89, May.
3. Anton Kartashov and Larysa Globa Overview of the Approaches to Managing Distributed Storage and Access to Cloud Data//Proceedings of International Conference on Applied Innovation in IT. Volume 11, Issue 2, pp. 19-29. (DOI:10.25673/112990).