УДК 004. 6:004.77 **doi: 10.26906/SUNZ.2024.4.084**

А. Касерес, Л. С. Глоба

Національний технічний університет України «КПІ імені І. Сікорського», Київ, Україна

ПІДХІД ДО ФОРМУВАННЯ МУЛЬТИХМАРНОГО СЕРЕДОВИЩА НА ОСНОВІ БАГАТОКРИТЕРІАЛЬНОГО АНАЛІЗУ ТА ОНТОЛОГІЧНОГО МОДЕЛЮВАННЯ

Анотація. На сьогоднішній день використання декількох хмар для збереження і обробки даних вже є розповсюдженою практикою по всьому світу. Кожний постачальник хмарних ресурсів пропонує для цього різні рішення, що ускладнює вибір користувачеві постачальника послуг, адже він, у свою чергу, прагне досягти максимальної ефективності за найменших витрат. У даній роботі досліджуються багатохмарні парадигми: мультихмарні шлюзи зберігання даних, платформи керування даними та хмарно-незалежні бібліотеки. В даному дослідженні пропонується на основі багатокритеріального аналізу (MCDA) методами AHP і TOPSIS обрати кращий підхід для корпоративних і академічних рішень. Критеріями слугують вартість, зусилля щодо впровадження, безпека і продуктивність. Крім того, запропоновано онтологічну модель для формування гнучкого сценарію вибору мультихмарних підходів, оригінальність якої полягає у використанні вагових коефіцієнтів отриманих з практичних експериментів із застосуванням AHP та TOPSIS методів, та сфокусованих на специфіці використання мультихмарних середовищ у корпоративному та академічному середовищах. Це дозволяє отримати найбільш ефективні способи поєднання хмарних сервісів для конкретного сценарію застосування. Проведені експериментальні дослідження дозволяють зробити висновок, що для академічних задач найбільше підходять мультихмарні шлюзи зберігання даних, проте для корпоративних середовищ - хмарно-незалежні бібліотеки і платформи керування даними.

Ключові слова: хмарні обчислення, мульти-хмара, сумісність даних, MCDA, онтологічна модель.

Вступ

У сфері хмарних обчислень багатохмарна архітектура є ефективним рішенням для компаній, які прагнуть отримати вигоду, суміщаючи переваги різних постачальників хмарних послуг. Відповідно до звіту Flexera 2024, спостерігається зростання інтеграції даних між хмарами до 45% з 37% у 2023-2024 роках, оскільки організації шукають найкращий варіант для прикладних програмних систем та аналізу даних [1]. У цьому звіті підкреслюється тенденція до зростання впровадження багатохмарних стратегій, а також вказується на вирішальну роль взаємодії даних у повному використанні потенціалу багатохмарних архітектур.

Зростання впровадження мультихмарних технологій разом із зростаючим попитом на ефективну сумісність даних створює складну проблему для організацій. Різноманітність хмарних платформ, кожна з яких має специфічні для постачальника API, служби та механізми безпеки, збільшує складність мультихмарних архітектур [2]. Ця складність вимагає структурованого та стратегічного підходу до прийняття рішень. Рішення щодо сумісності мають не тільки відповідати технічним вимогам інтеграції, але й відповідати ширшим організаційним цілям щодо масштабованості, продуктивності та вартості [3]. Використання багатохмарного підходу потребує методів, які дозволяють організаціям застосовувати найбільш ефективні стратегії. Серед доволі великої кількості підходів мультихмарної взаємодії [4-6] організаціям необхідно обрати саме ті, які найкращим чином відповідають їхнім потребам. Основними критеріями, які можуть бути важливі на етапі вибору тих чи інших засобів можна назвати такі: швидкодія, безпека, складність впровадження, вартість. Якщо швидкодію та вартість простіше розрахувати, то такі параметри як безпека та складність впровадження є доволі непростими з точки зору їх визначення, для чого застосовують різноманітні інтелектуальні методи. Такі обчислення в реальних умовах є затратними.

Таким чином, метою даної роботи є створення онтологічної моделі на основі багатокритеріального аналізу, яка дозволить визначитися з гнучким сценарієм формування набору інструментарію для реалізації мультихмарного доступу в конкретній організації. Онтологічна модель визначає оптимальний за комплексним критерієм набір технологій для академічних і корпоративних рішень. У якості методів багатокритеріального аналізу використано метод TOPSIS, результати якого порівняно з аналогічним методом AHP. У якості досліджуваних підходів обрано мультихмарні шлюзи зберігання даних (на прикладі S3Proxy), платформи керування даними (Apache NiFi) та хмарно-незалежні бібліотеки (Apache Libcloud).

Огляд сучасного стану підходів мультихмарної взаємодії

В даний час роль мультихмарних технологій продовжує зростати. Чисельні статті говорять про значні переваги їх використання [7–9]: незалежність від постачальника; економічна ефективність; покращена безпека. Але разом з тим, перш ніж розгортати мультихмарну стратегію, організація має врахувати такі фактори: вибір інструментарію; підготовку персоналу; моніторинг і оновлення налаштувань.

Наразі існує досить багато мультихмарних підходів. Розглянемо деякі з них.

Інструменти стандартизації API. Це група рішень, які дозволяють приховати відмінності API різних хмарних постачальників, тобто використовувати той самий робочий процес для керування декількома провайдерами, полегшуючи керування хмарами [10, 11]. До вказаних інструментів можна віднести JClouds, Pulumi, Apache Libcloud, Terraform та інші.

Мультихмарні шлюзи зберігання (Multi-Cloud Storage Gateways (MSG)). Дозволяють отримати доступ до хмарного сховища між різними постачальниками через єдиний інтерфейс, тобто діють як міст між локальними програмами та віддаленими хмарними

 © Касерес А., Глоба Л. С., 2024

центрами зберігання [12]. Окрім ліквідації розриву між застарілими рішеннями для зберігання даних і хмарними технологіями вони надають низку додаткових переваг. Серед таких – спрощений доступ до файлів для користувачів, спрощене керування сховищем і зниження витрат [13]. Лідерами на ринку MSG можна назвати: web-сервіси Amazon, Microsoft, Oracle, NetApp, Inc, Dell EMC. Серед прикладів можна виділити MinIO, S3Proxy, Hitachi Content Platform Anywhere Edge, AWS Storage Gateway та інші.

Контейнерізація та оркестровка контейнерів. Оркестровка хмарних ресурсів (процес автоматизованого керування кількома робочими навантаженнями в кількох хмарних рішеннях) стосується складних операцій, таких як вибір, розгортання, моніторинг і керування ресурсами під час виконання [14]. Контейнеризація дозволяє програмам працювати узгоджено в різних хмарних середовищах. Інструменти оркестровки контейнерів додатково полегшують керування цими контейнерами в хмарах. Представниками даного підходу є Docker Swarm, Google Kubernetes, Apache Mesos, тощо.

Платформи керування даними (Data Management Platform (DMP)). Такі платформи дозволяють переміщувати, перетворювати та керувати даними в кількох хмарних середовищах без прив'язки до певного постачальника хмарних технологій [15]. Підходять для організацій, яким необхідно обробляти велику кількість різнорідних даних. Прикладами є Azure Synapse Analytics, Google Cloud BigQuery, Talend, Apache NiFi.

Також серед менш поширених підходів мультихмарної взаємодії можна назвати Multi-Cloud Service Mesh, Гібридну хмару та багатохмарну мережу (Hybrid Cloud and Multi-Cloud Networking), Міжхмарні конвеєри CI/CD (Cross-Cloud CI/CD Pipelines).

Але разом з тим, вибір конкретного технічного рішення в конкретних умовах їх застосування вимагає окремих досліджень через різноманітність вимог до продуктивності, безпеки та витрат, а також необхідність враховувати попередній досвід використання аналогічних рішень у подібних середовищах.

В дослідженні зосередимось на прикладах Apache Libcloud (хмарно-незалежна бібліотека), S3Proxy (MSG) та Apache NiFi (DMP), тому що дані технології є open-source, тобто доступні для безкоштовного використання, є гнучкими – дозволяють адаптувати програмне забезпечення кінцевого користувача до структури мультиклаудного середовища під різні потреби і націлені на автоматизацію бізнес-процесів в ньому.

Вибір методу багатокритеріального аналізу

При виборі мультихмарної стратегії організацію цікавить ефективність, яку можливо від цього отримати, а це безпосередньо пов'язано з вартістю обслуговування, швидкістю обробки даних, безпечністю, а також складністю налаштувань системи [16]. Таким чином, кожен з цих підходів оцінюватиметься за критеріями продуктивності, вартості, зусиль впровадження і безпеки.

Розглянемо чотири критерії, за якими необхідно обрати той підхід, який найкраще задовольнить потреби організації. Кожна організація ставить свої вимоги, тому необхідно створити модель, яка б враховувала специфічні потреби конкретного користувача.

Розглянемо побудову такої моделі на прикладі потреб академічних і корпоративних організацій. Очевидно, що у бюджетних організацій типу навчальних і медичних закладів будуть інші пріоритети, ніж у великих корпорацій, тому пропонується застосувати ранжування означених критеріїв.

З огляду на звіт [1] вартість є найбільш вагомим фактором для організацій різних типів, щодо інших параметрів запропоновано на основі практичного досвіду таку послідовність від найвищого до найнижчого пріоритетів.

Академічні рішення: вартість, зусилля впровадження, безпека і продуктивність.

Корпоративні рішення: вартість, безпека, продуктивність, зусилля впровадження.

У дослідженні [17] було використано метод AHP для створення моделі вибору мультихмарного підходу, проте не було розділення пріоритетів для корпорації і академії. Крім того, у поточному дослідженні розглянуто додатковий метод для порівняння отриманих теоретичних результатів та гарантування їх достовірності. Виходячи з огляду джерел [18-21] обрано метод TOPSIS. Його суть полягає у виборі тієї альтернативи, яка є найближчою до ідеальної. Також він дозволяє враховувати важливість кожного критерію, так само як і метод AHP.

Постановка задачі

Серед підходів мультихмарної взаємодії необхідно обрати найбільш оптимальний для академічних і корпоративних задач, спираючись на критерії вартості, продуктивності, безпеки і зусиль щодо впровадження.

Вхідні дані:

Альтернативи: *A={CAL, SG, DMP},*

де *SG* – мультихмарні шлюзи зберігання даних, *DMP* – платформи керування даними, *CAL* – хмарно-незалежні бібліотеки.

Критерії: *CR={C, P, S, IE},*

де *C* – вартість, *P* – продуктивність (у значенні швидкості доступу до файлів), *S* – безпека, *IE* – зусилля щодо впровадження.

Вартість *C* може бути обчислена на основі таких складових: витрати на зберігання, витрати на екземпляр віртуальної машини та витрати на запити:

$$\mathrm{C} = \sum_{i=1}^{K} \mathrm{C}_i, \quad (1)$$

де *K* – кількість цінових факторів.

Продуктивність *P* у нашому випадку – це середній час передавання файлів різного розміру:

$$P = \frac{\sum_{i=1}^{n} T_i S_i}{\sum_{i=1}^{n} S_i}, \quad (2)$$

де S_i – розмір *i*-го файлу, T_i – час доступу до *i*-го файлу, *n* – кількість файлів.

Для оцінки безпеки варто врахувати такі аспекти як автентифікація та авторизація, шифрування даних, забезпечення відмовостійкості. У дослідженні використано ваги для таких параметрів безпеки *S*: W_a – керування ідентифікацією та доступом, W_s – шифрування, W_l – логування та моніторинг.

$$S = W_a + W_s + W_l. \quad (3)$$

Ваги безпекових параметрів наведені в табл. 1. Пропонується призначити ваги таким чином: 1 – базовий рівень безпеки, 2 – середній рівень, 3 – розширений. Зусилля впровадження *IE* – це час і ресурси, необхідні для розгортання рішення у мільтихмарному середовищі. Для визначення критерію складності розгортання було залучено три провідні консалтингові компанії з розробки хмарних технологій: Crayon [22], Nordcloud [23] і Tech5 [24].

Їх оцінка представлена у табл. 2 у вигляді витрачених людино-годин.

Таблиця 1 – **Ваги параметрів безпеки**

Параметр безпеки	Хмарно-незалежні бібліотеки (вага)	Мультихмарні шлюзи зберігання даних (вага)	Платформи керування даними (вага)
Керування ідентифікацією та доступом	1	1	3
Шифрування	1	1	2
Логування та моніторинг	1	1	3

Таблиця 2 – **Оцінка зусиль впровадження**

Консалтінгова компанія	Хмарно-незалежні бібліотеки (людино-години)	Мультихмарні шлюзи зберігання даних (людино-години)	Платформи керування даними (людино-години)
Nordcloud	30 - 50	40 - 60	40 - 60
Crayon	40 - 50	50 - 60	50 - 60
Tech-5	50 - 70	30 - 40	30 - 40

Для академічних і корпоративних задач вказані критерії мають різні пріоритети. Так, вартість є найважливішим параметром для обох видів задач. Проте, для академічних установ наступним за важливістю параметром є зусилля впровадження через обмеженість доступу до необхідних спеціалістів або їх навчання. Для корпоративних установ цей фактор не є проблематичним, тому має найменшу вагу. Наступним параметром обрано безпеку, так як в академічному середовищі також можливо обробляти досить вагомі дані, проте вимоги до безпеки є менш жорсткими, ніж у корпораціях. В той же час останні мають високі вимоги до стандартів безпеки і по пріоритету дана характеристика буде займати наступне місце після вартості. Продуктивність для академічних задач опиняється на останньому місті, для корпорації – займає місце після безпеки.

Таким чином отримуємо таку пріоритетність критеріїв:

- Для академічних рішень від найбільш вагомого: *C, IE, S, P*.
- Для корпоративних рішень від найбільш вагомого: *C, S, P, IE*.

Вихідні результати:

Для визначення оптимального підходу обчислити загальну оцінку для академічних і корпоративних рішень:

$$E_j = \sum_i w_i * x_{ij}, \quad (4)$$

де w_i – вага критерію CR_i, x_{ij} – оцінка альтернативи A_j за критерієм CR_i

Вибираємо підхід з максимальною оцінкою для кожної категорії задач.

$$Opt_{korp} = max\ E_j,\ \ Opt_{akad} = max\ E_j\ . \quad (5)$$

Оптимальний підхід пропонується обирати за допомогою методів багатокритеріального аналізу AHP і TOPSIS.

Для обох методів першим кроком є створення таблиці (табл. 1) з альтернатив (рядки таблиці) і критеріїв (стовбці таблиці).

Алгоритм методу AHP:

1. Для кожного критерію створюється матриця попарних порівнянь альтернатив, де кожний елемент набуває значення від 1 до 9, причому однакові альтернативи приймають значення 1. Іншими цифрами показано перевагу однієї альтернативи над іншою. Чим більше число, тим більша перевага. Так як критеріїв чотири, має бути чотири матриці.
2. Далі створюється матриця попарного порівняння критеріїв, де аналогічно оцінюється перевага одного критерія над іншим. Для корпоративного і академічного підходів ці матриці будуть різними.
3. Нормалізація матриць.
4. Обчислення вектору пріоритетів, який покаже ранжування альтернатив (найбільшому значенню відповідає краща альтернатива).
5. Перевірка чутливості отриманих параметрів. Якщо значення не задовольняють певному пороговому значенню, необхідно провести корегування матриць попарного порівняння.

Алгоритм методу TOPSIS:

1. Нормалізація початкової матриці.
2. Розподілення значень ваги пріоритетів у відповідності до важливості критеріїв.
3. Обчислення нормалізованої зваженої матриці. У нашому випадку їх буде дві: для корпорації і академії.
4. Вибір найкращих і найгірших рішень з отриманих таблиць.
5. Обчислення відстані до найкращого і найгіршого значення для кожної альтернативи.
6. Визначення близькості кожної альтернативи до ідеального значення. Найбільше значення дає найвищий пріоритет.

Використання методу TOPSIS для аналізу мультихмарних підходів

Для кожного підходу були визначені параметри, наведені у табл. 3. Параметр продуктивності визначено як середній час доступу до файлів різного розміру.

Таблиця 3 – **Загальна таблиця критеріїв**

Підхід	Середня продуктивність, с	Середня вартість за ТВ, USD	Зусилля впровадження, людино-години	Безпека (вагові коефіцієнти)
Хмарно-незалежні бібліотеки	0,27	117,3	48	3
Мультихмарні шлюзи зберігання даних	0,34	121,7	47	3
Платформи керування даними	7,24	126,1	87	8

У табл. 3 представлено середньозважене значення часу передавання для файлу розміром 1 МБ. Вартість розраховано на основі моделі ціноутворення AWS [25], де було враховано витрати на зберігання, на екземпляри та на запити. У таблиці представлено середньозважене значення за 1 ТБ. Зусилля впровадження визначено на основі думки експертів з області розробки. Даний параметр визначено як кількість годин, необхідних розробникам для розгортання даного підходу. Безпековий параметр визначений як сума ваг, які було призначено кожній технології на основі параметрів управління ідентифікацією та доступом, шифрування, логування та моніторингу.

Суть методу TOPSIS полягає у виборі тієї альтернативи, яка є найближчою до ідеальної. Першим кроком для цього є побудова матриці на основі результатів описаних вище експериментів. Маємо три альтернативи (Apache Libcloud, S3Proxy, Apache NiFi) і чотири критерії (Середня продуктивність, Середня вартість за ТБ, Зусилля впровадження і Безпека), тобто початкова матриця матиме вигляд табл. 3.

Наступний крок – нормалізація матриці (табл. 4), наведеної у табл. 3:

$$r_{i,j} = \frac{a_{i,j}}{\sqrt{\sum_{k=1}^{m} a_{ik}^2}} \quad (6)$$

де a_{ij} – значення елементів початкової матриці, r_{ij} – нормалізоване значення.

Призначаємо ваги критеріїв у відповідності до визначених раніше пріоритетів: найвищому пріоритету буде відповідати вища вага. Наприклад, для академії вартість є найбільш важливою, тому W_C=0,4. Найменш важливою є продуктивність, тому W_P=0,1. При цьому сума ваг має дорівнювати 1. Розподілимо ваги таким чином.

Для академічних рішень: W_P=0,1; W_C=0,4; W_{IE}=0,3; W_S=0,2 (табл. 5).

Для корпоративних рішень: W_P=0,2; W_C=0,4; W_{IE}=0,1; W_S=0,3 (табл. 6).

Обчислюємо значення нормалізованої зваженої матриці за формулою:

$$v_{ij} = W_i * r_{ij}, \quad (7)$$

Таблиця 4 – **Нормалізовані параметри**

Підхід	Середня продуктивність, с	Середня вартість за ТВ, USD	Зусилля впровадження, людино-години	Безпека (вагові коефіцієнти)
Хмарно-незалежні бібліотеки	0,037	0,556	0,437	0,331
Мультихмарні шлюзи зберігання даних	0,047	0,577	0,428	0,331
Платформи керування даними	0,998	0,598	0,791	0,883

Таблиця 5 – **Нормалізована зважена матриця для академічного підходу**

Підхід	Середня продуктивність, с	Середня вартість за ТВ, USD	Зусилля впровадження, людино-години	Безпека (вагові коефіцієнти)
Хмарно-незалежні бібліотеки	0,004	0,222	0,131	0,066
Мультихмарні шлюзи зберігання даних	0,005	0,231	0,128	0,066
Платформи керування даними	0,100	0,239	0,237	0,177

Таблиця 6 – **Нормалізована зважена матриця для корпоративного підходу**

Підхід	Середня продуктивність, с	Середня вартість за ТВ, USD	Зусилля впровадження, людино-години	Безпека (вагові коефіцієнти)
Хмарно-незалежні бібліотеки	0,007	0,222	0,044	0,099
Мультихмарні шлюзи зберігання даних	0,009	0,231	0,043	0,099
Платформи керування даними	0,200	0,239	0,079	0,265

Далі визначаємо за формулою 8 найкраще і найгірше рішення (табл. 7).

Таблиця 7 – **Значення кращого і гіршого варіантів**

Підхід	P	C	IE	S
V_{acad}^{+}	0,004	0,224	0,128	0,177
V_{acad}^{-}	0,100	0,237	0,237	0,066
V_{corp}^{+}	0,007	0,224	0,043	0,265
V_{corp}^{-}	0,200	0,237	0,079	0,099

Для параметру швидкодії, вартості і зусиль впровадження обираємо найменше значення, для безпеки – найбільше.

Вибір значень робимо з табл. 5 і 6.

$$V^{+} = (v_1^{+}, v_2^{+}, \dots, v_n^{+}),\ V^{-} = (v_1^{-}, v_2^{-}, \dots, v_n^{-}). \quad (8)$$

Відстань до найкращого і найгіршого рішення визначається за формулами:

$$E_i^{+} = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^{+})^2}\ ; \quad (9)$$

$$E_i^- = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^-)^2} \ . \quad (10)$$

Визначаємо відносну близькість кожної альтернативи до ідеального рішення:

$$R_i^* = \frac{E_i^-}{E_i^+ + E_i^-} \quad (11)$$

Чим більше значення R_i^*, тим краще альтернатива. Отримуємо значення, наведені у табл. 8.

Таблиця 8 – **Результати**

Підхід	R_{acad}^*	Ранг	R_{corp}^*	Ранг
Хмарно-незалежні бібліотеки	0,566	2	0,542	1
Мультихмарні шлюзи зберігання даних	0,567	1	0,539	2
Платформи керування даними	0,431	3	0,458	3

З табл. 8 видно, що для академічних рішень мультихмарні шлюзи зберігання даних мають найбільше значення, тому їм надається перший пріоритет, на другому місці – хмарно-незалежні бібліотеки і останнє місце залишається за платформами керування даними. Для корпоративних рішень бачимо, що найкращий результат мають хмарно-незалежні бібліотеки, далі мультихмарні шлюзи зберігання даних і найгірший результат знову показали платформи керування даними.

Результати експерименту

Аналіз отриманих результатів. Отримані результати методом TOPSIS показують, що для академічних рішень найкращим підходом є мультихмарні шлюзи зберігання даних. Для корпоративних рішень це хмарно-незалежні бібліотеки. Втім, обидва ці підходи мають зовсім невеликий розрив. Найгірший показник мають платформи керування даними.

У дослідженні [17] вже було розглянуто створення багатокритеріальної моделі методом AHP – аналітичний процес ієрархії, тому нижче наведено дані, необхідні для побудови моделі для академічних і корпоративних рішень. У цьому методі для кожного критерію необхідно створити порівняльні матриці. Для цього використовується шкала від 1 до 9. Чим краща альтернатива, тим більшу вагу вона отримує. Рівні альтернативи отримують 1 [26]. Наприклад, матриця порівнянь продуктивності для наведених даних має вигляд, наведений у табл. 9. Для корпоративних і академічних рішень визначимо матриці альтернатив (табл. 10, 11). Після цього розраховуємо вектор пріоритетів для визначення рангу підходів.

Таблиця 9 – **Матриця порівнянь для продуктивності**

Підхід	Хмарно-незалежні бібліотеки	Мультихмарні шлюзи зберігання даних	Платформи керування даними
Хмарно-незалежні бібліотеки	0,007	0,222	0,044
Мультихмарні шлюзи зберігання даних	0,009	0,231	0,043
Платформи керування даними	0,200	0,239	0,079

Таблиця 10 – **Матриця альтернатив для академічних рішень**

Критерій	Продуктивність	Вартість	Зусилля впровадження	Безпека
Продуктивність	1	1/4	1/3	1/2
Вартість	4	1	2	3
Зусилля впровадження	3	1/2	1	2
Безпека	2	1/3	1/2	1

Таблиця 11 – **Матриця альтернатив для корпоративних рішень**

Критерій	Продуктивність	Вартість	Зусилля впровадження	Безпека
Продуктивність	1	1/3	2	1/2
Вартість	3	1	4	2
Зусилля впровадження	1/2	1/4	1	1/3
Безпека	2	1/2	3	1

Таблиця 12 – **Вектори пріоритетів для академічних і корпоративних рішень**

Підхід	Академічне рішення	Ранг	Корпоративне рішення	Ранг
Хмарно-незалежні бібліотеки	0,3132	3	0,3058	3
Мультихмарні шлюзи зберігання даних	0,3715	1	0,3072	2
Платформи керування даними	0,3153	2	0,3869	1

На основі проведеного аналізу методом AHP кращим вважається варіант із найвищим балом. Згідно з оцінками, вектор пріоритету для академічних завдань становить [0,3132 0,3715 0,3153]. Таким чином, багатохмарні шлюзи сховищ стають оптимальним вибором для академічних установ. Вектор пріоритету для корпоративних завдань становить [0,3058 0,3072 0,3869], що вказує на те, що DMP є кращим варіантом для таких середовищ.

Порівняння результатів MCDA/AHP і MCDA/TOPSIS

Зведемо ранжування за обома методами у табл. 13, 14.

Таблиця 13 – **Порівняння результатів ранжування підходів для академічних рішень за методами TOPSIS та AHP**

Підхід	Ранг TOPSIS	Ранг AHP
Хмарно-незалежні бібліотеки	2	3
Мультихмарні шлюзи зберігання даних	1	1
Платформи керування даними	3	2

Таблиця 14 – **Порівняння результатів ранжування підходів для корпоративних рішень за методами TOPSIS та AHP**

Підхід	Ранг TOPSIS	Ранг AHP
Хмарно-незалежні бібліотеки	1	3
Мультихмарні шлюзи зберігання даних	2	2
Платформи керування даними	3	1

З табл. 13 видно, що обидва методи показали мультихмарні шлюзи зберігання даних як найкращий результат для академічних рішень, щодо ранжування інших підходів – результати розійшлися.

Для корпоративних рішень підхід хмарно-незалежних бібліотек методом TOPSIS показав найкращий результат, в той час як за AHP варто надати перевагу платформам керування даними. Друге місце за обома методами посідають мультихмарні шлюзи зберігання даних.

Згідно з [19] AHP має досить великий час розрахунків і сильний суб'єктивний вплив, у той час як TOPSIS простіший у реалізації і має менший суб'єктивний вплив на результати. Практичні дослідження підтверджують це, створюючи відповідні моделі для корпоративних і академічних рішень.

Маючи набір альтернатив $A=\{CAL, SG, DMP\}$ і критеріїв $CR=\{C, P, S, IE\}$, формалізуємо дану предметну область, запропонувавши онтологічну модель для формування найкращих стратегій вибору мультихмарних підходів.

Побудова онтологічної моделі

Дослідивши різні методи мультихмарної взаємодії, побудуємо онтологічну модель [27, 28], яка дозволить наочно показати основні поняття і зв'язки вказаної предметної області, та спростити процес формування сценарію вибору найкращого стеку технологій для побудови мультихмарного середовища під умови конкретного підприємства.

Онтологія має різноманітні форми представлення, при цьому складається з ієрархічного опису концептів [29]. Представлення онтології у вигляді орієнтованого графа дозволяє одночасно відображати велику кількість об'єктів і зв'язків між ними [30].

Вузли графа представляють ключові концепції або сутності, зокрема багатохмарні підходи, критерії, значення векторів пріоритетів. Стрілки вказують на зв'язки або взаємодію між поняттями.

Кожен вузол перетікає від однієї сутності до іншої. Параметри утворюють ієрархію оцінювання, де підходи знаходяться на одному рівні, а критерії і очікувана оцінка на інших.

Виходячи з отриманих моделей багатокритеріального аналізу для корпоративних і академічних рішень у зв'язках онтології проставлено ваги пріоритетів критеріїв.

Представимо онтологію у вигляді:

$$O=<C, A, V, I, R>, \quad (9)$$

де C – набір усіх ключових сутностей або концепцій (концептуальні класи), залучених до процесу прийняття рішень щодо вибору багатохмарного підходу; у розглядуваному випадку це мультихмарні шлюзи зберігання даних, платформи керування даними та хмарно-незалежні бібліотеки:

$$C=\{CAL, SG, DMP\}, \quad (10)$$

де SG – мультихмарні шлюзи зберігання даних, DMP – платформи керування даними, CAL – хмарно-незалежні бібліотеки; при використанні інших підходів мультихмарної взаємодії набір C може бути розширено;

A – набір атрибутів і характеристик, що описують класи понять:

$$A=\{C, P, S, IE\}, \quad (11)$$

де C – вартість, P – продуктивність, S – безпека, IE – зусилля щодо впровадження;

V – результуючі значення у вигляді рангу пріоритетів:

$$V=\{V_1, V_2,...,V_n\}, \quad (12)$$

де V_1 – найвищий пріоритет, а V_n – найнижчий;

I – множина екземплярів класів, представляють набір варіантів мульти-хмарної взаємодії:

$$I=\{ S3Proxy, Apache\ NiFi, Apache\ Libcloud\}; \quad (13)$$

R – набір відношень між класами, які вказують на зв'язки або взаємодію між поняттями.

Кожен підхід оцінюється на основі його ефективності щодо кожного критерію, пов'язуючи їх із потенційними результатами.

У представленій онтології можна виділити такі зв'язки:

1. Відношення між концепціями та екземплярами $R(C_i, I_j)$, де кожен екземпляр I_j пов'язаний з концепцією C_i.

2. Відношення між екземплярами та атрибутами $R(I_i, A_j) \rightarrow V_k$, де для кожного екземпляра I_i (наприклад, S3Proxy) існує відношення з атрибутами A_j (наприклад, вартість C, продуктивність P тощо), що призводить до оцінки V_k (результуючого значення).

3. Відношення між атрибутами та результатами $R(A_i, V_j)$ відображають зв'язок між кри-

теріями та відповідними результатами ранжування або пріоритетами.

На рис. 1 представлена описана онтологія у вигляді орієнтованого графу. Верхній рівень вузлів являє собою класи мультихмарних підходів. Наступний рівень відображає конкретні технології. Далі йде рівень оціночних критеріїв. Останній рівень відображає результати.

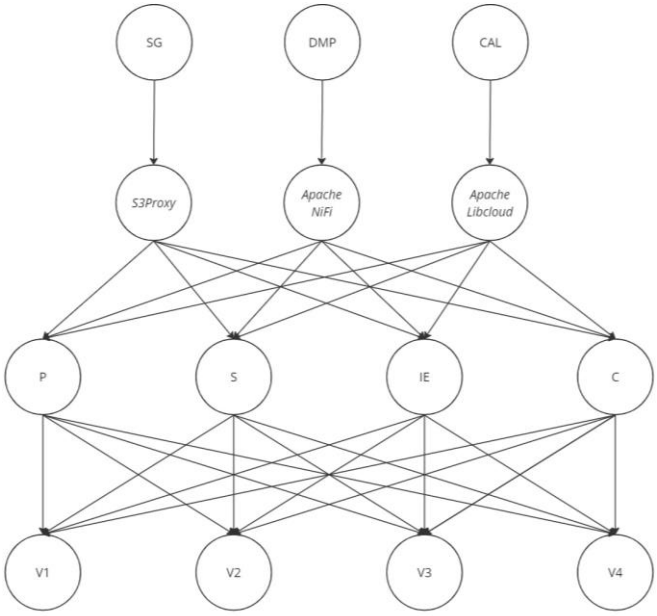

Рис. 1. Онтологічна модель вибору мультихмарного підходу

Таким чином, дана модель узагальнює відомості, необхідні для вибору мультихмарного підходу. Більш того, дану модель можна розширити на кожному ієрархічному рівні, де можна додати інші мультихмарні стратегії та критерії.

Бізнес-процес оцінки альтернативних рішень на основі онтологічної моделі можна розділити на такі етапи:

1. Визначення мети і вибір концепцій під конкретний сценарій.
2. Визначення критеріїв оцінки. В описаній моделі - це продуктивність, безпека, зусилля впровадження і вартість.
3. Ранжування альтернатив на основі оцінок.
4. Для кожної альтернативи проводиться оцінка по вказаним критеріям.
5. Ранжування альтернатив на основі оцінок.
6. Прийняття рішення щодо впровадження конкретного підходу.

Використання онтології дозволяє охопити важливі аспекти прийняття рішення, спрощено і наочно візуалізувати процес оцінки, а також визначитися з гнучким сценарієм формування набору інструментарію для реалізації мультихмарного доступу в конкретній організації.

Висновки

Дана робота аналізує методи мультихмарної взаємодії з метою побудови гнучкого сценарію формування набору інструментарію для реалізації мультихмарного доступу в конкретній організації, для чого застосовано методи багатокритеріальної оптимізації та онтологічна модель.

Виходячи з результатів багатокритеріального аналізу на основі методів AHP і TOPSIS, отримано такі результати: мультихмарні шлюзи зберігання даних найкраще підходять для академічних рішень, в той час як для корпоративних рішень більше підходять хмарно-незалежні бібліотеки і платформи керування даними.

Визначено групу технологій для мультихмарної взаємодії на основі методів TOPSIS і AHP, що дозволяє зробити висновки:

– метод AHP є більш трудомістким (потрібен великий час для побудови і розрахунків) і дуже чутливим до користувацьких налаштувань,

– метод TOPSIS набагато простіший в реалізації, при цьому результати набагато менш чутливі до користувацьких рішень.

У подальших дослідженнях планується зробити аналогічний аналіз для динамічного розподілу обчислювального навантаження у мульти-клауд системах.

Підсумком роботи є узагальнення мультихмарних підходів і зв'язків з їх параметрами, представлене у вигляді онтологічної моделі, що дозволяє підприємствам підвищити ефективність отриманих технічних рішень щодо мультихмарної взаємодії, дає можливість автоматизованим шляхом проводити оцінку альтернативних рішень на основі статистичної інформації, яка визначає наявний досвід роботи в різних хмарних середовищах.

СПИСОК ЛІТЕРАТУРИ

1. "Flexera 2024 State of the Cloud Report," Accessed on: 15.04.2024. [Online]. Available: https://info.flexera.com/CM-REPORT-State-of-the-Cloud.
2. J. Carvalho, D. Vieira, and F. Trinta, "Dynamic Selecting Approach for Multi-cloud Providers," in *Cloud Computing – CLOUD 2018*, M. Luo and L. J. Zhang, Eds., Cham: Springer, 2018, vol. 10967, Lecture Notes in Computer Science, https://doi.org/10.1007/978-3-319-94295-7_3.
3. P. Wang, C. Zhao, W. Liu, Z. Chen, and Z. Zhang, "Optimizing Data Placement for Cost-Effective and High Available Multi-Cloud Storage," *Computing and Informatics*, vol. 39, pp. 51–82, 2020. doi: 10.31577/cai20201-251.
4. J. Alonso, L. Orue-Echevarria, V. Casola, and et al., "Understanding the challenges and novel architectural models of multi-cloud native applications – a systematic literature review," *J. Cloud Computing*, vol. 12, p. 6, 2023. https://doi.org/10.1186/s13677-022-00367-6.
5. M. Boch et al., "A Systematic Review of Data Management Platforms," in *Information Systems and Technologies*, A. Rocha, H. Adeli, G. Dzemyda, and F. Moreira, Eds., Cham: Springer, 2022, vol. 469, Lecture Notes in Networks and Systems, https://doi.org/10.1007/978-3-031-04819-7_2.

6. H. A. Imran et al., "Multi-Cloud: A Comprehensive Review," in *2020 IEEE 23rd International Multitopic Conference (INMIC)*, Bahawalpur, Pakistan, 2020, pp. 1-5, doi: 10.1109/INMIC50486.2020.9318176.
7. S. Tu, "Multicloud Connectivity: A Complete Guide," Megaport, 2022. [Online]. Available: https://www.megaport.com/blog/multicloud-connectivity-complete-guide/
8. "Why Do Companies Use Multi-Cloud?" [Online]. Available: https://prosimo.io/why-do-companies-use-multi-cloud/.
9. N. S. Kravets and M. S. Chernyshov, "Problems of Multi-Cloud Solutions Deploying and Managing," *Bionika Intellekta*, no. 1(96), pp. 63–68, 2021. doi: 10.30837/bi.2021.1(96).10. [Online]. Available: https://openarchive.nure.ua/server/api/core/bitstreams/7ffc0f49-0a95-4228-8bc4-5d0cbcab6fed/content.
10. "Terraform Docs Overview," [Online]. Available: https://developer.hashicorp.com/terraform/docs.
11. "Apache Libcloud," Libcloud. [Online]. Available: https://libcloud.apache.org/.
12. "Cloud Storage Gateway," *Techopedia*, Accessed on: 25.04.2024. [Online]. Available: https://www.techopedia.com/definition/26537/cloud-storage-gateway.
13. "The Top 7 Cloud Storage Gateway Solutions," [Online]. Available: https://expertinsights.com/insights/top-cloud-storage-gateway-solutions/.
14. O. Tomarchio, D. Calcaterra, and G. D. Modica, "Cloud resource orchestration in the multi-cloud landscape: a systematic review of existing frameworks," *J. Cloud Computing*, vol. 9, p. 49, 2020. https://doi.org/10.1186/s13677-020-00194-7.
15. "What is Data Management Platform (DMP)?" *Oracle.* [Online]. Available: https://www.oracle.com/cx/marketing/data-management-platform/what-is-dmp/#how-to-choose-the-right-dmp.
16. M. Farid, H. S. Lim, C. P. Lee, and R. Latip, "Scheduling Scientific Workflow in Multi-Cloud: A Multi-Objective Minimum Weight Optimization Decision-Making Approach," *Symmetry*, vol. 15, no. 2047, 2023. https://doi.org/10.3390/sym15112047.
17. A. Cacerec, L. Globa, "AHP-driven multi-criteria analysis of multi-cloud data management techniques", unpublished.
18. M. Aruldoss, T. M. Lakshmi, and V. P. Venkatesan, "A Survey on Multi Criteria Decision Making Methods and Its Applications," *American Journal of Information Systems*, vol. 1, no. 1, pp. 31-43, 2013. doi: 10.12691/ajis-1-1-5.
19. A. Buriachenko and N. Kuts, "Multi-Criterion Analysis Methods," *ResearchGate*, 2021. [Online]. Available: https://www.researchgate.net/publication/348774582_MULTI-CRITERION_ANALYSIS_METHODS.
20. V. S. Korendovych, "Application of Multi-Criteria Analysis for Priority Selection," *Journal: Collection of the Scientific Papers of the Centre for Military and Strategic Studies of the National Defence University of Ukraine*, no. 2-60, pp. 129-136, 2018. doi: 10.33099/2304-2745/2017-2-60/129-136.
21. M. Hanine, O. Boutkhoum, A. Tikniouine, and et al., "Application of an Integrated Multi-Criteria Decision Making AHP-TOPSIS Methodology for ETL Software Selection," *SpringerPlus*, vol. 5, p. 263, 2016. https://doi.org/10.1186/s40064-016-1888-z.
22. Crayon. [Online]. Available at: https://www.crayon.com.
23. Nordcloud. [Online]. Available at: https://nordcloud.com.
24. Tech-5. [Online]. Available at: https://tech-5.de.
25. AWS, 2024. *AWS Prices*. Available at: https://aws.amazon.com/ru/pricing/?aws-products-pricing.sort-by=item.additionalFields.productNameLowercase&aws-products-pricing.sort-order=asc&awsf.Free%20Tier%20Type=*all&awsf.tech-category=*all
26. I. Linkov and E. Moberg, *Multi-Criteria Decision Analysis: Environmental Applications and Case Studies*. Accessed on: 08.05.2024. [Online]. Available: https://books.google.pl/books?hl=uk&lr=&id=LqdH2G6xk1AC&oi=fnd&pg=PP1&dq=Multiple+Criteria+Decision+Analysis&ots=rqQXIy452V&sig=fFSu7G4k4Lnxv55eO3dRQSWBarI&redir_esc=y#v=onepage&q=Multiple%20Criteria%20Decision%20Analysis&f=false.
27. A. Y. Gladun and K. O. Khala, "Computer Means, Networks, and Systems," *Kzms*, no. 18, pp. 5-10, 2019. [Online]. Available: http://nbuv.gov.ua/UJRN/Kzms_2019_18_3.
28. A. A. Kosolapov and Y. O. Pshinko, "Ontological Models in the Problems of Automation of Sorting Stations," *Artificial Intelligence*, no. 4, pp. 344–352, 2013. [Online]. Available: http://dspace.nbuv.gov.ua/handle/123456789/85188.
29. O. V. Zakharova, "Basic Principles for the Construction of Ontological Graph-Oriented Description of the Application Area," *Problemy Programuvannia*, no. 4, pp. 51-59, 2010. [Online]. Available: http://dspace.nbuv.gov.ua/bitstream/handle/123456789/50356/6_s_51-59.pdf?sequence=1.
30. S. Haiko and V. Prykhodniuk, "Representation of Educational Resources as Interactive Documents," in *Proceedings of the XII International Scientific-Practical Conference «INTERNET-EDUCATION-SCIENCE» (IES-2020)*, Ukraine, Vinnytsia, 26-29 May 2020, pp. 249–251. Vinnytsia: VNTU, 2020.

Received (Надійшла) 12.09.2024
Accepted for publication (Прийнята до друку) 06.11.2024

A multi-criteria and ontology-based approach to multi-cloud environment selection

Anton Caseres, Larysa. Globa

Abstract. Today, multi-cloud environments for data storage and processing have become a standard practice worldwide. Each cloud provider offers own unique solutions, making it challenging for end users to select a service provider, as they aim to maximize efficiency while minimizing costs. This paper explores multi-cloud paradigms, including multi-cloud storage gateways, data management platforms, and cloud-agnostic libraries. The study uses multicriteria decision analysis (MCDA) with AHP and TOPSIS methods to identify the optimal approach for corporate and academic applications, based on criteria such as cost, implementation effort, security, and performance. Additionally, an ontological model is proposed to create a flexible framework for selecting multi-cloud approaches. This model leverages weighting factors derived from practical experiments using AHP and TOPSIS, focusing on the specific requirements of multi-cloud environments in corporate and academic contexts. The model provides the most effective strategies for combining cloud services based on particular application scenarios.

Keywords: cloud computing, multi-cloud, data interoperability, MCDA, ontological model.