

## **PECULIARITIES OF USING THE DATA-CENTRIC APPROACH IN BUILDING AN INFORMATION INFRASTRUCTURE USING CLOUD TECHNOLOGIES**

### **ОСОБЛИВОСТІ ВИКОРИСТАННЯ ДАТАЦЕНТРИЧНОГО ПІДХОДУ ПРИ ПОБУДОВІ ІНФОРМАЦІЙНОЇ ІНФРАСТРУКТУРИ З ВИКОРИСТАННЯМ ХМАРНИХ ТЕХНОЛОГІЙ**

**Vadym Kapylevich<sup>1</sup>**  
**Andrii Fatalchuk<sup>2</sup>**

DOI: <https://doi.org/10.30525/978-9934-26-309-5-9>

В сучасному світі цифрова трансформація відіграє важливу роль у побудові корпоративної інфраструктури, де ІТ сервіси (інформаційні сервіси) відображають сутність, пов'язану з діями, що виконуються на запит або в інтересах споживача.

На цей час у провідних країнах світу вже розроблено національні цифрові стратегії (Digital Strategy), де окреслено технологічну перспективу використання штучного інтелекту (як інформаційних сервісів) через програми розбудови інфраструктури 5G та центрів обробки даних. Також, в деяких країнах запроваджено стратегії даних (Data Strategy), які описують політику створення та роботи з корпоративною моделлю даних. Це пов'язано із зростаючою кількістю даних для обробки, та сучасними тенденціями впровадження датацентричного підходу з використанням хмарних сервісів (сервісів зберігання та обробки даних, сервісів контейнеризації тощо).

Датацентричний підхід – це архітектурний підхід, в якому інформаційна інфраструктура створюється навколо загальнодоступного сховища даних. При такому підході, дані відіграють вагомішу роль ніж програмне забезпечення, та зміщується фокус на створення системи зберігання та обробки даних з єдиним сховищем даних, з фізичним або логічним доступом до нього.

Використовують дані декількох вимірів, об'єднуючи їх у багатовимірні бази даних:

1) Зірка (Star schema) – схема моделювання даних, при якій одна таблиця фактів асоційована із багатьма таблицями розмірності.

---

<sup>1</sup> The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Ukraine

<sup>2</sup> The National Defence University of Ukraine named after Ivan Cherniakhovskiy, Ukraine

2) Сніжинка (Snowflake schema) – схема моделювання даних на основі зіркової схеми, яка нормалізує таблиці розмірності у певну ієрархію. Дані в таких таблицях розбиваються на додаткові таблиці, що в свою чергу утворює форму, схожу на сніжинку [2, с. 1–7].

На відміну від сховищ даних Data Warehouse, попереднього покоління, де дані використовувались, в основному, для загального статистичного аналізу, а метадані не враховувались взагалі, друге покоління сховищ даних вже відрізняється можливістю проведення динамічного аналізу даних [2, с. 1–7]. Архітектура побудови сховищ даних (Data Warehouse 2.0) забезпечує баланс продуктивності навантаження та логічний доступ до даних, безперервно користуючись метаданими, які додатково описують (деталізують) фізично-збережені дані.

Доктринальні вказівки стосовно інтеграції комунікаційних та інформаційних систем (CIS) у спільні операції спонукають зарубіжні країни розвивати свої спроможності (спроможності мережі місії) у середовищі FMN, що в свою чергу вимагає врахування в них принципів таксономії СЗ, мережових стандартів та спроможностей Network-Enabled Capability (NEC), а також орієнтацію на інформацію. Це дозволяє командирів планувати та контролювати проведення спільних операцій та навчання за допомогою об'єднання мережевої та інформаційної інфраструктури.

Реалізація Програми FMN надає можливість зарубіжним країнам забезпечити розвиток початкових спроможностей, таких як: електронна пошта, спільна робота з документами, обмін текстовими, аудіо та відео повідомленнями, базові комунікаційні сервіси тощо, з поступовим нарощуванням спроможностей до більш складних сервісів із використанням хмарних обчислень.

Побудова систем симуляції та моделювання дій (simulation and modelling systems) відіграють важливу роль у наданні прогнозованої інформації для прийняття як ефективних управлінських рішень, так і більш ефективного управління ресурсами та повсякденної діяльності.

Велику увагу приділяють контейнерним сервісам, які забезпечують основу для розгортання хмар типу PaaS (Platform-as-a-Service) із наданням обчислювальних ресурсів та набору встановленого програмного забезпечення. Така контейнеризація передбачає безперервну інтеграцію та безперервну доставку програмного забезпечення із середовища розробки в операційне середовище, що значно полегшує оновлення програмних продуктів із середовища розробки в операційне середовище.

Сервіси зберігання передбачають доступ до спільних ресурсів інфраструктури для збереження великих масивів даних різної складності. Ці сервіси створюють умови для забезпечення логічного доступу до

сховища, абстрагуючись від топології фізичного зберігання даних. Реалізація вказаних сервісів зберігання даних при побудові інформаційної інфраструктури дозволяє передбачити розгортання хмар типу SaaS (Software-as-a-Service).

Розглядаючи використання датацентричного підходу можна виділити наступні переваги:

1) прогнозоване скорочення витрат на створення, впровадження та супроводження (модернізацію) програмного забезпечення (інформаційних систем);

2) унікальність даних та уникнення їх дублювання між елементами інформаційної інфраструктури; об'єднуюча структура даних, що включає всі можливі уявлення про область застосування;

3) підвищення якості надання інформації за рахунок логічної обробки даних та надання рекомендацій за рахунок попередньо-проведеного аналізу (вибірки даних).

До проблемних аспектів використання датацентричного підходу можна віднести:

1) розмите уявлення (не чітке розуміння) щодо загальної стратегії даних та корпоративної політики зберігання та обробки даних;

2) неналежні розрахунки щодо обчислювальних ресурсів та великого об'єму пам'яті для зберігання та динамічної обробки великих масивів даних;

3) недооцінка інтеграційних процесів, завдяки чому дані продовжують накопичуватися замість їх ефективного використання в режимі, наближеному до реального часу.

Сучасним прикладом, в основі якого є використання датацентричного підходу, є концепція проведення спільних операцій за допомогою фреймворка Об'єднаного командування і контролю (Joint All-Do main Command And Control), яку США представило у 2021 році. Такий фреймворк об'єднує платформи (сенсори) із різних доменів проведення операцій (таких як: повітря, земля, вода, космос та кіберпростір) для швидкого та спільного доступу до даних. Стрімке зростання даних, які генерують сенсори, потребують підвищення потужностей обробки даних та каналів передачі даних. Фреймворк JADC2 розділяє сховище та обробку даних. Разом з тим, сховище даних спроектовано так, щоб не допускати редагування вихідних даних, доступних для повторного використання через API [5].

Виходячи з проведеного аналізу можна стверджувати, що застосування датацентричного підходу з використанням хмарних сервісів дасть змогу підвищити якість використання (надання) інформації та заощадити витратні кошти на побудову інформаційної інфраструктури.

Перспектива росту кількості даних та подальша обробка великих масивів даних, їх динамічне оновлення (зміна) свідчить про актуальність використання датацентричного підходу при побудові інформаційної інфраструктури.

### **Список використаних джерел:**

1. Андрощук О. В., Голобородько М. Ю., Головченко О. В., Миронюк А. Б. Теоретичні особливості використання центрів обробки даних в приватних хмарах: вимоги побудови, види, переваги та недоліки, надійність. *Молодий вчений*. 2021. С. 1–4.
2. P. Adams, R. Djoko, S. Miller. A Datalake Solution for Time Series Data. URL: [https://sjmiller8182.github.io/Warehousing-Stock-Tweet-Data/reports/Project\\_Paper.pdf](https://sjmiller8182.github.io/Warehousing-Stock-Tweet-Data/reports/Project_Paper.pdf)
3. Taylor D. Difference Between Fact Table and Dimension Table / Guru99. 2022. URL: <https://www.guru99.com/fact-table-vs-dimension-table.html>
4. Pan D. Metadata Version Management for DW 2.0 Environment. *Journal of Convergence Information Technology*. 2010. P. 54–56.
5. Principles // The Data-Centric Manifesto. URL: <http://datacentricmanifesto.org/principles/>