

ІНФРАСТРУКТУРА СХОВИЩА ДАНИХ ДЛЯ МОДЕЛЮВАННЯ ЕПІДЕМІЧНИХ ПРОЦЕСІВ

*Пирогов Павло Андрійович\**, студент групи 345а

*Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ»*

Хмарна система зберігання даних, або зберігання даних як послуга – це абстрактне поняття, яке відповідає системі зберігання даних, яку можна адмініструвати за вимогою через спеціальний інтерфейс. Цей інтерфейс абстрагує і місцезнаходження системи. Хмарні інфраструктури зберігання даних утворюють нові архітектури, які підтримують різні рівні обслуговування поверх потенційно великої групи користувачів і географічно розподілених накопичувачів.

Загальна архітектура дає змогу виділити деякі характеристики сучасної хмарної архітектури зберігання даних. Розглянемо детальніше кожен з цих характеристик з аналізом її впливу на архітектуру сховища даних.

Керованість. Хмарова система зберігання має бути самокерованою. Повинна мати можливість додавати нові накопичувачі та знаходити і автоматично виправляти помилки.

Метод доступу. Загальноприйнятими є API Web-сервіс реалізованих на принципах REST, що оснований на об'єктно-орієнтованій схемі, розробленій поверх протоколу HTTP. Проблема API Web-сервісів полягає в тому, щоб скористатися хмаровою системою зберігання потребує інтеграції з додатком.

Продуктивність. Можна аналізувати з багатьох аспектів, але основне – це переміщення даних. Проте існує проблема в транспортному протоколі TCP. TCP керує потоком даних на основі підтвердження прийому пакетів з віддаленого вузла. Втрата або затримка пакетів дозволяють керувати перевантаженням, що ще більше обмежує продуктивність для уникнення глобальних мережових проблем. TCP підходить для переміщення невеликих даних, але для доставки великих обсягів даних – час обміну даними збільшується.

Кількість користувачів. Це стосується різних рівнів хмарової системи зберігання, від рівня додатка, до рівня зберігання. Використання багатьма користувачами поширюється на мережову інфраструктуру, яка з'єднує користувачів з накопичувачами, забезпечуючи гарантовану якість обслуговування і виділену смугу пропускання для конкретного користувача.

Масштабованість. Можливість нарощувати ресурси зберігання це поліпшує економічну ефективність для користувача. Масштабованість повинна забезпечуватися не тільки для самої системи зберігання, але й для її пропускну здатності.

Управління. Здатність контролювати і управляти тим, як зберігаються дані, і пов'язаними з цим витратами. Постачальники хмарових послуг пропонують засоби управління, які забезпечують користувачам підвищений контроль над витратами.

Ефективності зберігання. Це важлива характеристика інфраструктури зберігання, враховуючи акцент на економію. Щоб зробити систему зберігання ефективнішою, потрібно зберігати більше даних. Рішенням є скорочення обсягу даних. Є два способи: стиснення і вилучення всіх дублікатів даних.

Вартість. Одна з найпомітніших особливостей хмарового зберігання даних. Це економія на придбанні накопичувачів, їх енергопостачання, ремонт, а також на управління зберіганням. Якщо розглядати хмарове зберігання з цього погляду, воно може виявитися вигідним за певних умов.

Таблиця 1 – Характеристики хмарної архітектури зберігання даних

Характеристика	Опис	Було обрано для реалізації
Керованість	Здатність керувати системою за наявності мінімальних ресурсів	Самокерованість
Метод доступу	Протокол, через який надаються послуги хмарового зберігання даних	Web-сервіс реалізованих на принципах REST
Продуктивність	Вимірюється пропускну здатністю і часом затримки	TCP, тому що дані не великі
Кількість користувачів	Підтримка багатьох користувачів	Велика кількість користувачів
Масштабованість	Можливість поступового нарощування для задоволення нових вимог або обробки підвищеного навантаження	Масштабованість системи та пропускну здатності
Управління	Можливість управляти системою – зокрема, вибираючи вартість, продуктивність або інші характеристики	Статистика по використанню
Ефективності зберігання	Міра ефективності використання накопичувачів	Стиснення і вилучення всіх дублікатів даних.
Вартість	Міра вартості зберігання даних	Ціна за гігабайт.

\* Виконано в рамках проєкту Національного фонду досліджень України 2020.02/0404 «Розробка інтелектуальних технологій оцінки епідемічної ситуації для підтримки прийняття управлінських рішень у сфері біобезпеки населення».