

УДК 004.658

Мухін В.Є., д.т.н.; Корнага Я.І., к.т.н.

## МОДЕЛЬ РЕАЛІЗАЦІЇ УПРАВЛІННЯ РЕСУРСАМИ ГЕТЕРОГЕННИХ БАЗ ДАНИХ В ХМАРНОМУ СЕРЕДОВИЩІ

**Mukhin V.Ye, Kornaha Ya.I. The model for the control of the resources of the heterogeneous databases in the cloud environment.** Due to the permanent growth in the volume of processed data in the databases these data should be stored in the heterogeneous cloud computing environment. This environment requires the special control model for the large data resources control as well as to support the data mining, storing and updating, the load balance, the security mechanisms and monitoring. In the paper is described the main problems for the control of the heterogeneous distributed database in the cloud environment. There is suggested the joint control model for the heterogeneous distributed databases allows support a database data query in the cloud environment. The basic components of this model are described as well as the basic levels of the model implementation in the cluster system. The experiment researches have shown that the suggested model provides a high-performance query processing for the heterogeneous distributed databases, and the security control module enables perform the events monitoring in the database in the dynamic mode.

**Keywords:** heterogeneous distributed database; cloud environments, dynamic control

**Мухін В.Є., Корнага Я.І. Модель реалізації управління ресурсами гетерогенних баз даних в хмарному середовищі.** Постійне зростання обсягу оброблюваних даних в базах даних викликає необхідність їх перенесення в гетерогенну систему зберігання даних на основі середовища хмарних обчислень. Для даної системи потрібна спеціальна модель управління ресурсами для забезпечення управління великими ресурсами даних, а також підтримки пошуку, зберігання та оновлення даних, балансу навантаження, управління безпекою системи та її моніторингу. Запропоновано нову модель підтримки системи управління гетерогенними базами даних в хмарному середовищі.

**Ключові слова:** гетерогенні розподілені бази даних, хмарне середовище, динамічне управління

**Мухин В.Е., Корнага Я.И. Модель реализации управления ресурсами гетерогенных баз данных в облачной среде.** Постоянный рост объема обрабатываемых данных в базах данных вызывает необходимость их переноса в гетерогенную систему хранения данных на основе среды облачных вычислений. Для данной системы требуется специальная модель управления ресурсами для обеспечения управления большими ресурсами данных, а также поддержки поиска, хранения и обновления данных, баланса нагрузки, управления безопасностью системы и ее мониторинга. Предложена новая модель поддержки системы управления гетерогенными базами данных в облачной среде.

**Ключевые слова:** гетерогенные распределенные базы данных, облачная среда, динамическое управление

### Вступ

В хмарному середовищі велика кількість бізнес-даних зберігаються в різних базах даних. Організація їх розташування повинна бути прозора для користувачів та гетерогенні бази даних розподілені в різних хмарних вузлах [1-3]. Для ефективного управління різнорідними даними, утворюється великий пул баз даних, який забезпечує ефективне використання послуг передачі даних в хмарному середовищі та реалізує інтеграцію даних і операцій з різних СУБД [1-3].

Існує велика різноманітність гетерогенних баз даних, які можуть бути розміщені в хмарному середовищі, вони включають в себе, як реляційні бази даних так і не реляційні бази даних (NoSQL), а репозиторії дозволяють зберігати різні типи файлів [2-4]. Для покращення послуг передачі даних гетерогенні бази даних повинні бути поєднані динамічно, тому дослідження управління гетерогенними базами даних в області хмарних обчислень потребують постановки нових вимог [2-4].

Динамічне управління гетерогенними розподіленими базами даних включає в себе опис, організацію, відкриття та узгодження, планування, моніторинг та ін. У хмарному середовищі повинен бути механізм, який дозволить ефективно управляти даними відповідно до їх

розміщення в СУБД, тому необхідно розробити політику ресурсу контролю та управління гетерогенними розподіленими базами даних, яке забезпечить ефективну роботу [2-5].

В цій статті представлена модель управління гетерогенними розподіленими базами даних в хмарному середовищі, яка пропонує управління розміщенням СУБД, а також пошуком та зберіганням даних.

### 1. Проблеми управління ресурсами гетерогенних розподілених баз даних в хмарному середовищі

Для ефективного управління гетерогенними розподіленими базами даних в хмарному середовищі, бази розбиваються для досягнення масштабованості і відмово стійкості. Традиційні реляційні СУБД не можуть задовольнити високу масштабованість, доступність та продуктивність, тому все частіше використовують не реляційні бази даних, так звані бази даних NoSQL [2-6]. Не реляційні бази даних мають перевагу в підтримці розподіленого прозорого зберігання даних, яке може бути виконане на дешевшому обладнанні та дозволяє вільно зберігати пару ключ-значення [3-6]. Багато компаній розпочали розробку NoSQL баз даних для хмарних технологій як для власного використання так і для розповсюдження. Основні не реляційні СУБД показані в таблиці 1.

Табл. 1. СУБД NoSQL

Компанії	СУБД
Google	BigTable
Microsoft	Microsoft SQL Azure
Oracle	Oracle Cloud
Open source	Hbase

Гетерогенні розподілені не реляційні бази даних не підтримують управління транзакціями та більшість з них засновані за рахунок проектів з відкритим кодом, тому вони не досконалі [4-7]. З розвитком хмарних технологій розвивається і методи обробки даних в не реляційних базах даних.

В даний час дослідження і застосування гетерогенної управління ресурсами даних в хмарних обчислень повинні вирішити наступні задачі [2-7]:

1. Опис неоднорідного ресурсу даних в хмарному середовищі (дані ізольовані і розкидані). В гетерогенних розподілених базах даних не вистачає семантичної інформації, важко визначити ресурс даних і вилучення інформації.

2. Відкриття та узгодження гетерогенних ресурсів. Немає технології відкриття, яка була б адаптована до динамічних змін в хмарному середовищі.

3. Динамічне управління гетерогенними розподіленими базами даних.

4. Моніторинг в режимі реального часу гетерогенних розподілених баз даних. Немає єдиного механізму, який своєчасно виявляє і діагностує ризики, які можуть існувати в хмарному середовищі.

5. Доступ до даних гетерогенних розподілених баз. Бази даних, які розміщені в хмарному середовищі, відрізняються від традиційних баз даних методами підключення, структурою розміщення даних та пошуковими методами.

6. Спільне управління гетерогенними розподіленими базами даних в хмарному середовищі. Кожна СУБД може містити часткову інформацію, тому потрібно враховувати не окремі елементи, а всю структуру в цілому. Потрібно розглядати питання про перетворення різнорідних даних в єдиний формат та спільного управління ресурсами різнорідних СУБД.

7. Безпека даних в гетерогенних розподілених баз даних хмарного середовища. При інтеграції даних повинні бути надійні механізми забезпечення безпеки функціонування всієї системи.

## 2. Модель реалізації управління гетерогенними розподіленими базами даних в хмарному середовищі

Для управління гетерогенними розподіленими базами даних розроблена модель реалізації спільного управління для підтримки обробки запитів до баз даних в хмарному середовищі. Архітектура моделі (рис. 1) включає в себе:

- рівень підготовки запитів користувачів;
- рівень управління розподіленими гетерогенними БД;
- рівень мережевого ресурсу БД;
- фізичний рівень розміщення БД;
- рівень управління безпекою.

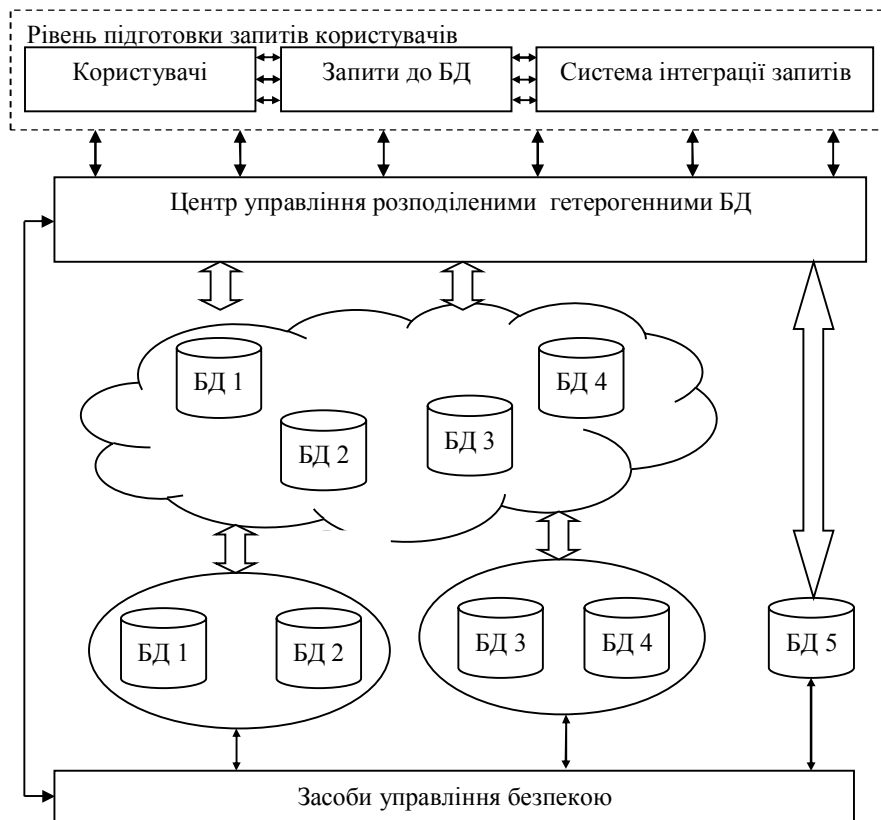


Рис. 1. Архітектура моделі управління гетерогенними розподіленими базами даних в хмарному середовищі

Модель реалізації спільного управління містить наступні компоненти:

1. Фізичний рівень зберігання БД: зберігає всі різномірні БД у різних хмарних вузлах, що дозволяє створювати складну структуру зберігання, у тому числі для класичних реляційних баз даних (БД1 та БД2), баз даних NoSQL (БД3 та БД4) і інші БД (інші БД).

2. Мережевий ресурс БД: окремі хмарні вузли об'єднуються в складну гетерогенну структуру. В хмарному середовищі мережевий рівень має абстрактну структуру, де фізичні вузли об'єднанні в логічні для формування спільної системи управління обробки запитів, що дозволить покращити управління гетерогенними БД.

3. Управління хмарними БД: для уніфікації формату доступу до БД, центр управління форматує дані з БД у відповідний формат для усунення семантичної неоднозначності. Центр управління БД підтримує спільне управління відображених даних, їх перетворення та злиття і т.д.

4. Рівень управління безпекою: відповідає за безпеку загальної архітектури розподілених баз даних, він включає в себе моніторинг баз даних і хмарних ресурсів. Моніторингу БД забезпечує безпечний доступ до даних і перевірку користувачів, що формують запити.

5. Рівень підготовки запитів користувачів: інтегрує запити з різних БД хмарного середовища відповідно до вимог прикладного програмного забезпечення користувача та забезпечує обмін запитами користувачів з БД.

### 3. Реалізація моделі в кластері

*Фізичний рівень зберігання.*

Хмарне середовище містить фізичну пам'ять в кластері даних та об'єднання кластерів і звичайні сервери, тому гетерогенні розподілені бази даних можуть бути розміщені в будь-якій частині. Для інтеграції СУБД в розподілені хмарні масиви використовується класична файлова систему HDFS.

*Мережевий ресурс БД.*

В хмарному середовищі ресурси формуються у вузли, які складають єдину мережу, аналогічно формується і система розміщення гетерогенних розподілених баз даних. Для отримання доступу до певного вузла БД застосовуються спеціальні алгоритми, які виконують динамічну підтримку мережеских джерел даних. Процедури обробки запитів в БД повинні функціонувати спільно з процедурами обробки інформації в хмарному середовищі для запобігання підвищення навантаження.

*Управління безпекою БД.*

Безпека обробки запитів є важливим питанням в гетерогенних розподілених базах даних, які розміщені в хмарному середовищі. Дослідження показують, що система безпеки хмарних технологій в певних випадках може завадити виконанню міграції корпоративних баз даних в кластер. Модуль управління безпекою гетерогенних розподілених баз даних в хмарному середовищі взаємодіє з процедурами управління даними і завдяки активному моніторингу всіх БД ресурси захищені від зовнішніх атак.

Рис. 2 показує, що модуль моніторингу окремого вузла кластера виконує моніторинг ресурсу вузла і накопичує інформацію про його стан, а модуль моніторингу самого кластеру збирає події моніторингу з різних вузлів. Отримані дані передаються в базу даних, яка використовується адміністратором безпеки для контролю критичних подій та станів гетерогенних розподілених баз даних.

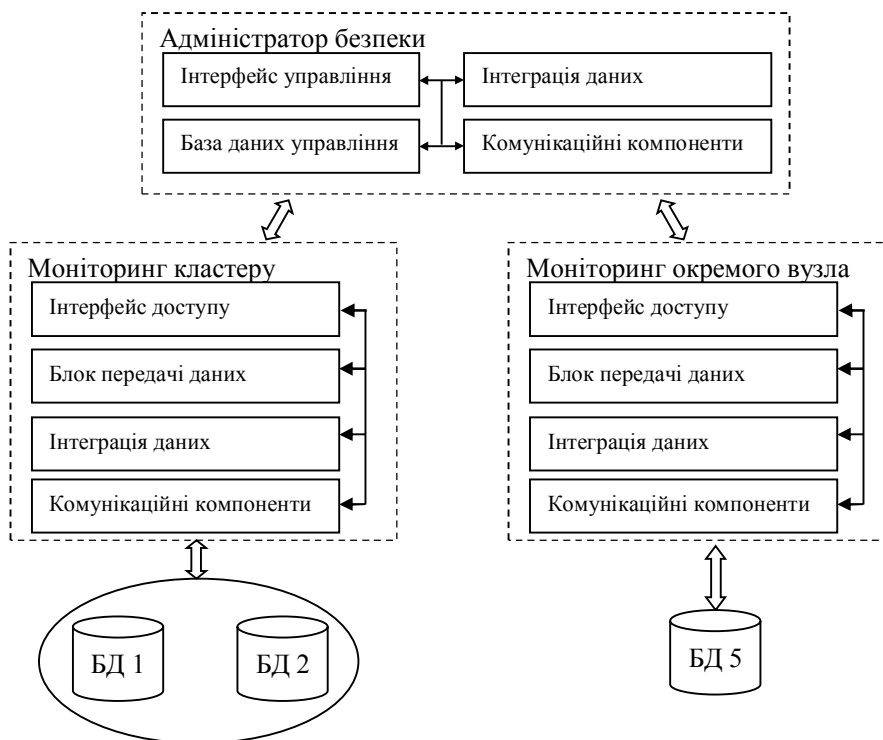


Рис. 2. Схема моніторингу гетерогенних розподілених баз даних хмарного середовища

Для перевірки моделі реалізації спільного управління гетерогенними розподіленими базами даних проведено експериментальні дослідження з застосуванням реляційних БД та NoSQL БД, при цьому хмарне середовище базувалося на класичних кластерних платформах. При застосуванні двох розподілених баз даних швидкість обробки запитів виявилася меншою ніж при застосуванні восьми вузлів баз даних. Слід зазначити, що окремо порівнювались реляційні розподілені БД та NoSQL БД.

Експеримент (рис. 3) довів, що при побудові розподілених баз даних в хмарному середовищі швидкість обробки запитів залежить від кількості вузлів: із збільшенням кількості вузлів збільшується кількість оброблених запитів за певний період часу, наприклад, за 1 хвилину. Також важливим є вибір типу бази даних, оскільки при роботі з реляційними базами даних швидкість обробки запитів виявляється меншою, ніж при роботі з NoSQL БД.

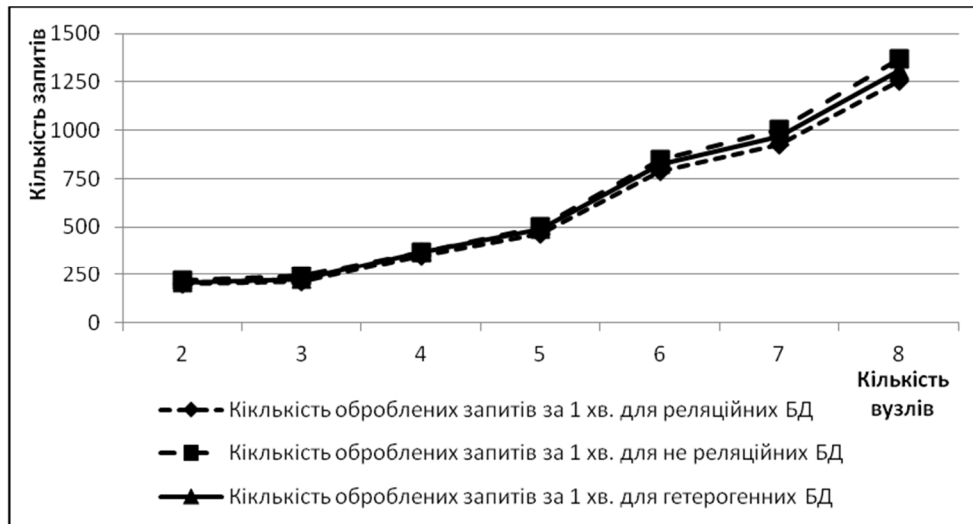


Рис. 3. Порівняння кількості оброблених запитів за 1 хв. для різної кількості вузлів

Для аналізу коректності роботи модулю управління безпекою проведено експеримент щодо визначення помилок I та II роду для різної кількості вузлів. Помилкою I роду вважається неправильне трактування загрози, а помилкою II роду - неправильне трактування звичайної події.

З результатів експерименту, який зображений на рис. 4 видно, що при збільшенні кількості вузлів помилки I та II роду збільшуються.

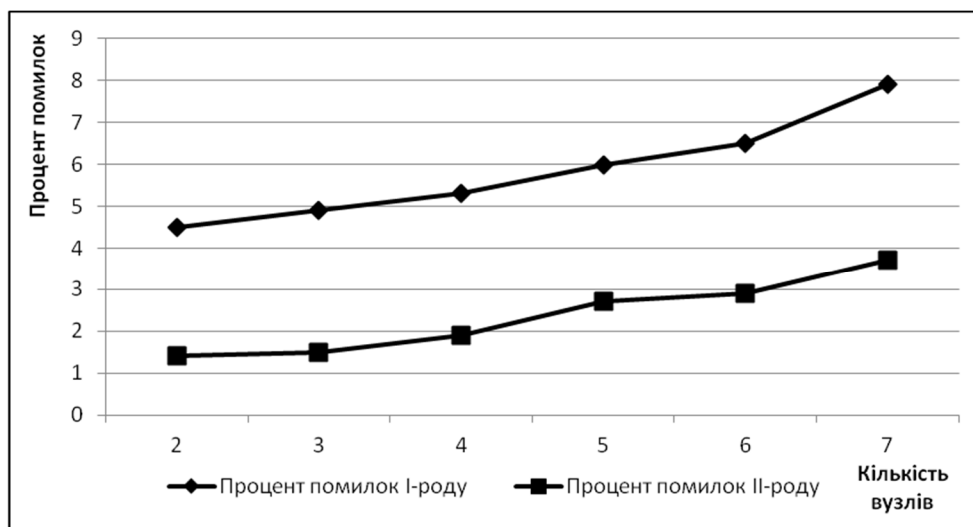


Рис. 4. Обробка подій моніторингу

## **Висновки**

Система управління гетерогенними розподіленими базами даних в хмарному середовищі є гнучкою, тому в останній час більшість баз даних підприємств переносяться на хмарні платформи. В статті запропоновано модель реалізації спільного управління ресурсами розподілених баз даних в кластері. Експериментальні дослідження показали, що запропонована модель має високу продуктивність обробки запитів для гетерогенних розподілених баз даних, при цьому модуль управління безпекою забезпечує проведення моніторингу подій в базі даних в динамічному режимі.

## **Література**

1. Sun T. Research on heterogeneous data management model in cloud environment / T. Sun, X. Wang // International journal of database theory and application. – 2013. – P. - 152.
2. Birman K.C. Toward a cloud computing research agenda / K.C. Birman, G. Renesse // ACM SIGACT News. – 2009. – PP. 68- 80.
3. Mell P. The NIST Definition of Cloud Computing / P. Mell, T. Grance // National Institute of Standards and Technology. – 2011. – 7 p.
4. Юдін О.К. Сучасні моделі корпоративних мереж на базі хмарних технологій / О.К. Юдін, Р.В. Зюбіна, Т.В. Зюбін // Інформатика. – 2014. – С. 86-91.
5. Яковицький І. Технологія хмарних обчислень як інструмент створення інформаційної інфраструктури управління / І. Яковицький // Комунальне господарство міст. – 2012. – №102 С. 320-327.
6. Коміссар Д.О. Хмарні технології безпеки / Д.О. Коміссар, Є.Ю. Луппол // Вісник східноукраїнського національного університету імені Володимира Даля. – 2013. – Ч. 1. – № 15(204). – С. 83-87.
7. Аулов І.Ф. Хмарні обчислення та аналіз питань інформаційної безпеки в хмарі / І.Ф. Аулов, І.Д. Горбенко // Прикладная радиоэлектроника. – 2013. – Т. 12. – № 2. – С. 194-201.

## *Автори статті*

**Мухін Вадим Євгенійович** – доктор технічних наук, доцент, професор кафедри обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут», Київ, Україна. Тел. +38 067 508 76 84. E-mail: v\_mukhin@mail.ru

**Корнага Ярослав Ігорович** – кандидат технічних наук, доцент, доцент кафедри технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут», Київ, Україна. Тел. +38 099 224 65 15. E-mail: slovyan\_k@ukr.net

## *Authors of the article*

**Mukhin Vadym Yevheniyovych** – doctor of science (technic), associate professor, professor of Department of Computer Engineering, National technical university of Ukraine “Kiev Polytechnic Institute”, Kyiv, Ukraine. Tel. +38 067 5087684. E-mail: v.mukhin@kpi.ua.

**Kornaha Yaroslav Ihorovych** – candidate of science (technic), associate professor, associate professor of Department of Computer Engineering, National technical university of Ukraine “Kiev Polytechnic Institute”, Kyiv, Ukraine. Tel. +38 099 224 65 15. E-mail: slovyan\_k@ukr.net

Дата надходження в редакцію: 24.11.2015 р.

Рецензент: д.т.н., проф. О.В. Барабаш