

Дослідження шляхів підвищення ефективності запитів до баз даних розташованих на декількох хостах

Дипломна робота на здобуття кваліфікації бакалавра
комп'ютерних наук

Поршнева Катерина Олегівна, КА-25

Керівник — к.т.н., старший викладач кафедри ММСА Мурга
Микола Олексійович

Вступ

Часто при розробці програмного забезпечення виникає необхідність мати доступ до даних, які не тільки розподілені мережею, а й можуть зберігатися у різних форматах та сховищах. В таких випадках, швидкий і постійний доступ до даних стає проблемою. Різноманітність даних не дозволяє робити запити до них через один універсальний спосіб, так як кожна СКБД вимагає свій власний формат запитів.

Для реалізації системи керування гетерогенними базами даних було обрано мультиагентний підхід. За допомогою створення агентів, які можуть приймати рішення щодо того звідки які дані потрібно отримати і як їх привести до єдиного формату, цю проблему можливо вирішити.

Структура дослідження

Об'єкт дослідження: гетерогенні дані в розподілених інформаційних системах.

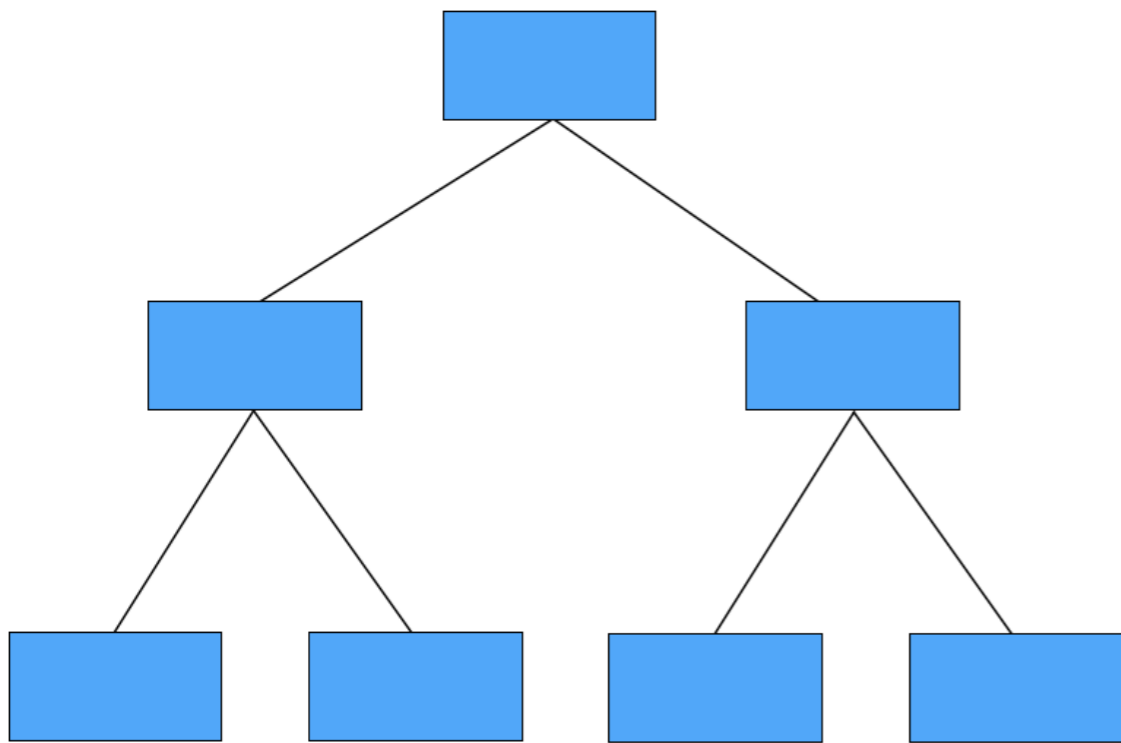
Предмет дослідження: моделі, методи та технології керування гетерогенними даними в розподілених інформаційних системах.

Мета дослідження: розробити та дослідити ефективність мультиагентної системи керування базами даних.

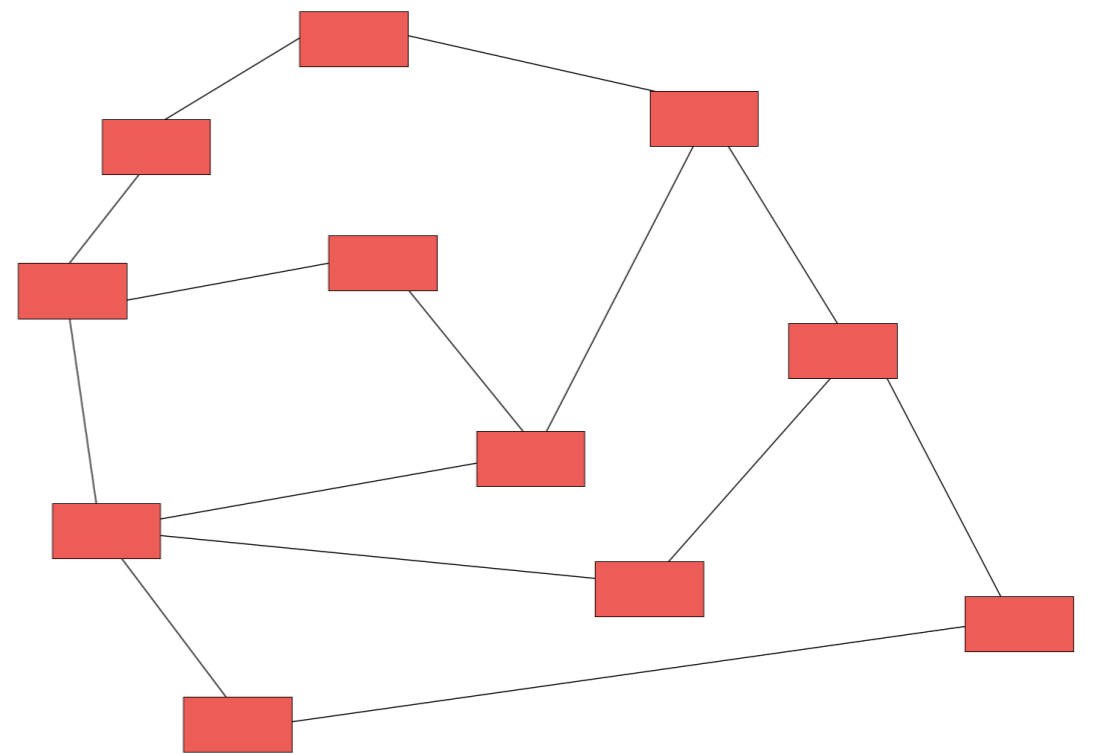
Завдання

1. Визначити основні вимоги до керування даними в розподілених інформаційних системах.
2. На підставі визначених вимог до керування даними в розподілених інформаційних системах, спроектувати та реалізувати платформу для створення і комунікації агентів, мультиагентну систему керування базами даних та інтерфейс адміністратора.
3. Дослідити властивості розробленої мультиагентної системи керування базами даних для простих запитів. Провести експериментальні дослідження впливу кількості баз даних, загальної кількості записів та типу запиту на час обробки системою.
4. Дослідити властивості розробленої мультиагентної системи керування базами даних для складних запитів. Порівняти ефективність системи для обробки складних та простих запитів.

Мультиагентна схема побудови ПЗ



Традиційна схема



Мультиагентна схема

Загальна схема розробленої системи

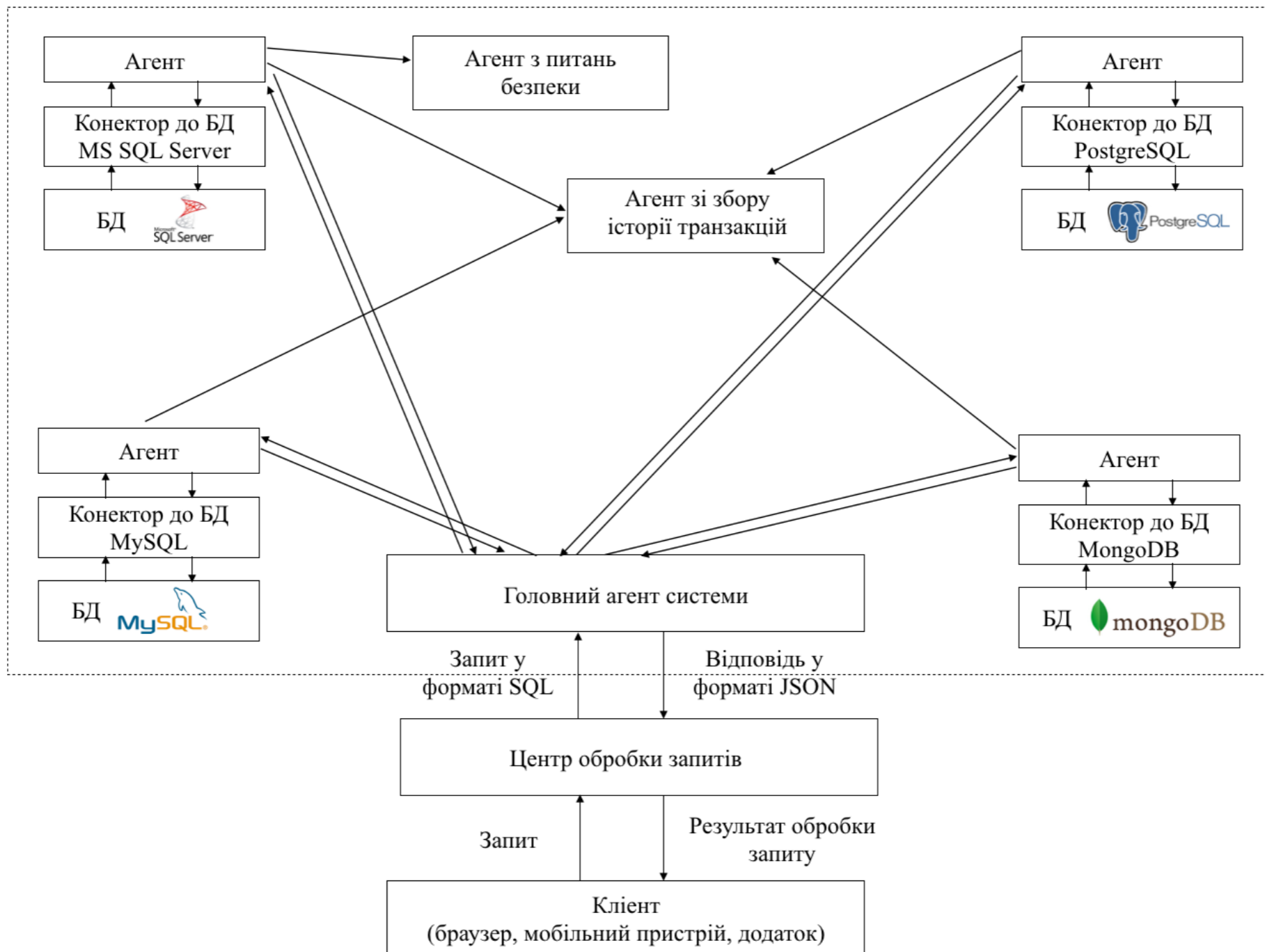
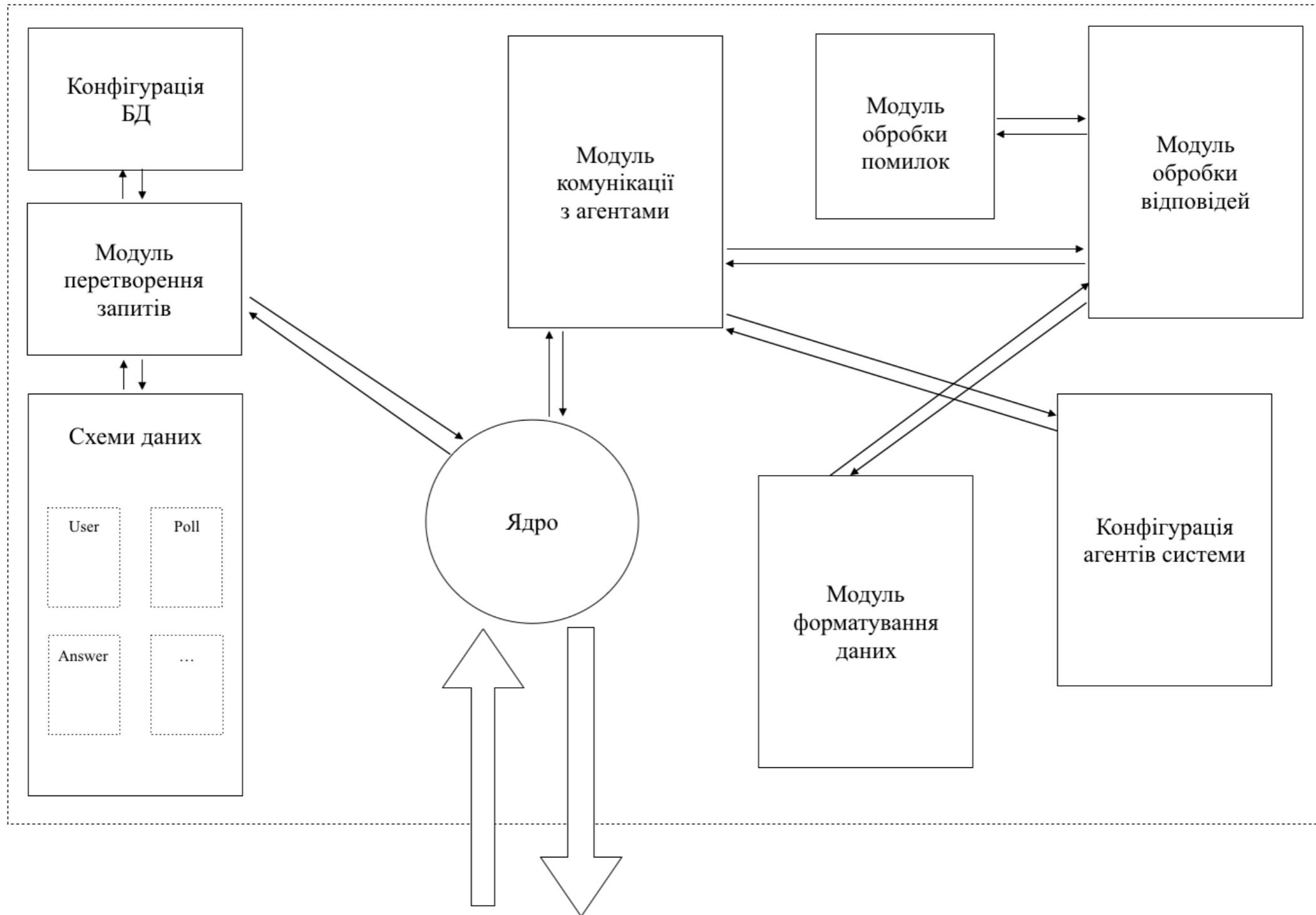
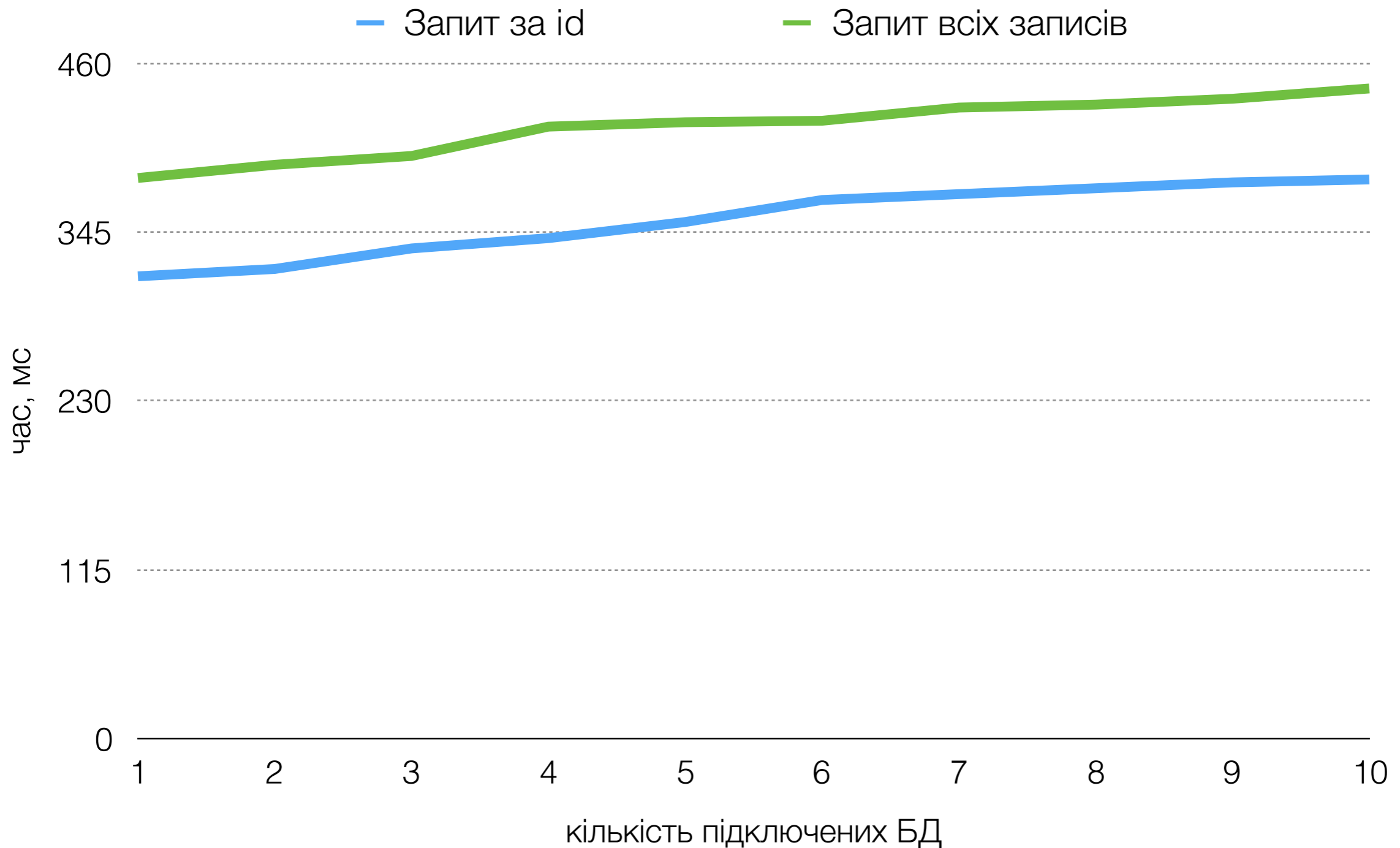


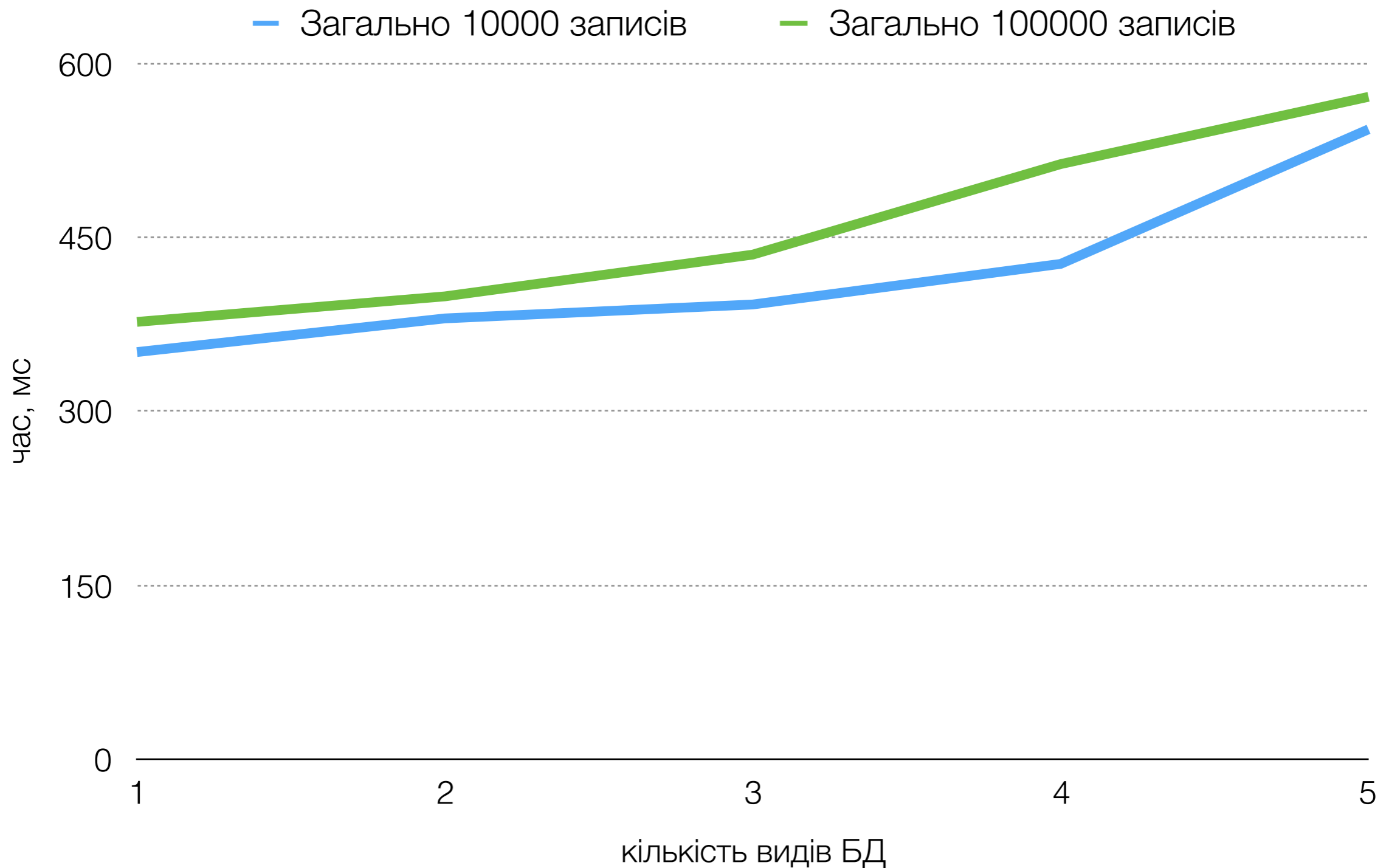
Схема роботи агента



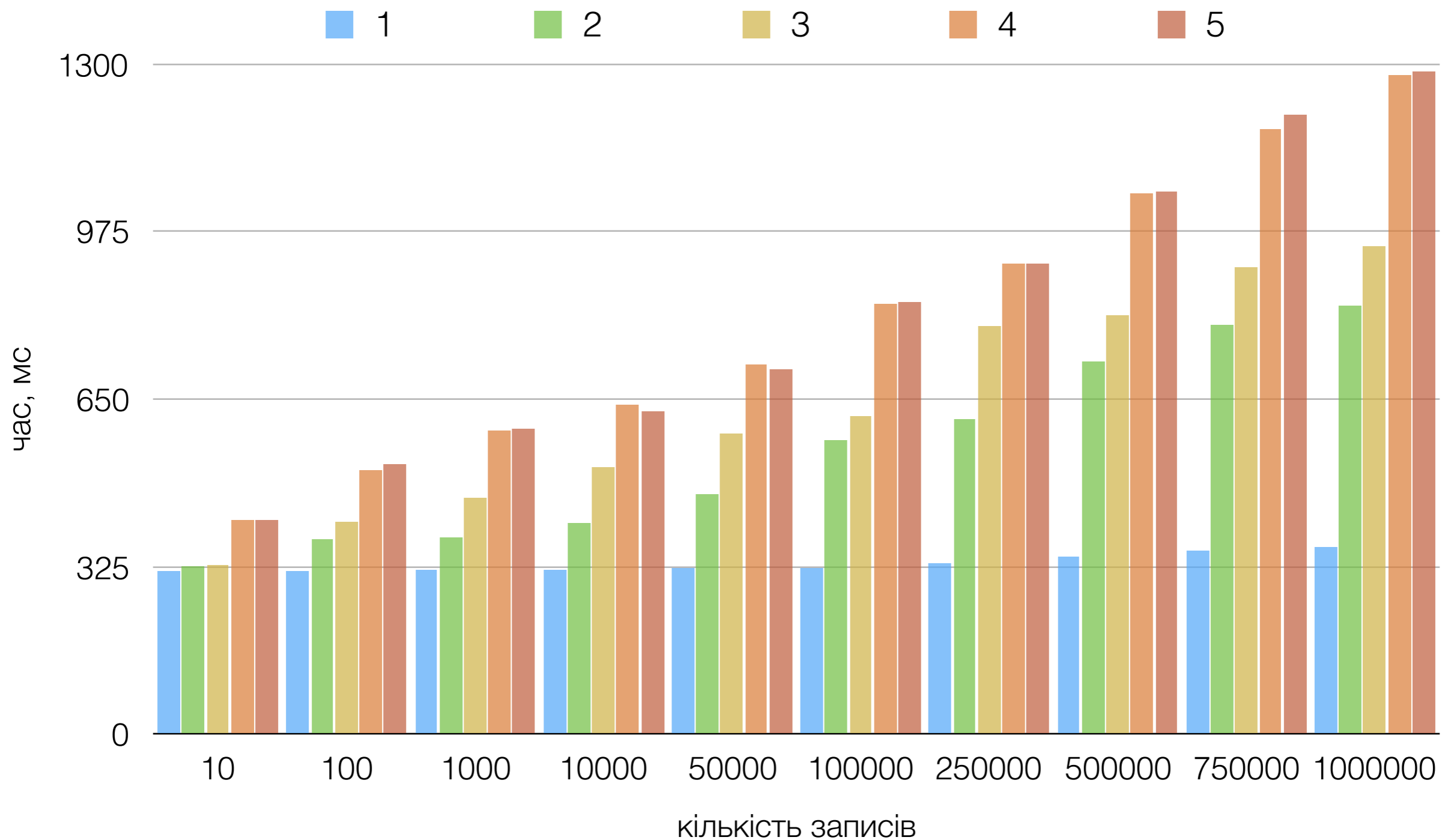
Залежність часу обробки запиту від кількості підключених БД



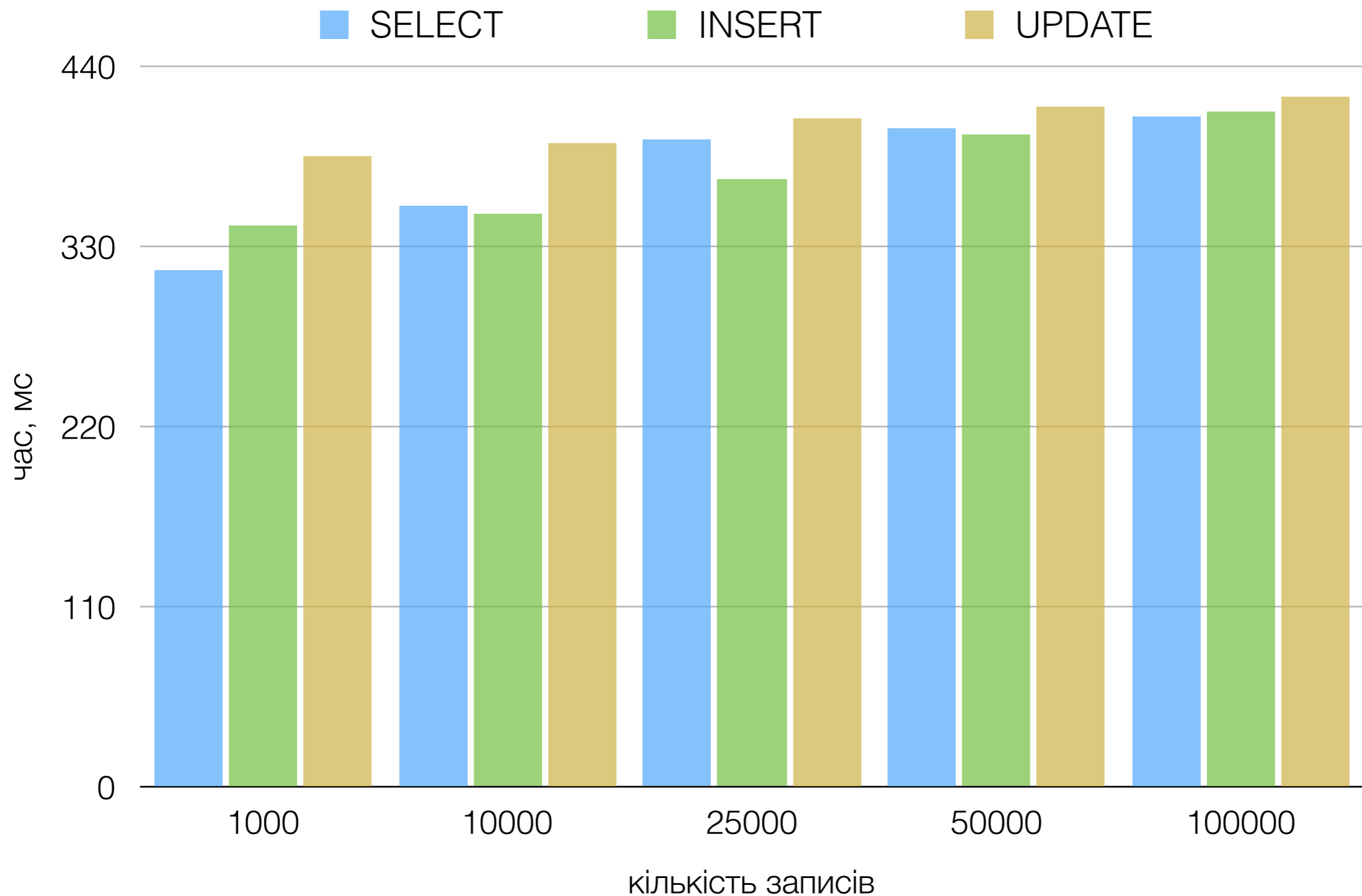
Залежність часу обробки запиту від кількості видів БД



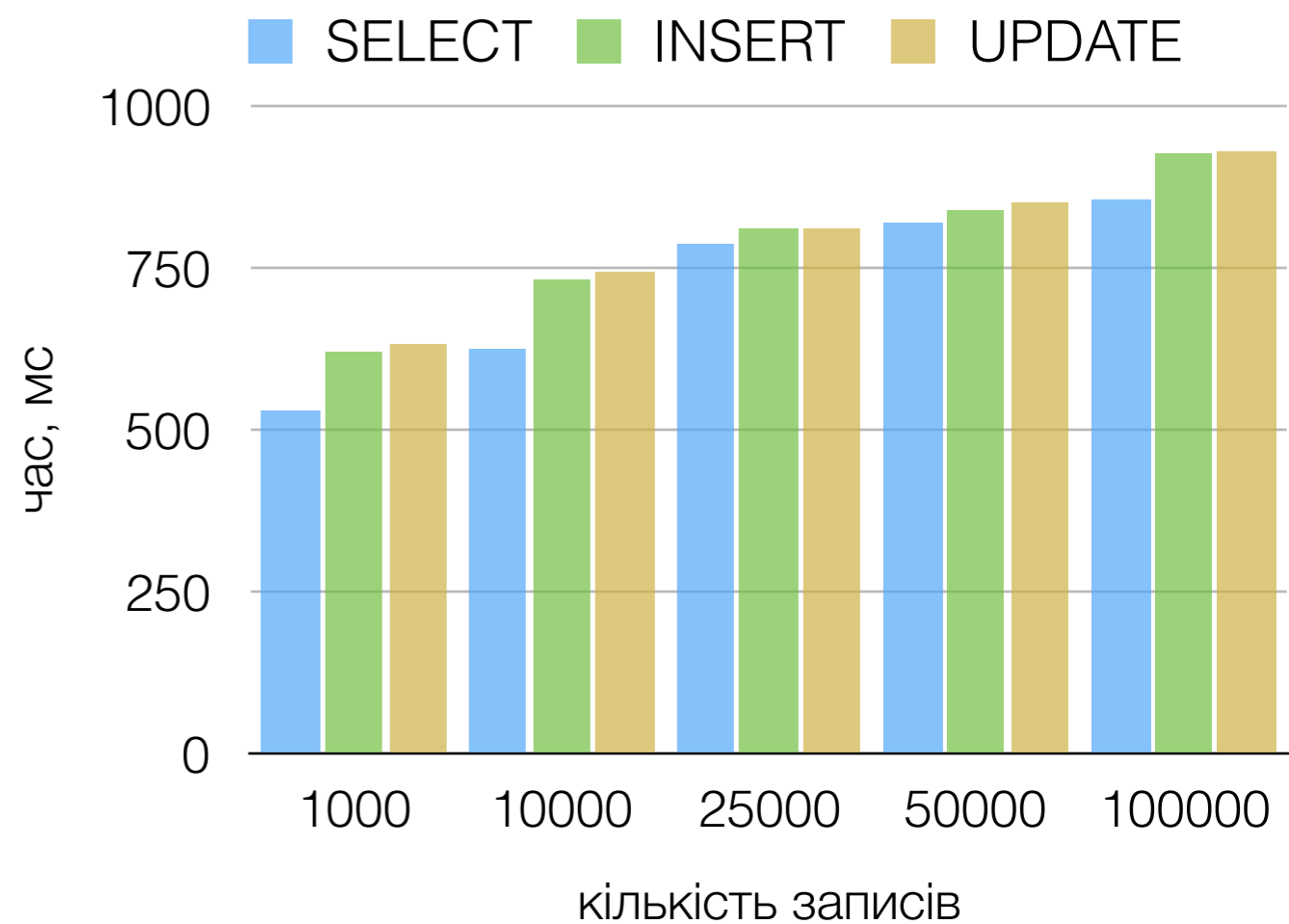
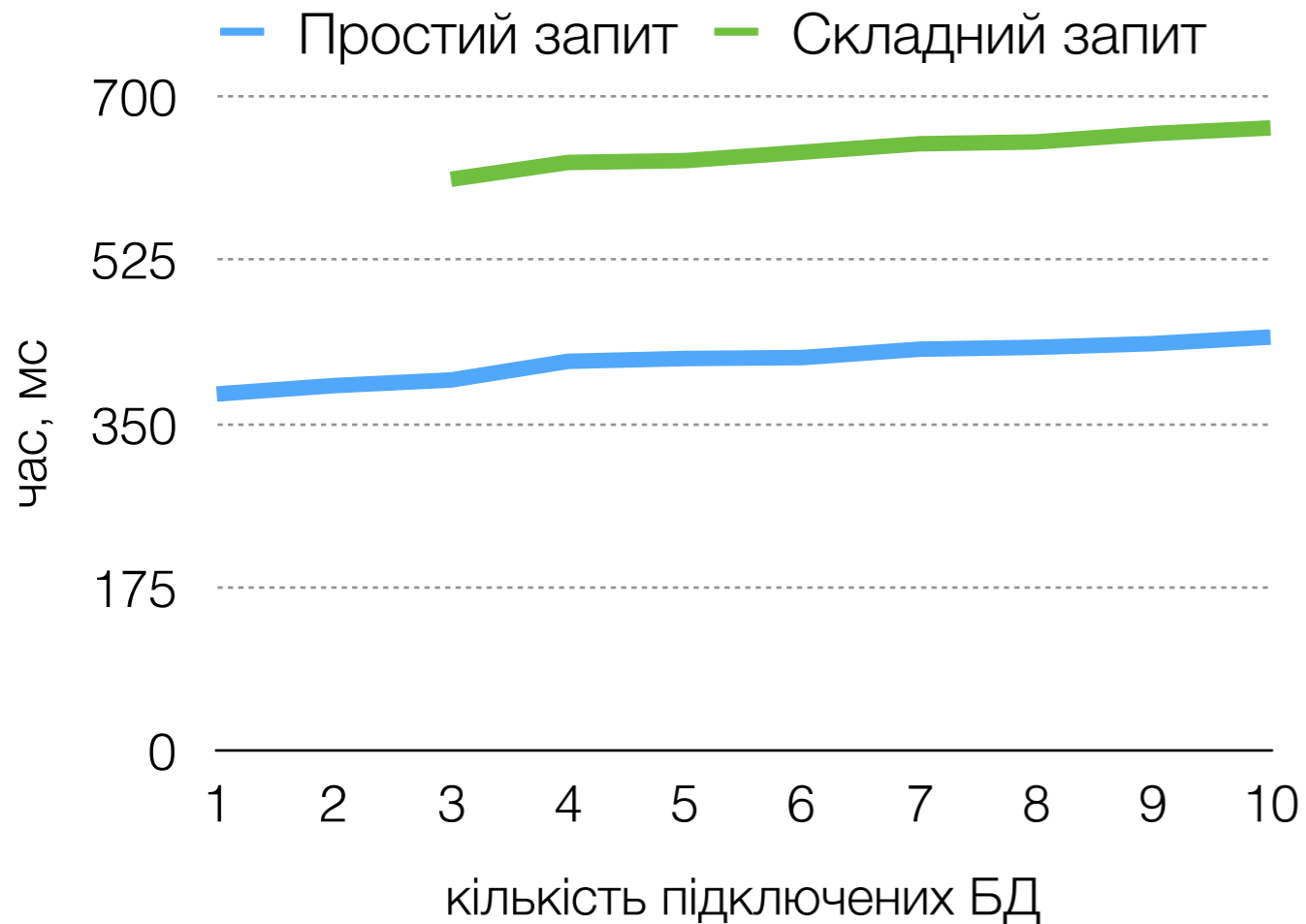
Залежність часу обробки запиту від загальної кількості записів



Залежність часу обробки запиту від типу запиту



Дослідження ефективності складних запитів



Висновки

- При збільшенні кількості підключених баз даних значення приросту часу рівне, в середньому, 2% з розрахунку на одну підключену базу даних, що можна пояснити необхідністю пересилання даних, та приведення їх до єдиного формату.
- Час обробки запитів залежить від кількості підключених видів баз даних, зріст часу на обробку запиту, насамперед, викликаний необхідністю приведення даних до єдиного формату та швидкістю обробки запитів конкретною БД, наприклад, при підключенні БД Oracle час обробки збільшився на 29%, а при підключенні інших БД в середньому збільшувався на 6%.

Висновки

- Експериментальне дослідження з варіюванням кількості записів від 100 до 1000000 показало середній приріст часу на обробку запитів 2.1% з розрахунку на додану 1000 записів в залежності від конфігурації системи та кількості підключених типів БД, приріст, в першу чергу, спровоковано часом обробки запиту самою базою даних та часом обробки та форматування запиту та результату системою при розподілених запитах по мережі.
- Найдовшими для обробки виявився запит типу UPDATE та INSERT, найшвидшим — SELECT, в середньому приріст складав 1.6% для всіх типів та складностей запитів.

Висновки

- Варіюванням кількості підключених вузлів та загальної кількості записів для обробки складних запитів у системі показало, в середньому, приріст відносно показників для простих запитів 50.3% для однакових конфігурацій системи, такий зріст, в першу чергу, пов'язаний з необхідністю агрегування даних тільки після того, як всі запити будуть завершені.
- Проведені експериментальні дослідження довели, що мультиагентний підхід для обробки запитів до баз даних має переваги своїми можливостями бути налаштованим під потреби конкретного користувача, зробити агентів інтелектуальними та збільшувати їх кількість без шкоди для швидкості обробки запитів.