

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕФЕКТИВНОСТІ ВИКОРИСТАННЯ РЕЛЯЦІЙНОЇ БД І MONGODB В ІНФОРМАЦІЙНІЙ СИСТЕМІ

Київський національний університет харчових технологій

Анотація

Проведено порівняльний аналіз реляційної бази даних з нереляційною і виділені основні переваги кожної з них, які дадуть нам можливість визначити, в яких випадках краще використовувати нереляційні бази даних, а в яких навпаки.

Ключові слова: бази даних, SQL, NoSQL, JSON, реляційні СКБД, нереляційні СКБД, MongoDB, SQL Server.

Abstract

A comparative analysis of relational databases with non-relational and highlights the main advantages of each, which will give us the opportunity to analyze and determine in which cases can be better to use non-relational database, in which vice versa.

Keywords: database, SQL, NoSQL, JSON, relational database, non-relational database, MongoDB, SQL Server.

Вступ

Протягом довгого часу, реляційні бази даних традиційно використовувалися для зберігання і пошуку структурованих даних. При правильному використанні, реляційні бази даних дають нам стабільність, ізоляцію і надійність. Але, в останні роки NoSQL-бази даних набувають все більшу популярність.

З еволюцією інтернету і мобільних пристроїв значно зріс обсяг даних, які необхідно зберігати і обробляти. В наш час, стає набагато складніше працювати з фіксованими структурами даних. Ще більше складнощів виникає з обробкою неструктурованих даних. Це говорить про те, що якщо вам потрібна практично необмежена масштабованість, вам потрібна нереляційна БД.

Метою роботи є вибір критеріїв та розробка і застосування алгоритмів порівняння реляційних та не реляційних баз даних для формування пропозицій, щодо ефективного використання кожного виду баз даних в інформаційних системах.

Результати дослідження

Дослідження та порівняльний аналіз проводились для СКБД MS SQL Server і MongoDB. Було проведено порівняльний і графічний аналіз реляційних і нереляційних баз даних. Порівняльний аналіз було здійснено за такими параметрами: функціонал, швидкодія, внесення та корегування даних, методи зберігання даних, обсяги пам'яті, що займає база даних.

Перевіримо швидкість роботи за наступним алгоритмом:

1. Спочатку проведемо масову вставку 1 млн. записів в одному потоці. Виміряємо, скільки часу займе ця операція.
2. Потім перевіримо роботу з базою даних з 10 потоків. В кожному потоці виконується наступна робота:
 1. 2% - Додавання нових записів
 2. 3% - Оновлення випадкового запису
 3. 1% - видалення випадкового запису
 4. 13% - вибірка випадкової сторінки з 20 записів

5. 81% - вибірка одного випадкового запису

Для порівняння швидкодії візьмемо SQL Server 2012 Enterprise.
Результати першої операції:

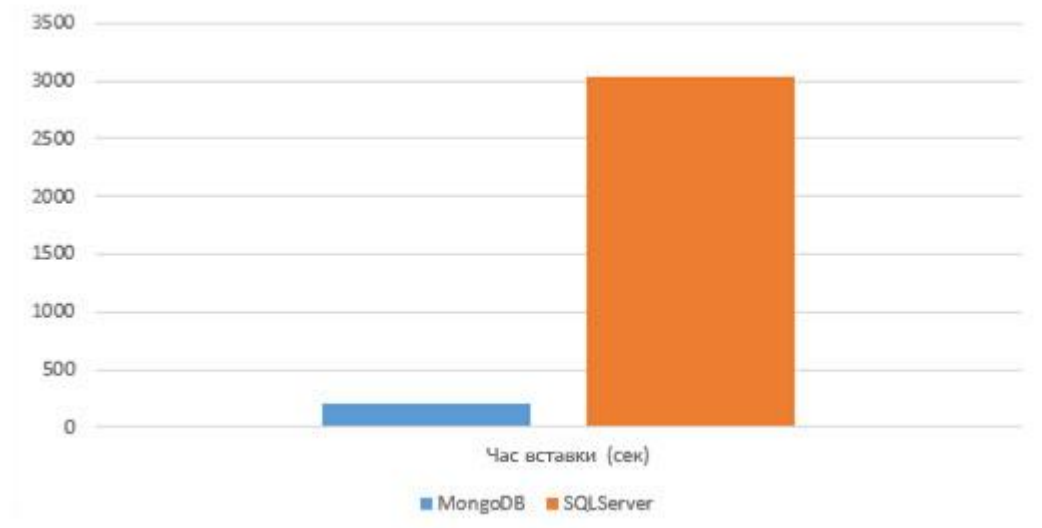


Рис.9. Час виконання 1 млн. записів в MongoDB & SQL Server

Вставка 1 млн. Записів на SQL Server - 3040 секунд (~ 50 хвилин), точно така ж процедура на MongoDB зайняла 214 секунд (~ 3.5 хвилини). **Різниця в 14 разів!** Це дуже сильний результат.

До речі, розмір бази даних: SQL Server - 96Мб, MongoDB - 384Мб. Різниця в 4 рази, причому не на користь MongoDB. Чим менше файл - тим менше операцій введення-виведення, тим простіше його кешування та резервування.

Звичайно, наведений тест некоректний. Вставляти 1 млн.записів в базу даних на SQL-сервер по одній - це абсолютно невірне рішення, необхідно використовувати оптимізацію. Масова вставка 1 млн. Записів за допомогою оператора Bulk Insert теж займає близько 3 хвилин, однак не завжди в програмі можна застосувати оператор масової вставки. Також можна вставляти дані «порціями» по 1000 записів у одній транзакції - це теж істотно прискорює вставку.

Проведемо другий тест, який набагато більше підходить до реальної роботи з типовим додатком: 3 вставки з різною кількістю елементів, 3 видалення по індексам, конкретним значенням і видалення усіх даних , 5 правок по всьому стовбцю, індексом, заданими параметрами, з перевіркою заданої умови і 15 різних вибірок - посторінкових і до одного запису, причому робота ведеться одночасно в декількох потоках.

Таблиця 1.1.

	<i>MongoDB</i>	<i>SQL Server</i>
1 потік	1 сек.	17 сек.
10 потоків	3 сек.	211 сек.

На простих операціях CRUD MongoDB демонструє дуже серйозний приріст продуктивності.

Однак більш глибоке вивчення аналітичних функцій - MAX / MIN, AVG, SUM, COUNT, DISTINCT, GROUP BY - показує високу трудомісткість в їх реалізації на MongoDB, особливо на великих базах.

Висновки

Реляційні або нереляційні бази даних - це не єдиний вибір, який розробникам систем належить зробити між різними видами баз даних. Доведеться ще вибрати і всередині них. Потрібно просто оцінити їх і чітко зрозуміти свої власні потреби.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Java development 2.0: Вторая волна разработки Java-приложений. MongoDB – хранилище данных NoSQL со всеми (полезными) признаками СУРБД.
2. Онлайн-руководство по MongoDB.
3. <http://www.json.org/json-ru.html> - Введение в JSON
4. <http://habrahabr.ru/post/119703/> - Возможности MongoDB
5. The Little MongoDB Book - Karl Seguin — разработчик с опытом во многих областях и технологиях.
6. <http://habrahabr.ru/post/217393/> - MongoDB от теории к практике. Руководство по установке кластера mongoDB
7. ОРГАНІЗАЦІЯ БАЗ ДАНИХ ТА ЗНАНЬ - О.М. М'ЯКШИЛО
8. http://itedu.org.ua/files/seminar_25_26_04/MongoDB__SoftTimeINFO.pdf - Язык запросов MongoDB.
9. <http://bugscatcher.net/archives/2956> - MONGODB. ФОКУС НА ТЕСТИРОВАНИЕ
10. <https://www.mongodb.org/> –офіційний сайт MongoDB
11. <http://squadette.ru/blog/2014/08/12/bazy-dannykh-i-nosql/> - Базы данных и NoSQL

Брацький Вадим Олегович – аспірант кафедри інформаційних управляючих систем, факультету автоматизації і комп'ютерних систем, Київський національний університет харчових технологій, м.Київ, e-mail: vadybratskyi@gmail.com

Науковий керівник: **М'якшило Олена Михайлівна** - канд. техн. наук, доцент кафедри інформаційних управляючих систем, Київський національний університет харчових технологій, м.Київ.

Bratskyi Vadim Olegovich. - postgraduate student of the Department of Information Control Systems, Faculty of Automation and Computer Systems, Kyiv National University of Food Technologies, Kiev, e-mail: vadybratskyi@gmail.com

Supervisor: **Myakshylo Elena Mikhailovna** - candidate. Sc. Associate Professor, Department of Information Control Systems, Kyiv National University of Food Technologies, Kyiv.