

УДК 004.6

Бахметьев В.В.¹, Киричек Г.Г.²

¹ студ. гр. КНТ-518, НУ «Запорізька Політехніка»

² канд. техн. наук, доц., НУ «Запорізька Політехніка»

ДОКУМЕНТО-ОРІЄНТОВАНІ СКБД MONGODB ТА COUCHDB

Потрібність у використанні СКБД (система керування базами даних) в більшості видів діяльності у ІТ-сфері не є секретом. Усе ще дуже популярними залишаються реляційні СКДБ на моделі SQL, але напрямок

нереляційних документо-орієнтованих представників СКДБ зайняв значну частину усього ІТ-простору, де використовуються СКБД загалом: напрями та задачі, що потребують інструментарій для зберігання ієрархічних структур даних (документів) та мають вирішені проблеми масштабованості і доступності за рахунок повної або часткової відмови від вимог атомарності та узгодженості даних. Вони знаходять своє застосування у системах управління вмістом, видавничій справі, документальному пошуку.

MongoDB є документо-орієнтованою системою управління базами даних, «заточена» під веб-додатки та інфраструктуру Інтернету, розроблена на мові C++. Модель даних та стратегія їх постійного зберігання спроектовані для досягнення високої пропускної спроможності читання та запису та забезпечує просту масштабованість з автоматичним переходом на резервний ресурс у разі відмови [1].

Працює з набором реплік, тобто може містити дві чи більше копії даних різних вузлах. Кожен екземпляр набору реплік може будь-якої миті виступати у ролі основної чи допоміжної репліки. Усі операції запису та читання за замовчуванням здійснюються з основною реплікою. Допоміжні репліки підтримують копію даних у актуальному стані. Уся система масштабується горизонтально, використовуючи техніку сегментування об'єктів баз даних - розподіл їх частин за різними вузлами кластера. Адміністратор вибирає ключ сегментування, який визначає, за яким критерієм дані будуть рознесені вузлами. Завдяки тому, що кожен вузол кластера може приймати запити забезпечується балансування навантаження (рис.1).

CouchDB є документо-орієнтованою системою управління базами даних з відкритим вихідним кодом, розроблена програмним фондом Apache на мові Erlang. Основна увага приділяється простоті використання. CouchDB не зберігає дані та зв'язки в таблицях. Натомість кожна база даних це набір незалежних документів. Кожен з яких містить власні дані та незалежну схему (рис.1).

Програма може отримати доступ до кількох баз даних, наприклад, що зберігається на мобільному телефоні користувача та сервері. Метадані документа містять інформацію про версію, дозволяючи об'єднувати дані та вирішувати будь-які суперечності, які могли виникнути в момент, коли бази даних були роз'єднані. Для управління конкурентним доступом використовується механізм REST API, завдяки чому можна уникнути необхідності блокування файлу бази даних під час запису. Вирішення конфліктів відноситься до сфери відповідальності логіки докладання, вирішення конфлікту зазвичай включає об'єднання даних в один документ, старий видаляється [3].

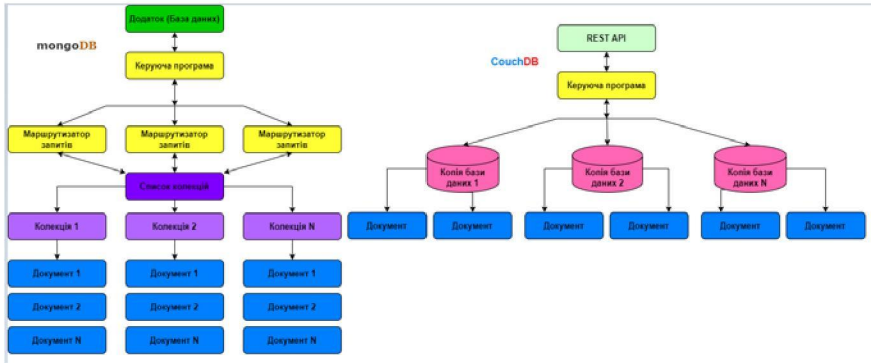


Рисунок 1 – Схема роботи MongoDB та CouchDB.

Обидві СКБД розроблялися для схожих задач, але для різних їх специфікацій, тому мають суттєві відмінності. Доречно розбити їх на елементи.

Ліцензування: CouchDB – обмеження на використання і модифікацію відсутні, публікація змін вихідного коду не потрібна; MongoDB – обов'язкова публікація змін вихідного коду, наявність ліцензії окрім обмеженого доступу.

Інтерфейс та модель даних: CouchDB – використовує інтерфейс на основі HTTP/REST з даними формату JSON, дуже інтуїтивно зрозумілий і добре розроблений; MongoDB - використовує бінарний протокол BSON і протокол користувача через TCP/IP, разом дозволяє швидше розшифровувати дані.

Структура БД: CouchDB – усі документи рівнозначні і між собою відрізняються тільки за полями ідентифікації та ревізії id - rev; MongoDB – документи діляться на колекції за типом та мають лише ідентифікаційне поле id.

Вирішення колізій: CouchDB – при оновленні документу потрібне визначення правильного значення rev; MongoDB – не має відповідного поля, але дозволяє відправляти лише оновлені поля, а не весь документ.

Індексація: CouchDB – проводить тільки при виклику запиту та оновлює весь документ, що займає багато часу; MongoDB – оновлює індекси при редагуванні [2].

Зберігання даних: CouchDB – дані зберігаються в одному файлі, їх додавання можливе тільки в кінець, старі ревізії зберігаються, а їх видалення не блокує БД, що гарантує цілісність файлу; MongoDB – дані зберігаються у вигляді багатьох memo-маркованих файлів – колекцій, які включають в себе документи, що потребує великої кількості ОЗУ, СКБД намагається оновлювати документ в місці його розташування.

Робота з файлами: CouchDB – дозволяє прикріпляти файли до документа і працювати з ними незалежно; MongoDB – реалізована потокова передача файлів великого розміру.

Реплікація: CouchDB – підтримує реплікацію типу master-master з функціями вирішення конфліктів; MongoDB – підтримує реплікацію master-master, але з обмеженою кількістю машин в множині реплікації.

Швидкість роботи: CouchDB – суттєво залежить від внутрішніх процесів і розміру документів; MongoDB – критично швидша за CouchDB через інший підхід до зберігання і роботи з документами.

Підтримка мобільних платформ: CouchDB – можна запускати на пристроях Apple iOS та Android, представлена підтримка мобільних пристроїв; MongoDB – мобільна підтримка не представлена.

Розмір БД: CouchDB – може збільшуватись, але краще, якщо розмір і структура визначені заздалегідь; MongoDB для швидкозростаючих баз даних.

Запити: CouchDB – запити використовують функції map-reduce; MongoDB – використовує Map/Reduce (JavaScript), створюючи колекцію + мову запитів на основі об'єктів [1, 3].

Паралелізм та послідовність: CouchDB – слідує за MVCC, загалом несуперечлива; MongoDB – оновлюється на місці, строго несуперечлива.

Безпека: CouchDB – автентифікація OAuth та базова HTTP, власні права доступу; MongoDB – має слабший захист та три групи прав доступу для адміністраторів та користувачів.

Висновки. Незважаючи на те, що обидві СКБД побудовані на нереляційній основі і є документо-орієнтованими, між ними більше відмінностей, ніж у багатьох СКБД, що побудовані на принципі SQL. Тобто, для їх використання потрібно точно розуміти завдання і можливість проблем майбутньої БД для правильного вибору. Якщо нам потрібна база даних, яка працює на мобільних пристроях, потрібна реплікація master-master або надійність одного сервера, то CouchDB. А якщо потрібна максимальна пропускна здатність і швидкість або у нас швидко зростаюча база даних, то MongoDB.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Редмонд Э., Уилсон Д. Р. Семь баз данных за семь недель. Введение в современные базы данных и идеологию NoSQL. 2013, 384 с.
2. Бэнкер К. MongoDB в действии. Litres, 2022, 394 с.
3. Anderson J. C., Lehnardt J., Slater N. CouchDB: the definitive guide: time to relax. – " O'Reilly Media, Inc.", 2010. 300 p.