



УДК 332.3:528

ЗБЕРЕЖЕННЯ ГЕОПРОСТОРОВИХ ДАНИХ ДЛЯ ВИРІШЕННЯ ЗАВДАНЬ ІЗ ЗЕМЛЕУСТРОЮ

ЩЕПАК Віра¹, КАРЮК Алла²

Ключові слова

геопросторові дані,
бази даних,
землеустрій

Анотація

У науковій роботі представлені основні вимоги до сховищ бази даних та проблеми зберігання геопросторових даних для вирішення завдань із землеустрою. Підвищення ефективності використання ГІС при розробці комплексних планів розвитку територій спонукає до вибору таких баз даних, які можуть забезпечити високу швидкість запису та зчитування просторових даних. У роботі розглянуто основні поняття та характеристики нереляційної бази даних NoSQL. Проведено аналіз варіантів розміщення бази даних NoSQL та виокремлено переваги розміщення бази даних NoSQL на хмарах. Проведено аналіз типів NoSQL баз даних по функціональному застосуванню для збереження геопросторових даних для вирішення завдань із землеустрою.

Основні вимоги до сховищ БД та можливості зберігання геопросторових даних при вирішення завдань із землеустрою направлені на забезпечення високої швидкості запису та зчитування просторових даних. Основні вимоги до сховищ даних це простота, відмовостійкість, масштабування, продуктивність, сумісність і можливість отримання аналітичних зрізів даних [1]. База даних SQL – це реляційна база даних, розглядається як математична модель зберігання і розподілу даних. Вся інформація представлена у вигляді об'єктів, які розташовані в таблицях певної структури. Кожен об'єкт складається із властивостей, пов'язаних між собою ключами. При цьому деякі властивості можуть перебувати в окремих таблицях. Історично склалося так, що більшість Веб-додатків корпоративного рівня використовують реляційні бази даних [2].

При розробці комплексних планів розвитку території використання ГІС повинно забезпечувати високу швидкість запису та зчитування просторових даних [3, 4, 5]. При зміні значень певних властивостей може змінюватися загальна структура елементів, що визначає

¹ Національний університет «Полтавська політехніка імені Юрія Кондратюка», доцент кафедри автомобільних доріг, геодезії та землеустрою, к.т.н., ORCID: [0000-0003-2185-1181](https://orcid.org/0000-0003-2185-1181), e-mail: kanameshch@gmail.com

² Національний університет «Полтавська політехніка імені Юрія Кондратюка», доцент автомобільних доріг, геодезії та землеустрою, к.т.н., ORCID: [0000-0003-4839-024X](https://orcid.org/0000-0003-4839-024X), e-mail: kariuk15@ukr.net

ці властивості. Для вирішення цієї проблеми було створено NoSQL сховище, як нереляційну базу даних. База даних NoSQL може використовуватись для зберігання великої кількості геопросторових даних. Бази даних NoSQL – нереляційні бази даних, в яких весь обсяг інформації представлений без схеми, зокрема, у вигляді пари «ключ – значення», де значення – це набір всіх властивостей об'єкта. NoSQL дозволяє швидко записувати дані і звертатися до них та легко масштабуватися горизонтально. Така модель дає більше гнучкості при роботі з різними типами об'єктів. NoSQL бази даних не використовують загальний формат запиту, такий як у SQL, застосовуючи власну систему запитів [2]. Бази даних NoSQL можна масштабувати на тисячі серверів, хоча іноді і з втратою узгодженості даних. Щоб уникнути з'єднань і супутніх їм помилок і тим самим поліпшити обробку дуже великих наборів даних, в NoSQL запропоновано кілька альтернатив реляційної моделі. NoSQL бази даних особливо актуальними сьогодні тому, що вони особливо ефективні для роботи з великими наборами розподілених даних, що робить їх ефективними для великих проєктів зі зберігання даних і аналітики.

Проведено аналіз варіантів розміщення бази даних NoSQL геопросторових даних для вирішення завдань в землеустрої. Бази даних NoSQL для збереження геопросторових даних для вирішення завдань в землеустрої можна розташовувати як на локальних серверних платформах так і на хмарах. Основними перевагами розміщення бази даних NoSQL на локальних серверних платформах є безпека ваших даних більша, ніж при використанні хмарних технологій [2].

Сьогодні стало можливим використовувати ресурси навіть зарубіжних операторів центрів обробки даних, які пропонують послуги зі зберігання даних «в хмарі», розміщуючи їх в декількох дата-центрах, пов'язаних між собою високошвидкісними каналами зв'язку.

NoSQL бази даних можна розділити за їх особливостям функціонального застосування [2]. Поділ баз даних по функціональному застосуванню зображено на рисунку 1. Даний поділ не є вичерпуючий, в своїй класифікації аналізуємо тільки найбільш поширені бази NoSQL.

Проведений аналіз типів NoSQL баз даних по функціональному застосуванню показав, що для збереження просторових даних ГІС необхідно використовувати ті бази даних, які відповідають поставленим завданням.

Сховище ключ-значення. Представником такого сховища є база даних Redis Це сховище, яке використовує асоціативний масив (відомий як карта або словник) як основну модель даних. В цій моделі дані представляються як колекція пар типу «ключ-значення», таких, що кожен можливий ключ з'являється в колекціях не більше одного разу. Модель такого типу є однією із найпростіших, і більш розвинені моделі зазвичай реалізовані як їх розширення. Модель типу «ключ-значення» може бути розширена до скінченно впорядкованої, що підтримує ключі в лексикографічному порядку. Це розширення є обчислювально потужне в тому, що воно може ефективно видобувати селективні діапазони ключів. Моделі типу «ключ-значення» можуть використовувати різні рівні узгодженості, починаючи від випадкової і завершуючи серіалізованими даними. Деякі бази даних підтримують впорядковані ключі. Існують різні реалізації апаратного забезпечення. Деякі користувачі підтримують дані в пам'яті (RAM), в той час як інші використовують (SSD) накопичувачі або класичні (HDD) диски.

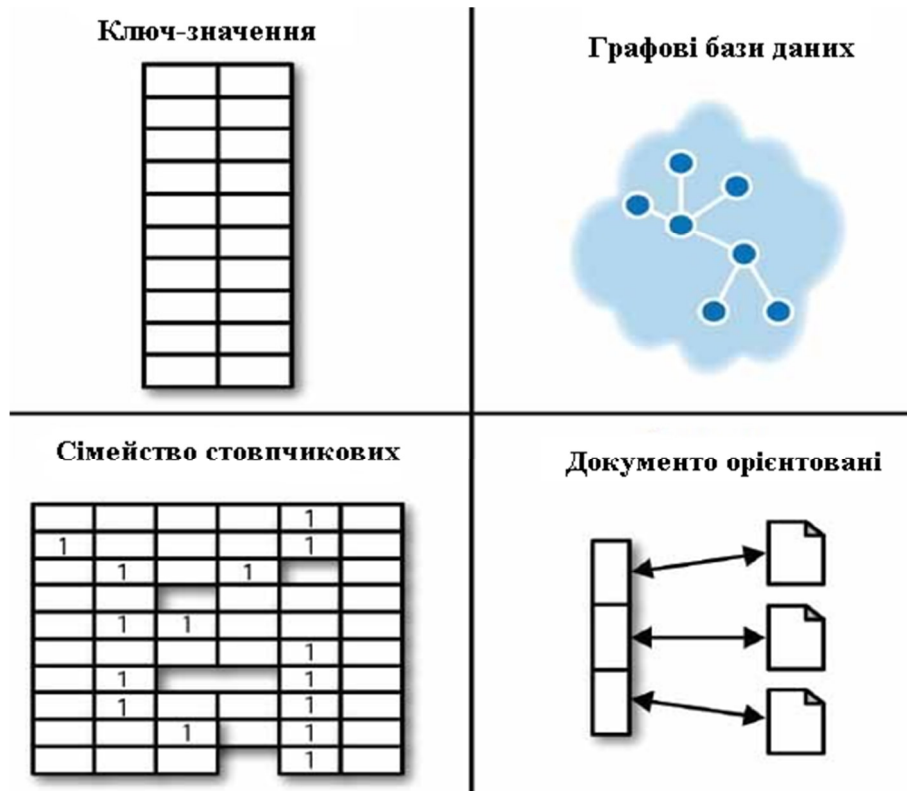


Рисунок 1 – Поділ баз даних по функціональному застосуванню

Сховища типу «ключ-значення» (key-value stores) найкраще підходять для постійного обміну даними декількома процесами або мікросервісами в додатку. Redis (Remote Dictionary Server), ярий приклад сховища типу «ключ-значення», зберігає дані в пам'яті, доступ до даних здійснюється по ключу. Копія даних може зберігатися на диску. Цей підхід забезпечує продуктивність, в десятки разів перевищує продуктивність реляційних СУБД [6].

Сховище на основі графів. Представником такого сховища є база даних Neo4j. Це графова база даних з відкритим вихідним кодом. Дані зберігаються у власному форматі збереження вузлів і ребер (спеціалізовані сховища графів). Цей тип баз даних розроблений для даних, де дані можуть бути представлені у вигляді складових вершин графу, об'єднаних скінченним числом зв'язків між ними. Механізми обчислення графів дозволяють виконувати глобальні графові обчислювальні алгоритми для великих наборів даних. Вони призначені для вирішення таких завдань, як ідентифікація кластерів даних. Таким типом даних можуть бути соціальні мережі, мережі громадського транспорту, карти доріг тощо. Самою поширеною графовою базою даних є Neo4j - це NoSQL графова СУБД з відкритим вихідним кодом, реалізована на Java і Scala. Проектом займається американська компанія Neo Technology [7].

Стовпчикове сховище. Представником такого сховища є база даних Apache Cassandra. У таких системах дані зберігаються у вигляді матриці з великими даними, рядки і стовпці якої використовуються як ключі. Це масштабована база даних з відкритим кодом, відмінно підходить для задач управління великим числом неструктурованих, напівструктурованих і структурованих даних в хмарах. При потребі збирати дані дуже швидко і у великих обсягах для аналітики, то ефективним буде сімейство стопчиків або, як їх ще називають – бази даних з широким значенням стовпця (wide column store).

Типовим застосуванням цього типу СУБД є Веб-індексування, а також завдання, пов'язані з великими даними, зі зниженими вимогами до узгодженості. Кассандра забезпечує безперебійну доступність, лінійну масштабність. Вона може використовуватися на декількох серверах без єдиної точки відмови. Задача, яка вирішується розробленою архітектурою Cassandra, це можливість обробки великого обсягу даних інформації і тисяч паралельних операцій користувачів в секунду [1].

Документо-орієнтоване сховище. Представником документно-орієнтованої бази даних є база даних MongoDB. Центральним поняттям такої моделі даних є «документ». Документо-орієнтовані бази зберігають інформацію у вигляді ієрархічних структур даних. Може йтися про об'єкти з довільним набором атрибутів. В реляційній БД буде розбито на кілька взаємопов'язаних таблиць, в нереляційних може зберігатися у вигляді цілісної сутності. База даних MongoDB допускає набагато більшу вкладеність і складність структури даних (наприклад, документ вкладений в документ, вкладений в документ). Документи знімають обмеження вкладеності першого і другого рівнів типу ключ-значення в розподілених сховищах. В цілому, можна описати як завгодно складну структуру даних як документ і зберегти в такій БД.

Область використання документо-орієнтованих баз даних включають персоналізацію, аналітику в реальному часі, великі дані (Big Data), каталоги продуктів / активів, мобільні додатки, концентратори даних, управління контентом і додатки для спілкування та спільної роботи. MongoDB підходить для ведення статистики та для зберігання просторових даних.

Висновок

У роботі проведено аналіз бази даних NoSQL, зокрема, розглянуто основні вимоги до сховищ даних та проблеми зберігання просторових даних. Проведено аналіз варіантів розміщення бази даних NoSQL для вирішення завдань у землеустрої, визначено переваги розміщення бази даних NoSQL на локальних серверних платформах та на хмарах. Проаналізовано NoSQL-бази даних по функціональному застосуванню, а саме бази даних: ключ-значення, графову, сімейство стовпчиків та документо-орієнтовану базу даних. Розглянуто найбільш поширені NoSQL бази даних MongoDB, Cassandra, Redis та Neo4j в контексті відповідності до потреб зберігання геопросторових даних.

Література

- [1] Ion Lungu, Manole Velicanu & Iuliana Botha. (2019). Database Systems – Present and Future. *Economic Informatics Department, Academy of Economic Studies*. Bucharest, Romania, vol. 13, 1/2009. – P. 84–99.
- [2] Marin Fotache, Dragos Cogean. (2013). NoSQL and SQL Databases for Mobile Applications. *Informatica Economica*. Vol., 2. February 2013, – P. 5–14.
- [3] Таратула Р.Б. (2017). Особливості геоінформаційного забезпечення земельно-інформаційної системи. *Збалансоване природокористування*. – 2017. – № 2. – С. 118–123.
- [4] Потапова Н.А., Лавринчук Д.Д. (2019). Інформаційні технології в управлінні земельними ресурсами. *Ефективна економіка*. 2019. № 11. [Електронний ресурс]. – Режим доступу: <http://www.economy.nayka.com.ua/?op=1&z=7412>. DOI: [10.32702/2307-2105-2019.11.65](https://doi.org/10.32702/2307-2105-2019.11.65)
- [5] Vyawahare H.R., Karde P.P. & Thakare V.M. (2017). Brief Review on SQL and NoSQL. *Thakare International Journal of Trend in Scientific Research and Development*. Vol. 2, Issue 1. Nov-Dec 2017, – P. 968–971.
- [6] Velicanu M., Lungu I., Muntean M. & Ionescu S. (2016). *Sisteme de baze de date – teorie și practică*, Ed. Petrion, Bucharest, 2016, 170 p.

- [7] Patrícia Cavoto, André Santanchè. (2016). ReGraph: Bridging Relational and Graph Databases. *Thesis presented to the Institute of Computing of the University of Campinas*, [Електронний ресурс]. – Режим доступу: <https://lis-unicamp.github.io/wp-content/uploads/2016/02/Cavoto2015c.pdf>

STORAGE OF GEOSPATIAL DATA FOR SOLVING PROBLEMS WITH LAND STRUCTURE

SHCHEPAK Vira, KARIUK Alla

Abstract. The scientific work presents the main requirements for database storage and the problems of storing geospatial data for solving problems with land management. Increasing the efficiency of using GIS in the development of complex plans for the development of territories encourages the selection of such databases that can provide high speed of recording and reading of spatial data. The paper examines the main concepts and characteristics of a non-relational NoSQL database. An analysis of options for placing a NoSQL database was carried out and the advantages of placing a NoSQL database on the cloud were highlighted. An analysis of the types of NoSQL databases by functional application for saving geospatial data for solving problems with land management was carried out.

Keywords: geospatial data, databases, land management