

Анализ и применение мультимодельной БД для построения хранилища данных для маркетплейса

Солдатов А.Ф. (Университет ИТМО)

Научный руководитель - доцент, к.ф.-м.н., Ромакина О.М. (Университет ИТМО)

Введение. Важным аспектом при проектировании сервисов электронной коммерции, в частности маркетплейсов является выбор СУБД для построения хранилища данных для этого сервиса. С каждым годом количество обрабатываемых создаваемых данных увеличивается более чем на 20% в год [1], у разработчиков стоит выбор использования отдельных моделей данных СУБД для конкретных задач, либо использование мультимодельной БД (в дальнейшем ММСУБД), которая содержит в себе несколько моделей хранения данных. В настоящее время ММСУБД используются небольшими продуктовыми компаниями, стартапами и интернет-магазинами. Большинство компаний используют подход polyglot persistence [2,3], который позволяет использовать дополнительную модель хранения данных в виде СУБД для решения отдельной задачи. Из-за этого разработчикам приходится работать с большим набором различных СУБД, которые сложно контролировать и поддерживать данные в согласованном состоянии. В связи с этим актуальными задачами исследования является поиск подходящей ММСУБД и ее сравнительный анализ с отдельными мономодельными СУБД.

Основная часть. В большинстве архитектур ММСУБД представлены две или три из четырех основных моделей хранения данных: реляционная, ключ-значение, графовая и документная [4]. За счет этого, ММСУБД могут обрабатывать данные вне зависимости от их структуры, что позволяет работать со структурированными, полуструктурированными и неструктурированными данными. Кроме этого, ММСУБД используют единый язык запросов [5, 6], который позволяет работать с каждой моделью хранения данных и облегчает работу разработчикам хранилища данных, так как им не придется изучать отдельный язык для работы с разными СУБД, как это происходит с подходом polyglot persistence.

В целях сравнения отдельных ММСУБД в исследовании будут рассмотрены такие их характеристики, как:

- производительность;
- масштабируемость;
- сложность в эксплуатации;
- поддерживаемость.

Это исследование поможет определить возможность использования ММСУБД для таких сфер, как электронная коммерция и маркетплейсы.

Заключение. В ходе выполнения работы были изучены основные подходы проектирования хранилищ данных, проведен сравнительный анализ характеристик ММСУБД с соответствующими отдельными мономодельными СУБД.

Список используемых источников

1. Statista, Объем данных/информации, созданных, собранных, скопированных и потребленных во всем мире с прогнозом до 2025 года [Электронный ресурс], - URL: <https://www.statista.com/statistics/871513/worldwide-data-created/>
2. Omar Lajam, Salahadin Mohammed, Revisiting Polyglot Persistence: From Principles to Practice, 2022, https://thesai.org/Downloads/Volume13No5/Paper_99-Revisiting_Polyglot_Persistence_From_Principles_to_Practice.pdf
3. Zdepski, Cristofer, Tarcizio Alexandre Bini and Simone Nasser Matos. PDDM: A Database Design Method for Polyglot Persistence, - 2020 - , - URL: <https://www.semanticscholar.org/paper/PDDM%3A-A-Database-Design-Method-for-Polyglot-Zdepski-Bini/97ab17cbaa37a967698b1b7c0fbe2cb8a2c5a418>

4. Justas Kazanavičius, Dalius Mažeika and Diana Kalibatiene, An Approach to Migrate a Monolith Database into Multi-Model Polyglot Persistence Based on Microservice Architecture: A Case Study for Mainframe Database, - 2022 -, <https://www.mdpi.com/2076-3417/12/12/6189/pdf>
5. SurrealQL, [Электронный ресурс], - 2023 -, - URL: <https://docs.surrealdb.com/docs/surrealql/overview>
6. AQL Data queries, [Электронный ресурс], - 2023 -, , - URL: <https://docs.arangodb.com/3.11/aql/data-queries>