

РАЗРАБОТКА ГРАФИЧЕСКОЙ НОТАЦИИ МОДЕЛИРОВАНИЯ NOSQL БАЗ ДАННЫХ

Клемешева А. С. (Университет ИТМО)

Научный руководитель – доцент, кандидат физико-математических наук,
Ромакина О. М. (Университет ИТМО)

Введение. Важным этапом проектирования баз данных является моделирование структуры данных. Моделирование проводится на концептуальном, логическом и физическом уровнях абстракции [1], формируются схемы, визуализирующие организацию данных. Разработанная визуализация помогает на ранних этапах разработки замечать ошибки и проблемы в структуре данных и находить способы повышения производительности базы данных. Кроме того, сформированные схемы облегчают последующую работу с базой данных, направленную на ее развитие и дополнение. Для моделирования реляционных баз данных разработаны и повсеместно активно используются различные нотации и методологии моделирования, существуют утвержденные стандарты, например, IDEF1x. Благодаря этому разработчики обладают не только удобным инструментарием моделирования, но и возможностью чтения документации, составленной другими разработчиками, что облегчает взаимодействие между специалистами.

Разработчики NoSQL баз данных сталкиваются с проблемой отсутствия единой нотации и подхода к визуализации организации неструктурированных и полуструктурированных данных. Это связано с распространенным мнением о том, что NoSQL базы данных являются бессхемными, что на деле не совсем верно. Данные в таких базах организованы согласно некоторой подразумеваемой схеме, но она менее жесткая, благодаря чему обеспечивается гибкость и адаптивность базы данных к новым требованиям. Исследователями предлагаются различные варианты решения проблемы. В основном существующие на данный момент разработки [2], [3], [4] касаются одной или двух моделей данных из наиболее распространенных (к которым относятся документная, графовая, ключ-значение и колоночная модели) и опираются на нотации, предназначенные для моделирования реляционных баз данных. Однако они не дают возможности моделирования других моделей данных, и не могут применяться в мультимодельным базам данных.

Основная часть. В структуре данных нереляционных баз данных существуют зависимости и взаимосвязи, характерные и для реляционных баз данных, поэтому разработанная графическая нотация для NoSQL должна включать паттерны нотаций, разработанных для моделирования реляционных баз данных. Это не только облегчит ее изучение разработчиками, но и обеспечит возможность моделирования мультимодельных баз данных, одна из моделей которых является реляционной.

Упомянутые выше четыре распространенные NoSQL модели данных можно в упрощенном виде представить как структуру ключ-значение, где ключ может быть составным, а значение – вектором характеристик. От этого объединяющего фактора можно отталкиваться при моделировании баз данных перечисленных выше моделей.

- 1) Ключ-значение. Представляет собой простую связку ключей и значений разных типов.
- 2) Графовые. Включают узлы с ключом и вектором характеристик и ребра, где в качестве ключа выступает пара вершин, а возможные веса и прочие свойства ребра также моделируются вектором характеристик.
- 3) Документные. У каждого документа есть ключ или уникальный идентификатор, чаще всего формируемый автоматически, набор наименований атрибутов и их значений, что можно рассматривать как совокупность пар ключ-значение. Коллекция, в свою очередь, также представляет собой структуру ключ-значение,

где ключом является наименование коллекции, а значением – совокупность документов, входящих в коллекцию.

- 4) Колоночные. Здесь используется концепция пространства ключей, подобная схеме в реляционных моделях.

В реляционных базах данных первичному ключу также соответствует набор наименований атрибутов и их значений.

Заключение. Изучены существующие нотации реляционных и NoSQL баз данных и определены ключевые графические паттерны, которые могут быть применены в единой нотации. Изучены наиболее распространенные модели данных NoSQL, рассмотрены их специфические особенности, предложен принцип моделирования этих моделей данных.

Список использованных источников

1. Технологии проектирования баз данных. / Осипов Д. Л. – М.: ДМК Пресс, 2019. – 498 с.
2. Smirnov M. V., Tolmasov R. S. Graphical Notation for Document Database Modeling // *Otkryt. obraz.* (Mosk.). 2021. Т. 25. № 5. С. 50–60.
3. De Sousa V. M., Cura L. M. D. V. Logical Design of Graph Databases from an Entity-Relationship Conceptual Model // *Proceedings of the 20th International Conference on Information Integration and Web-based Applications & Services.* 2018. С. 183–189.
4. Palagashvili A. M., Stupnikov S. A. Reversible Mapping of Relational and Graph Databases // *Pattern Recognit. Image Anal.* 2023. Т. 33. № 2. С. 113–121.