

SQL vs NoSQL

Максим КУКЛЕВ, Анна ЗИНГАН

Технический Университет Молдовы, Департамент Программной Инженерии и Автоматики

Аннотация: В данной работе проводится сравнение использования реляционных и нереляционных баз данных, основанных на управлении SQL, либо NoSQL.

Ключевые слова: базы данных, реляционные базы данных, нереляционные базы данных, workflow, сравнение.

Введение

Мир технологий баз данных держится на двух больших колоннах: реляционных и нереляционных баз данных, или SQL и NoSQL. Эти две структуры кардинально отличаются друг от друга в строении, замысле, проектировании, подходах к хранению информации. Однако какие из различий можно вывести на первый план, чтобы продемонстрировать разницу между этими двумя видами? Какую из моделей выбирать для своего проекта?

Говоря об общих сведениях, известно, что **реляционные БД** хранят структурированные данные, представляющие объекты реального мира. Например, информация о человеке, об ассортименте супермаркета и так далее. Всё это сгруппировано в таблицы, формат которых задаётся в процессе проектирования (рис. 1b).

Нереляционные, в свою очередь, выглядят несколько иначе. Базы с информацией о документах могут быть выстроены иерархически. Могут быть объекты с произвольным набором атрибутов, а то, что в реляционной БД разбивается на взаимосвязанные таблицы, здесь может содержаться в виде целостной сущности. (рис. 1a). Также известно, что нереляционные базы лучше поддаются масштабированию.

1. Преимущества SQL и NoSQL БД

Как правило, при выборе следует склониться в сторону SQL-баз, когда ваш проект подходит по двум следующим параметрам:

- База данных должна соответствовать требованиям ACID (атомарность, непротиворечивость, изолированность и долговечность), что позволяет уменьшить риск неожиданного поведения системы и обеспечить целостность базы. NoSQL, напротив, ставят в приоритет гибкость и скорость, нежели целостность хранимой информации;
- Данные должны быть структурированы, структура не подвержена частым изменениям. Если организация не находится в стадии экспоненциального роста, вероятно, не найдётся веских причин выбирать БД, вольно обращающуюся с данными и нацеленную на обработку огромных объёмов информации.

При использовании NoSQL-баз выделяют следующие преимущества:

- Хранение больших объёмов неструктурированной информации. База данных NoSQL не накладывает ограничений на введённые данные. Помимо этого, можно добавлять новые типы данных в процессе работы;
- Использование облачных вычислений и хранилищ. Данные должны легко распределяться между несколькими серверами для масштабирования;
- Быстрая разработка. Для agile-проектов не подходит реляционная модель, так как она замедляет работу.

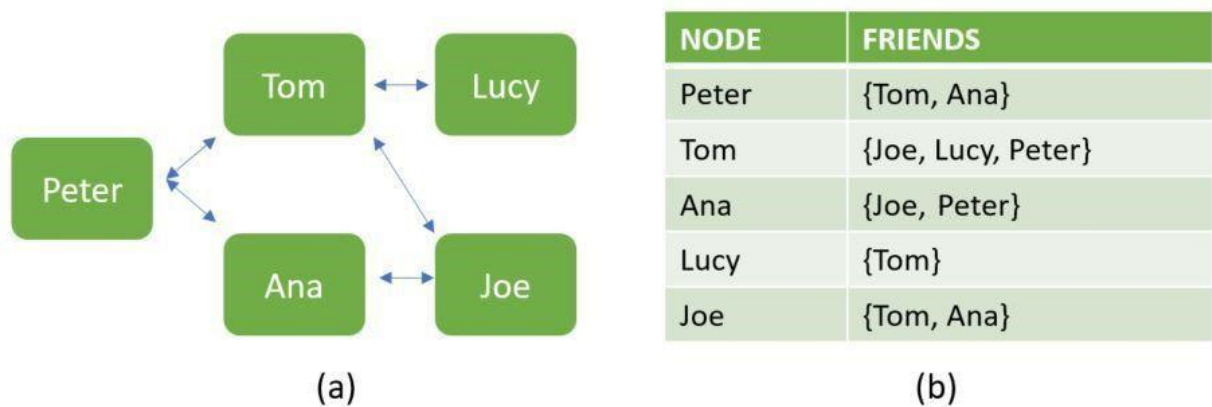


Рисунок 1 Представление данных в нереляционных (a) и реляционных (b) базах данных.

Как более продвинутый пример, для демонстрации того, когда SQL предпочтительнее NoSQL, рассмотрим особенности применения в NoSQL-базах алгоритмов уплотнения. Проблема заключается в том, что в некоторых NoSQL-базах (например, в CouchDB и HBase) постоянно приходится формировать так называемые sstables — строковые таблицы в формате ключ-значение, отсортированные по ключу. В такие таблицы, которые сохраняются на диск, данные попадают из таблиц, хранящихся в памяти, при их переполнении и в других ситуациях. При интенсивной работе с базой создание таблиц, со временем, приводит к тому, что подсистема ввода-вывода устройства хранения данных становится узким местом для операций чтения данных. Как результат, чтение в NoSQL-базе происходит медленнее, чем запись, что сводит на нет одно из главных преимуществ нереляционных баз данных. Именно для того, чтобы уменьшить этот эффект, системы NoSQL используют, в фоновом режиме, алгоритмы уплотнения данных, пытаясь объединить множество таблиц в одну. Но и сама по себе эта операция весьма ресурсоёмкая, система работает под повышенной нагрузкой.

Заключение

В итоге стоит сказать, что в современном мире нет противостояния между реляционными и нереляционными базами данных. Вместо этого стоит говорить об их совместном использовании для решения задач, на которых та или иная технология показывает себя лучше всего. Кроме того, всё сильнее наблюдается интеграция этих технологий друг в друга. Например, Microsoft, Oracle и Teradata сейчас предлагают некоторые формы интеграции с Hadoop для подключения аналитических инструментов, основанных на SQL, к миру неструктурированных больших данных.

Библиография

1. Alon Brody, *SQL или NoSQL — вот в чём вопрос*. - [Электронный ресурс]. – Режим доступа: <https://habr.com/company/ruvds/blog/324936/>
2. Alon Brody, *SQL vs NoSQL: The Differences Explained*. - [Электронный ресурс]. – Режим доступа: <https://blog.panoply.io/sql-or-nosql-that-is-the-question>
3. NoSQL. - [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/NoSQL>