



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004

## СРАВНЕНИЕ SQL И NOSQL: ПРЕИМУЩЕСТВА, НЕДОСТАТКИ И ОБЛАСТИ ПРИМЕНЕНИЯ

<sup>1</sup>Гасанов И.З., Ликсаков М.В.

ФГБОУ ВО "Национальный Исследовательский Университет "Высшая Школа Экономики", Москва, Россия (101000, Москва, ул. Мясницкая, д. 20), e-mail: <sup>1</sup>gasanov2727@yandex.ru

В данной работе представлен сравнительный анализ SQL и NoSQL баз данных. В статье выделяются сильные стороны SQL в обработке структурированных данных и сложных транзакций, в то время как для NoSQL выявлены преимущества масштабируемости и гибкости при управлении разнообразными наборами данных. Также были приведены примеры сфер для использования рассматриваемых подходов. Наконец, в работе подчеркивается важность выбора типа базы данных в соответствии с конкретными требованиями проекта и предлагается дальнейшее изучение гибридных моделей и отдельных типов NoSQL.

Ключевые слова: SQL, NoSQL, ACID, СУБД, базы данных.

## COMPARISON OF SQL AND NOSQL: ADVANTAGES, DISADVANTAGES, AND USE CASES

<sup>1</sup>Hasanov I.Z., Leksakov M.V.

National Research University Higher School of Economics, Moscow, Russia (101000, Moscow, Myasnitskaya Street, 20), e-mail: <sup>1</sup>gasanov2727@yandex.ru

The paper presents a comparative analysis of SQL and NoSQL databases. The article highlights the strengths of SQL in processing structured data and complex transactions, while emphasizing the advantages of NoSQL in terms of scalability and flexibility for managing diverse datasets. The study also examines examples of domains suitable for each approach. Finally, the importance of choosing the appropriate database type based on specific project requirements is emphasized, and further research on hybrid models and individual types of NoSQL databases is recommended for future investigations.

Keywords: SQL, NoSQL, ACID, RDBMS, database.

Безусловно, сегодня объем данных и их сложность растут с невероятной скоростью. Если отсечь сгенерированную человечеством информацию за последние несколько лет, то останется всего около 10% данных за все время нашего существования. Поэтому, вместе с данными растут и технологии для работы с ними. Так, в 1974 году был создан язык структурированных запросов SQL, который активно применяется для табличных баз данных (БД), в то время как понятие NoSQL, или "не только SQL", появилось лишь в 1998 году и предоставляет возможности для обработки информации разного формата. Значимость данного сравнительного анализа очевидна, так как современный мир в каком-то смысле основан на данных, и эффективное управление информацией, а также готовность к ее росту и разнообразию помогут государственным и негосударственным организациям успешно вести свою деятельность.

Итак, начнем с обзора реляционных систем управления базами данных (РСУБД), то есть тех, в которых данные хранятся в виде таблиц (relational – имеющий отношения). В таких СУБД каждая строка соответствует записи, а каждый столбец – атрибуту с определенным типом данных, что позволяет задать структуру и предопределяет шаблон для каждой ячейки. “Реляционные базы данных основаны на реляционной модели – интуитивно понятном и простом способе представления данных в таблицах” [1, с.343]. Также в последнее время набирают популярность и колоночные СУБД, в которых, наоборот, данные организованы по столбцам вместо строк. Такие СУБД предназначены для эффективных операций чтения и аналитики (OLAP сценариев), и это частный случай реализации реляционной модели, не изменяющий общей концепции структуры данных.

Говоря о полезных свойствах реляционных БД, следует отметить такие особенности, как атомарность, согласованность, изолированность и долговечность. Все перечисленные свойства формируют общеизвестную аббревиатуру ACID и позволяют гарантировать целостность данных. Атомарность означает, что группа операций будет выполнена либо полностью, либо не выполнена совсем (значит, например, в банке не будет транзакций с частичным успехом, когда денежные средства со счета списаны, но не зачислены реципиенту). Согласованность гарантирует, что транзакции будут соблюдать все ограничения целостности, определенные в базе данных. Следующее свойство – изолированность – предотвращает конфликты и несогласованности данных, которые могут возникнуть при одновременном доступе нескольких транзакций к одной и той же информации. Наконец, долговечность означает, что успешно завершённые транзакции сохраняют свои изменения даже в случае сбоя или перезагрузки системы.

Таблица 1 – Базы данных SQL

База данных	Год создания	Особенности
Oracle Database	1979	Многопользовательская, в памяти, машинное обучение
Microsoft SQL Server	1989	Поддержка XML, транзакционная репликация, моментальная репликация
PostgreSQL	1989	Расширяемость, MVCC, процедурные языки
MySQL	1995	Репликация, разделение, хранимые процедуры

SQLite	2000	Без сервера, с нулевой конфигурацией, транзакционная
--------	------	--

Наиболее популярными РСУБД с открытым исходным кодом являются MySQL и PostgreSQL, при этом последний считается подходящим решением для более сложных операций с большими объемами данных. Microsoft SQL Server и Oracle Database, с другой стороны, являются примерами коммерческих баз данных, которые предлагают расширенные функции, такие как поддержка XML, репликация транзакций и обработка в памяти. “Главным недостатком является высокая стоимость, которая не позволяет использовать ее в малых или средних проектах” [2, с.3]. Самым простым решением можно считать SQLite, который предоставляет движок базы данных без сервера и с нулевой конфигурацией (Таблица 1).

Основным преимуществом баз данных SQL является высокий уровень согласованности, что делает их надежным выбором для приложений, где целостность данных имеет первостепенное значение. Они предназначены для сложных запросов и транзакций, которым требуются свойства ACID. SQL отлично подходит для обработки структурированных данных с фиксированной схемой, он имеет широкую поддержку сообщества и большого числа опытных разработчиков.

Несмотря на свои сильные стороны, данный подход имеет и определенные ограничения. Так, базы данных SQL в первую очередь предназначены для вертикального масштабирования, что предполагает увеличение мощности одного и того же сервера, а следовательно, в определенный момент приведет к трудностям роста. Кроме того, хотя свойства ACID и повышают целостность данных, они могут снижать производительность в сценариях, где требуется высокая доступность или обработка больших наборов данных в режиме реального времени.

Рассмотрим теперь нереляционные базы данных. Отличие состоит в том, что их схема является динамической и может меняться в любой момент времени. “NoSQL – это система управления нереляционной базой данных, база данных быстрого поиска информации и переносимая” [3, с.23]. NoSQL появилась для удовлетворения потребностей современных приложений с большими и сложными данными. Эти базы данных обычно отказываются от свойств ACID в пользу BASE, повышая масштабируемость и гибкость. Базы данных NoSQL подразделяются на четыре типа: БД документов, БД с ключами и значениями, а также БД с широкими столбцами и графические БД. В целом, конструкция NoSQL обеспечивает горизонтальное масштабирование и эффективную обработку неструктурированных данных, что является преимуществом в ряде современных приложений.

Таблица 2 - Базы данных NoSQL

База данных	Дата выхода	Особенности
CouchDB	2005	Репликация, RESTful интерфейс, конечная согласованность
Neo4j	2007	ACID, хранилище графов, язык запросов Cypher
MongoDB	2009	Специальные запросы, индексация, репликация
Apache Cassandra	2008	Распределенная репликация с высокой доступностью в нескольких центрах обработки данных
Redis	2009	В памяти, Pub/Sub, транзакции

MongoDB (документно-ориентированная БД) и Apache Cassandra (хранилище с широкими столбцами) предлагают такие функции, как специальные запросы, высокая доступность и репликация в нескольких центрах обработки данных. В то время как БД Redis (хранилище структур данных в памяти) славится своей производительностью и поддержкой различных структур данных, а CouchDB предоставляет интерфейс RESTful и определяет приоритеты репликации данных и возможной согласованности. Наконец, Neo4j выделяется как графическая БД, которая превосходно справляется с управлением взаимосвязями данных и предлагает мощный язык графических запросов Cypher (Таблица 2).

Базы данных NoSQL обладают рядом преимуществ по сравнению с традиционными РБД. Наиболее существенным является их способность к горизонтальному масштабированию, то есть добавлению большего количества серверов в сеть, что делает их эффективным решением для обработки больших наборов данных. Разнообразные модели NoSQL позволяют осуществлять индивидуальное управление данными, оптимизируя производительность в зависимости от конкретных вариантов использования. Кроме того, БД NoSQL часто следуют модели BASE («в основном доступны», «мягкое состояние», «конечная согласованность»), это обеспечивает более высокую доступность и отказоустойчивость.

Безусловно, БД NoSQL не лишены и недостатков. Отсутствие фиксированной схемы может привести к несогласованности и дублированию данных. Хотя модель BASE и полезна, она может нарушить согласованность данных в распределенных системах. Более того, в базах данных NoSQL отсутствуют стандартизированные интерфейсы, что приводит к различиям на разных платформах и меньшей поддержке сообщества по сравнению с SQL.

В мире веб-приложений и мобильных устройств базы данных NoSQL служат базовой технологией для платформ социальных сетей, игровых приложений и потоковых сервисов, которым часто приходится управлять огромными объемами пользовательских данных. Также NoSQL используют приложения со сложными взаимосвязями и сетевыми данными, такие как системы рекомендаций, сайты социальных сетей и системы обнаружения мошенничества. Ярким примером является Netflix – ведущий потоковый сервис для управления историей просмотров своих клиентов и рекомендациями по фильмам. Сервис обрабатывает огромное количество пользовательских данных и улучшает пользовательский опыт за счет персонализации контента.

В заключение следует отметить, что дискуссия о сравнении SQL и NoSQL, конечно, не заканчивается тем, что один из них превосходит другой; вместо этого подчеркиваются их различные возможности и оптимальные условия. SQL, с его высокой согласованностью и поддержкой сложных запросов, превосходит в сценариях, где данные структурированы, а транзакции являются сложными. Она служит основой для многих традиционных финансовых систем и систем учета запасов. В это же время, NoSQL, с его гибкой схемой, масштабируемостью и различными моделями данных, хорошо приспособлен для решения сложностей современных приложений, работающих с разнообразными и быстро меняющимися данными.

Задумываясь о перспективах развития, отметим, что все больший интерес вызывают гибридные модели, которые используют сильные стороны как SQL, так и NoSQL. Кроме того, поскольку NoSQL сам по себе является широкой категорией, более глубокое изучение отдельных типов баз данных NoSQL может дать лучшее представление об их уникальных преимуществах и подходящих вариантах использования. Эта растущая область технологий баз данных продолжает развиваться, предоставляя организациям множество возможностей для эффективного управления своими данными.

### **Список литературы**

1. Трофименко М. С., Кумратова А. М., Пополитов А. В. Изучение принципов работы NoSQL и SQL баз данных и их сравнение на примере Postgres Pro и MongoDB //Цифровизация экономики: направления, методы, инструменты. – 2021. – С. 342-347.
2. Савоськин И. В., Фирсов А. О. Исследование способов применения NoSQL и реляционных баз данных //E-Scio. – 2019. – №. 6 (33). – С. 101-108.
3. Ахмедов К. М. Базы данных SQL и NoSQL //Вестник магистратуры. – 2020. – №. 5-3 (104). – С. 23-25.

### **References**

1. Trofimenko M. S., Kumratova A. M., Popolitov A. V. Studying the principles of operation of NoSQL and SQL databases and their comparison on the example of Postgres Pro and MongoDB // Digitalization of the economy: directions, methods, tools. - 2021. - pp. 342-347.
  2. Savoskin I. V., Firsov A. O. Study of ways to use NoSQL and relational databases // E-Scio. – 2019. – no. 6 (33). - pp. 101-108.
  3. Akhmedov K. M. SQL and NoSQL databases // Bulletin of the Magistracy. – 2020. – no. 5-3 (104). - pp. 23-25.
-