

УДК 004.65

DOI: 10.18413/2518-1092-2022-8-3-0-7

Гармашев М.А.
Резников Н.Г.

**ТЕХНОЛОГИЯ NOSQL: ТЕХНИЧЕСКИЕ ОСОБЕННОСТИ,
ПЕРСПЕКТИВЫ РАЗВИТИЯ И СРАВНИТЕЛЬНЫЕ
ПРЕИМУЩЕСТВА ПЕРЕД SQL**

Белгородский государственный национальный исследовательский университет,
ул. победы, 85, Белгород, 308015, Россия

e-mail: garmashev@bsu.edu.ru

Аннотация

В данной обзорной статье рассматриваются концепции, технические особенности и перспективы развития технологии NoSQL, подчеркиваются ее преимущества по сравнению с обычными базами данных SQL. Базы данных NoSQL привлекли внимание благодаря своей способности эффективно обрабатывать огромные объемы неструктурированных и полуструктурированных данных. Авторы проводят краткий анализ ключевых характеристик баз данных NoSQL через частичное сравнение с классической SQL. В статье рассматриваются преимущества технологии NoSQL по сравнению с традиционными базами данных SQL, такие как повышенная производительность, упрощенное моделирование данных, повышенная масштабируемость и поддержка распределенных сред. В заключение перечислены будущие перспективы технологии NoSQL и рост используемости в современных приложениях, ориентированных на большие данные.

Ключевые слова: информация; аналитика; интернет; хранилище; ключи; значения; база данных; графовая база данных; распределенная база данных; горизонтальное масштабирование

Для цитирования: Гармашев М.А., Резников Н.Г. Технология NoSQL: технические особенности, перспективы развития и сравнительные преимущества перед SQL // Научный результат. Информационные технологии. – Т.8, №3, 2023. – С. 56-62. DOI: 10.18413/2518-1092-2022-8-3-0-7

Garmashev M.A.
Reznikov N.G.

**NOSQL TECHNOLOGY: TECHNICAL FEATURES,
DEVELOPMENT PROSPECTS AND COMPARATIVE
ADVANTAGES OVER SQL**

Belgorod State National Research University, 85 Pobedy St., Belgorod, 308015, Russia

e-mail: garmashev@bsu.edu.ru

Abstract

This review article examines the concepts, technical features, and future developments of NoSQL technology, highlighting its advantages over conventional SQL databases. NoSQL databases have gained attention due to their ability to efficiently process huge volumes of unstructured and semi-structured data. The authors provide a brief analysis of the key characteristics of NoSQL databases through a partial comparison with classical SQL. This article discusses the benefits of NoSQL technology over traditional SQL databases, such as improved performance, simplified data modeling, increased scalability, and support for distributed environments. Finally, the future prospects of NoSQL technology and its increasing use in modern big data applications are listed.

Keywords: information; analytics; Internet; storage; keys; values; database; graph database; distributed database; horizontal scaling

For citation: Garmashev M.A., Reznikov N.G. NoSQL technology: technical features, development prospects and comparative advantages over SQL // Research result. Information technologies. – Т.8, №3, 2023. – P. 56-62. DOI: 10.18413/2518-1092-2022-8-3-0-7

ВВЕДЕНИЕ

В последние годы наблюдается значительный рост использования технологии NoSQL благодаря ее способности обрабатывать большие объемы неструктурированных данных и обеспечивать высокую масштабируемость и производительность. Программисты заинтересованы в использовании технологии NoSQL, поскольку она обеспечивает гибкое моделирование данных, более быстрые циклы разработки и горизонтальную масштабируемость. Базы данных NoSQL также поддерживают различные типы данных и хорошо подходят для таких случаев использования, как аналитика в реальном времени, системы управления контентом и приложения с быстро меняющимися требованиями к данным. Базы данных NoSQL используются различными организациями в России, включая технологические компании, стартапы, государственные учреждения, исследовательские институты и предприятия в различных секторах, таких как финансы, телекоммуникации, электронная коммерция и производство. В число известных компаний, использующих базы данных NoSQL в России, входят Яндекс, Mail.Ru Group, Сбербанк, Газпром, Ростелеком и ВКонтакте. NoSQL (что означает «не только SQL») – это категория систем баз данных, которые отличаются от традиционных реляционных баз данных. Базы данных NoSQL не были изобретены одним человеком, а возникли как ответ на требования современных веб-приложений и больших данных. Различные компании и разработчики внесли свой вклад в разработку баз данных NoSQL, включая Apache Cassandra, MongoDB и CouchDB.

КРАТКИЙ АНАЛИЗ ПЕРСПЕКТИВЫ ТЕХНОЛОГИИ NOSQL

1. Гибкость схемы:

Базы данных NoSQL предлагают гибкую модель данных, позволяющую разработчикам хранить и извлекать данные, не придерживаясь предопределенной схемы. Такая гибкость учитывает меняющиеся требования к данным и облегчает быструю разработку и итерацию, что делает NoSQL идеальным для гибкой разработки и приложений больших данных.

2. Масштабируемость и высокая производительность:

Базы данных NoSQL предназначены для горизонтального масштабирования, то есть они могут обрабатывать большие объемы данных, распределяя их по нескольким серверам или кластерам. Это обеспечивает высокую пропускную способность записи и чтения, что делает базы данных NoSQL хорошо подходящими для обработки огромных рабочих нагрузок и обслуживания приложений реального времени.

3. Репликация и доступность данных:

Большинство баз данных NoSQL предлагают автоматическую репликацию данных, гарантируя, что данные реплицируются на несколько узлов для обеспечения отказоустойчивости и высокой доступности. Эта функция предотвращает единые точки отказа и обеспечивает встроенную избыточность, гарантируя долговечность данных и бесперебойное обслуживание.

4. Запрос на гибкость:

Базы данных NoSQL используют различные модели запросов, такие как базы данных «ключ-значение», «документ», «семейство столбцов» и «граф». Каждая модель предлагает уникальные возможности запроса, подходящие для конкретных шаблонов доступа к данным и сценариев использования. Базы данных NoSQL часто обеспечивают эффективный поиск по ключам, широкие возможности выполнения запросов и поддержку сложных структур данных [1].

5. Распределенная и отказоустойчивая архитектура:

Базы данных NoSQL созданы с учетом распределенной архитектуры. Они используют такие методы, как сегментирование и секционирование, для распределения данных по нескольким серверам, обеспечивая горизонтальное масштабирование. Кроме того, базы данных NoSQL предназначены для корректной обработки сбоев серверов, обеспечивая высокую доступность и отказоустойчивость.

6. Поддержка неструктурированных и полуструктурированных данных:

Базы данных NoSQL превосходно справляются с обработкой неструктурированных и полуструктурированных данных, таких как документы JSON, XML или двоичный контент. Они обеспечивают гибкость хранения и индексирования этих разнообразных форматов данных, позволяя эффективно находить и анализировать сложные структуры данных.

7. Типы баз данных NoSQL

- Документоориентированные базы данных (например, MongoDB)
- Столбчатые базы данных (например, Apache Cassandra)
- Базы данных «ключ-значение» (например, Redis)
- Графовые базы данных (например, Neo4j)

СРАВНЕНИЕ БАЗЫ ДАННЫХ NOSQL И БАЗЫ ДАННЫХ SQL

1. Модель данных. Базы данных SQL используют структурированную и предопределенную схему с таблицами, строками и столбцами, тогда как базы данных NoSQL используют гибкую и динамическую модель данных, часто основанную на парах ключ-значение, документах, графиках или хранилищах с широкими столбцами.

2. Масштабируемость. Базы данных NoSQL предназначены для горизонтального масштабирования, что означает, что они могут обрабатывать большие объемы данных и высокую нагрузку трафика на нескольких серверах. Базы данных SQL обычно масштабируются вертикально, требуя более мощного оборудования по мере роста данных.

3. Гибкость схемы. Базы данных NoSQL обеспечивают гибкость схемы, позволяя легко добавлять, изменять и удалять поля в модели данных, не затрагивая существующие данные. Бадам данных SQL требуется предопределенная схема, а изменения могут потребовать модификации схемы и миграции данных.

4. Язык запросов. Базы данных SQL используют язык запросов SQL, стандартизированный язык для управления данными и манипулирования ими. Базы данных NoSQL часто имеют свои собственные языки запросов, которые могут различаться в зависимости от типа базы данных. Некоторые базы данных NoSQL также поддерживают запросы, подобные SQL.

5. Соответствие ACID. Базы данных SQL обычно соответствуют требованиям ACID (атомарность, согласованность, изоляция, долговечность), что обеспечивает согласованность и надежность данных. Базы данных NoSQL по умолчанию обеспечивают итоговую согласованность, отдавая приоритет высокой доступности и масштабируемости, а не строгой согласованности.

6. Варианты использования. Базы данных SQL хорошо подходят для приложений со структурированными и сложными связями, требующих строгой целостности данных, таких как транзакционные системы и приложения, использующие финансовые данные. Базы данных NoSQL часто используются в сценариях с большими объемами неструктурированных или полуструктурированных данных, обработкой данных в реальном времени и высокими требованиями к масштабируемости [2]. На рисунке представлен краткий анализ сравнения разных технологий.



Рис. Краткий анализ сравнения разных технологий
Fig. Brief comparison analysis of different technologies

Важно отметить, что это общие характеристики, и существуют различные типы баз данных NoSQL, каждая из которых имеет свои сильные и слабые стороны. Выбор между SQL и NoSQL зависит от конкретных требований и характера разрабатываемого приложения или системы.

НЕКОТОРЫЕ РАСПРОСТРАНЕННЫЕ ОШИБКИ ПРИ ПЕРЕХОДЕ ОТ БАЗ ДАННЫХ SQL К БАЗАМ ДАННЫХ NOSQL

Следует знать, что успешный переход от SQL к NoSQL требует понимания различий между этими типами баз данных и соответствующей адаптации подхода программиста.

1. Игнорирование моделирования данных. Базы данных NoSQL требуют другого подхода к моделированию данных по сравнению с SQL. Неспособность разработать соответствующую модель данных может привести к проблемам с производительностью и неэффективным запросам.

2. Чрезмерное использование денормализации. Хотя в некоторых случаях денормализация может улучшить производительность запросов, чрезмерное ее использование может привести к избыточности данных и увеличению сложности. Важно найти баланс и тщательно взвесить компромиссы.

3. Пренебрежение индексацией. Базы данных NoSQL часто предоставляют гибкие варианты схемы, но эта гибкость может привести к проблемам с производительностью, если индексы не используются эффективно. Правильная индексация имеет решающее значение для оптимизации производительности запросов.

4. Не учитывается масштабируемость. Базы данных NoSQL известны своей масштабируемостью, но важно планировать будущий рост. Неспособность разработать масштабируемую архитектуру может привести к трудностям при увеличении объема данных или рабочей нагрузки.

5. Недооценка проблем согласованности данных. Базы данных NoSQL, особенно ориентированные на высокую масштабируемость, могут обеспечивать более слабые гарантии согласованности по сравнению с традиционными базами данных SQL. Очень важно понять модель согласованности выбранной базы данных NoSQL и соответствующим образом спроектировать приложение.

6. Недостаточное знакомство с языками запросов и API. Базы данных NoSQL часто имеют свои собственные языки запросов и API, которые могут отличаться от SQL. Неспособность изучить и понять эти новые языки и API может привести к неэффективным или неверным запросам.

7. Предполагается, что NoSQL является универсальным решением: базы данных NoSQL превосходны в определенных случаях использования, но могут подходить не для всех сценариев. Важно тщательно оценить требования конкретного приложения и выбрать технологию базы данных, которая лучше всего соответствует техническому заданию [3].

NoSQL может различаться по сложности в зависимости от конкретной технологии баз данных и предыдущего опыта работы с базами данных. Некоторые базы данных NoSQL довольно просты в освоении, в то время как другие могут быть требовательны к обучению. Однако в целом базы данных NoSQL спроектированы так, чтобы быть гибкими и масштабируемыми, что может облегчить работу с ними в определенных случаях использования по сравнению с традиционными реляционными базами данных.

НЕДОСТАТКИ ТЕХНОЛОГИИ NOSQL

Важно отметить, что обозреваемые недостатки относятся не ко всем базам данных NoSQL, поскольку существуют разные типы и реализации с разными функциями и компромиссами.

1. Отсутствие стандартизированного языка запросов. В базах данных NoSQL часто отсутствует стандартизированный язык запросов, такой как SQL, что усложняет выполнение сложных запросов и анализ данных.

2. Ограниченная поддержка транзакций. Базы данных NoSQL могут не обеспечивать полные гарантии транзакций ACID (атомарность, согласованность, изоляция, долговечность), что делает их менее подходящими для приложений, требующих строгой согласованности данных.

3. Дублирование и денормализация данных. Базы данных NoSQL часто денормализуют данные и дублируют их в нескольких документах или коллекциях для оптимизации производительности чтения. Это может привести к увеличению использования дискового пространства и потенциальной несогласованности данных, если обновления выполняются неправильно.

4. Ограниченные инструменты и экосистема. По сравнению с реляционными базами данных базы данных NoSQL часто имеют меньший набор инструментов и экосистемную поддержку, что может усложнить такие задачи, как мониторинг, резервное копирование и восстановление, а также управление схемой.

5. Недостаточная зрелость и стабильность. Некоторые базы данных NoSQL являются относительно новыми и могут не обладать зрелостью и стабильностью более устоявшихся реляционных баз данных. Это может привести к потенциальным проблемам с ошибками, производительностью и целостностью данных [5].

6. Более крутая кривая обучения. Базы данных NoSQL часто требуют от разработчиков изучения новых моделей данных и языков запросов, что может быть более сложным и трудоемким по сравнению с традиционными базами данных на основе SQL [5].

ЗАКЛЮЧЕНИЕ

В данной обзорной статье представлен всесторонний обзор технологии NoSQL, ее технических особенностей и перспектив развития. В нем подчеркиваются преимущества баз данных NoSQL по сравнению с традиционными системами SQL, подчеркиваются их масштабируемость, гибкость и преимущества в производительности. Рост больших данных и потребность в аналитике в реальном времени создают возможности для баз данных NoSQL для эффективной обработки крупномасштабных и быстро меняющихся наборов данных. Институты и исследовательские организации России активно участвуют в изучении и развитии технологий управления данными, в том числе NoSQL. Потенциальное сотрудничество между научными кругами и промышленными

партнерами облегчит обмен знаниями и может привести к прогрессу в разработке NoSQL. Поскольку объем и сложность данных продолжают расти, технология NoSQL, вероятно, будет играть решающую роль в современном управлении данными и аналитике. Перспективы развития NoSQL в России могут быть многообещающими благодаря растущему спросу на масштабируемые решения для управления данными, процветающей экосистеме стартапов, академическим исследованиям и государственной поддержке.

Список литературы

1. Попов В.Б. Технологии "nosql" в алгоритмах анализа больших данных и искусственном интеллекте / В. Б. Попов, И. В. Гавриков // Проблемы информационной безопасности: V Всероссийская с международным участием научно-практическая конференция, Симферополь – Гурзуф, 14-16 февраля 2019 года / Крымский федеральный университет имени В.И. Вернадского. – Симферополь – Гурзуф: ИП Зуева Т.В., 2019. – С. 158-160. – EDN YXYRAD.
2. Абелян А.А., Шерстобитов Г.В., Уварова А.А., Плужник Е.Н. Нереляционные базы данных. Обзор технологии nosql. Студенческая наука для развития информационного общества: сборник материалов VII Всероссийской научно-технической конференции, Ставрополь, 26–28 декабря 2017 года. Том Часть 2. – Ставрополь: Северо-Кавказский федеральный университет, 2018. – С. 15-19. – EDN YOTOKG.
3. Матыцын А.В., Есина А.Р. Перспективы интеграции технологии nosql в производственный процесс предприятия. Экономическое развитие в XXI веке: тенденции, вызовы и перспективы: Сборник научных трудов IX Международной научно-практической конференции «Горизонты России» В двух частях.
4. Буслаева У.Ю. Тенденции развития noSQL-технологий и систем управления базами данных. Тридцатая годовичная сессия ученого совета Сыктывкарского государственного университета имени Питирима Сорокина: Сборник статей Национальной конференции. В 2-х частях, Сыктывкар, 01–28 февраля 2023 года / Отв. редакторы О.А. Сотникова, Н.Н. Новикова. Том Часть 2. – Сыктывкар: Сыктывкарский государственный университет им. Питирима Сорокина, 2023. – С. 276-278. – EDN VBWTJQ.
5. Быкова К.И., Кузьмичева Е.А., Павлов П.А. Преимущества и недостатки больших баз данных, проблемы их сжатия и разработки. Информационные технологии в образовательном процессе вуза и школы: Материалы XVI Всероссийской научно-практической конференции, Воронеж, 30 марта 2022 года / Редколлегия: Р.М. Чудинский (науч. ред.), В.В. Малев, А.А. Малева (отв. ред.). – Воронеж: Воронежский государственный педагогический университет, 2022. – С. 47-53. – EDN FFGFLD.

References

1. Popov V.B., Gavrikov I.V. "Nosql" technologies in big data analysis algorithms and artificial intelligence. Problems of information security: V All-Russian scientific and practical conference with international participation, Simferopol – Gurzuf, February 14-16, 2019 / Crimean Federal University named after V.I. Vernadsky. – Simferopol – Gurzuf: IP Zueva T.V., 2019. – P. 158-160. – EDN YXYRAD.
2. Abelyan A.A., Sherstobitov G.V., Uvarova A.A., Pluzhnik E.N. Non-relational databases. Review of nosql technology. Student science for the development of the information society: collection of materials of the VII All-Russian Scientific and Technical Conference, Stavropol, 26–28 December 2017. Volume Part 2. – Stavropol: North Caucasus Federal University, 2018. – pp. 15-19. – EDN YOTOKG.
3. Matytsyn A.V., Esina A.R. Prospects for integrating nosql technology into the production process of an enterprise. Economic development in the 21st century: trends, challenges and prospects: Collection of scientific papers of the IX International Scientific and Practical conference "Russian Horizons" In two parts
4. Buslaeva U.Yu. Trends in the development of noSQL technologies and database management systems. Thirtieth annual session of the Academic Council of Syktyvkar State University named after Pitirim Sorokin: Collection of articles of the National Conference. In 2 parts, Syktyvkar, February 01–28, 2023 / Rep. editors O.A. Sotnikova, N.N. Novikova. Volume Part 2. – Syktyvkar: Syktyvkar State University named after. Pitirim Sorokina, 2023. – pp. 276-278. – EDN VBWTJQ.
5. Bykova K.I., Kuzmicheva E.A., Pavlov P.A. Advantages and disadvantages of large databases, problems of their compression and development. Information technologies in the educational process of universities and schools: Materials XVI All-Russian Scientific and Practical Conference, Voronezh, March 30, 2022 / Editorial Board: R.M. Chudinsky (scientific editor), V.V. Malev, A.A. Maleva (responsible editor). – Voronezh: Voronezh State Pedagogical University, 2022. – P. 47-53. – EDN FFGFLD.

Гармашев Михаил Александрович, старший преподаватель кафедры гражданского права и процесса юридического института

Резников Никита Григорьевич, ассистент кафедры информационных и робототехнических систем института инженерных и цифровых технологий

Garmashev Mikhail Aleksandrovich, senior lecturer of the Department of Civil Law and Process of the Law Institute

Reznikov Nikita Grigorievich, assistant of the department of information and robotic systems of the Institute of Engineering and Digital Technologies