DOI 10.37882/2223-2966.2024.10-2.09

ПОДХОДЫ К ОБЕСПЕЧЕНИЮ БЕЗОПАСНОСТИ РАСПРЕДЕЛЕННЫХ БАЗ ДАННЫХ

APPROACH TO ENSURING SECURITY OF DISTRIBUTED DATABASES

A. Kolesnikov

Summary. The study presented in the article is devoted to the consideration of various approaches to ensuring the security of distributed databases. In the process of analysis, key problems of information protection in distributed databases are identified. Approaches to information protection related to backup, authentication, encryption and regular security audits are described. The author's basic scheme and algorithm for ensuring the security of a distributed database are presented. Various strategies and approaches are considered that allow protecting distributed databases from malicious actions and intrusions. It seems that the development of artificial intelligence technologies, blockchain, machine learning and other breakthrough solutions will improve the efficiency and simplify the protection of distributed databases, making them more reliable.

Keywords: distributed database, protection, encryption, copy.

Колесников Антон Александрович
Санкт-Петербургский Политехнический университет Петра Великого
anton.kolesnikov.science@mail.ru

Аннотация. Представленное в статье исследование посвящено рассмотрению различных подходов к обеспечению безопасности распределенных баз данных. В процессе анализа выделены ключевые проблемы защиты информации в распределенных базах данных. Описаны подходы к защите информации, связанные с резервным копированием, аутентификацией, шифрованием и регулярными аудитами безопасности. Представлена авторская базовая схема и алгоритм обеспечения безопасности распределенной базы данных. Рассмотрены различные стратегии и подходы, позволяющие защитить распределенные базы данных от злонамеренных действий и вторжений. Представляется, что развитие технологий искусственного интеллекта, блокчейна, машинного обучения и других прорывных решений позволит повысить эффективность и упростит защиту распределенных баз данных, сделав их более надежными.

Ключевые слова: распределенная база, защита, шифрование, копия.

По мере роста бизнеса и расширения потребностей в хранении информации, выходящих за рамки традиционной единичной базы данных, компании нуждаются в более сложном варианте подходов и методов работы с широким набором данных. Традиционные хранилища, построенные на основе реляционных баз данных и способные оперировать только структурированными сведениями, стали испытывать трудности при работе с большими объемами информации. И в данном случае для удовлетворения возникшего спроса свое широкое применение нашли распределенные системы, которые позволяют получить все преимущества традиционного хранилища, но при этом обеспечивают сохранность больших объемов информации и делают их доступными по всей компьютерной сети [1].

Системы распределенных баз данных стали краеугольным камнем современных приложений, особенно с развитием облачных вычислений, больших данных и аналитики в реальном времени. Об их популярности свидетельствует тот факт, что глобальный рынок облачных баз данных оценивался в 15,08 млрд долларов США в 2022 году, и ожидается, что к 2032 году он превысит примерно 68,38 млрд долларов США, а среднегодовой темп роста составит 16,32 % с 2024 по 2032 год (см. рис. 1).

Распределенные базы данных прекрасно справляются со своей задачей, обеспечивая внутренние данные организационной структурой, расширяя возможности совместного использования, повышая доступность, открывая новые возможности для модульного роста и увеличивая скорость обработки. Кроме того, они предлагают множество преимуществ, таких как масштабируемость, отказоустойчивость и географическое распределение. Однако, ни одно решение для хранения данных не обходится без недостатков, и распределенные базы данных также сталкиваются с уникальным набором проблем. И эти проблемы связаны с их уязвимостью к сторонним атакам, с высокой подверженностью взлома и нарушению конфиденциальности данных.

В данном контексте, актуальность вопросов безопасности распределенной базы данных, развития методов защиты информации и системы, которая ими управляет, не подлежит сомнению, что и послужило основанием для выбора темы данной статьи.

Фундаментальным и прикладным исследованиям в области безопасности различных систем, сетей и баз данных посвятили свои труды Налисник А.Н., Белов Ю.С., Полтавцева М.А., Зегжда Д.П., Калинин М.О., Foad Jalali, Mehran Alidoost Nia, Tatiana Ermakova, Meisam Abdollahi, Benjamin Fabian.

Общие концепции и модели безопасности информации, способы обеспечения разрешенных доступов получили свое развитие в публикациях Сухобокова А.А., Тру-

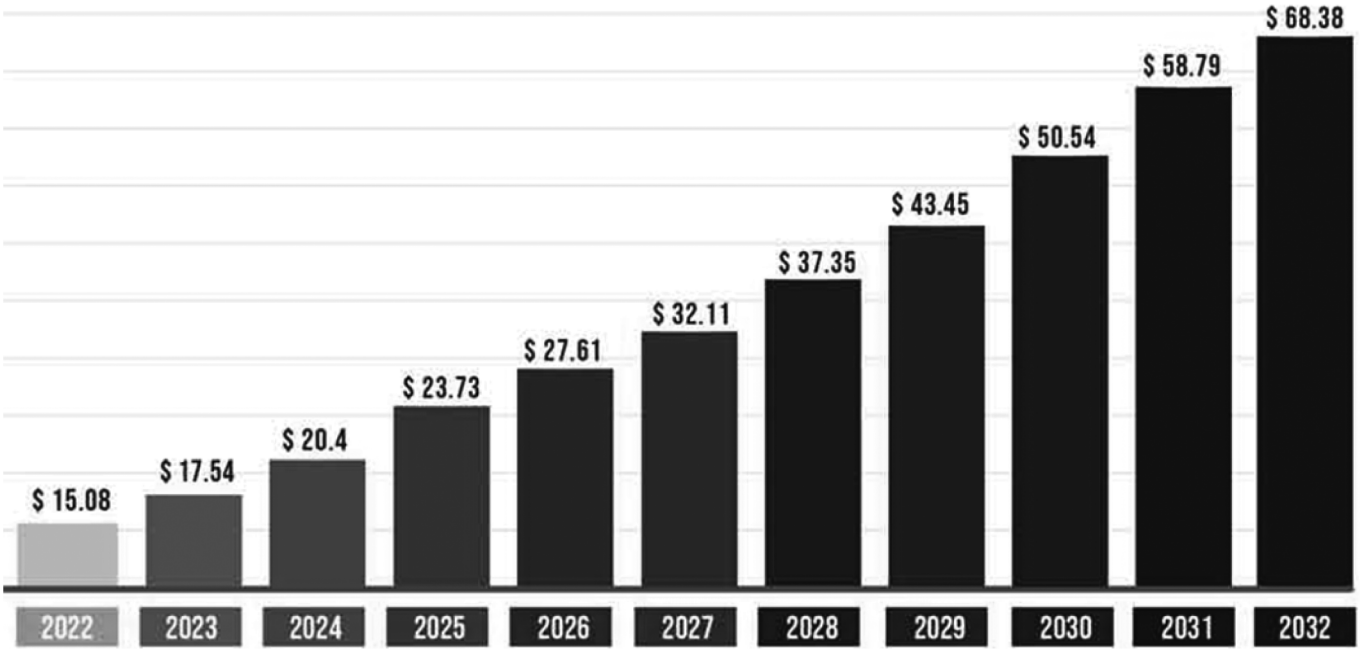

Рис. 1. Динамика глобального рынка облачных баз данных, млрд дол. [2]

фанова В.А., Столярова Ю.А., Садыкова М.Р., Елизарова О.О., Грозмани Е.С., Петрова С.В., Jafar A. Alzubi, Omar A. Alzubi, Ashish Singh, Tareq Mahmod Alzubi.

Однако, количество нарушений безопасности баз данных в настоящее время растет в геометрической прогрессии, а связанные с ними последствия становятся все более существенными. Это свидетельствует о том, что в данной предметной плоскости существует еще много нерешенных проблем, которые заслуживают отдельного внимания. Так, в более углубленной проработке нуждаются методы защиты архитектуры распределенных систем, которые могут включать межоблачные, гибридные и/или мультиоблачные сети. Кроме того, в отдельном обосновании нуждаются подходы защиты данных, которые могут использоваться к базам со специальными требованиями соответствия в регулируемых средах.

Таким образом, цель статьи заключается в изучении подходов и методов, которые используются в процессе обеспечения безопасности распределенных баз данных.

Распределенные системы баз данных состоят из нескольких баз данных, соединенных сетью, часто расположенных в разных физических местах. Вместо единого центрального хранилища данные распределяются между несколькими серверами. Для оптимизации производительности и надежности в системе могут использоваться различные стратегии, такие как чередование, разделение или репликация [3]. На фоне всех своих преимуществ и возможностей распределенные базы данных часто более подвержены рискам безопасности из-за своей природы. Несанкционированный доступ, утечка данных и другие слабые места могут поставить под угрозу всю систему. Кратко обозначим уязвимости распределенной базы данных, что позволит более четко выделить наиболее приемлемые подходы к ее защите.

1. Злоупотребление привилегиями базы данных — открытый, совместный характер распределенных баз данных может позволить неавторизованным пользователям получить доступ к данным, которые для них закрыты.
2. Угрозы, связанные с вредоносным программным обеспечением.
3. Физическая уязвимость — поскольку данные в распределенной системе хранятся на физических серверах, они подвержены тем же угрозам, что и любые физические устройства хранения: наводнение, пожар, удар молнии, другие физические повреждения, технические сбои и многое другое.
4. Обнаружение данных — чувствительные данные могут быть раскрыты при получении доступа неавторизованным пользователем, создании несанкционированных копий или краже устройства, содержащего эти данные.
5. Синхронизация данных — одно из самых больших преимуществ распределенной базы данных является также одним из ее самых слабых мест. Распределенные базы данных часто применяются для предоставления пользователям расширенного совместного доступа к данным на нескольких сайтах, а также для увеличения емкости хранилища [4]. Хотя эта функция повышает производительность и максимально увеличивает пространство для хранения, она также представляет собой один

из самых больших рисков нарушения целостности, конфиденциальности распределенной базы данных.

На сегодняшний день наработаны различные подходы к обеспечению безопасности распределенных баз данных.

Одним из самых популярных является надежное облачное решение для резервного копирования, которое позволяет:

— обеспечивать резервное копирование данных в неограниченном количестве баз данных;
— автоматически и непрерывно создавать резервные копии данных, чтобы учитывать каждый файл и его изменения;
— получить доступ и восстанавливать предыдущие версии файлов, чтобы отменить случайные изменения и удаления;
— реализовать упреждающую виртуализация серверов для полного устранения простоев, связанных с потерей данных;
— восстановить данные на том же устройстве или на новом в случае повреждения или потери устройства.

Помимо этого, перспективным подходом для защиты распределенной базы данных является использование надежных механизмов аутентификации, шифрования и регулярные аудиты безопасности, которые могут значительно повысить уровень надежности системы [5].

Аутентификация, контроль доступа, криптографические методы, системы на основе кворума, модели, основанные на доверии и т.д. — это многочисленные разработки, направленные на создание безопасных и доверенных распределенных систем. Обобщение различных аспектов и технологи безопасности в распределенных базах данных представлены в таблице 1.

Вышеизложенные методы и подходы, конечно, не являются исчерпывающим списком всего, что можно сделать и использовать для защиты распределенных баз данных.

В рамках проводимого исследования рассмотрим более детально практические аспекты обеспечения безопасности распределенных систем хранения информации. Для этого автором формализована соответствующая базовая схема и алгоритм.

На рис. 2 описаны запрос, обработка и ответ в распределенной базе данных.

Алгоритм обеспечения безопасности для распределенной базы данных приведен на рис. 3 в виде блок-схемы. Общие шаги алгоритма включают в себя:

Таблица 1.

Технологии обеспечения безопасности в распределенных базах данных (составлено автором)

№ п/п	Категория	Фокус
1	Подходы, основанные на аутентификации	Техника аутентификации пути
		Архитектура планирования с учетом требований безопасности
		Аутентификация удаленного клиента
		Пароли, цифровые сертификаты и конфиденциальность
		Криптография в серверах аутентификации
2	Безопасность на основе доверия	Управление рисками
		Система P2P
		Расширенная модель на основе теории D-S
		Контекстно-чувствительная модель доверия
3	Безопасность на основе политики	Модульные политики безопасности
4	Безопасность на основе шаблонов	Модель безопасности для распределенных систем
5	Безопасность на основе кворума	Распределенная система отказоустойчивости
6	Другие методы	Система на основе мобильных агентов
		Генетический алгоритм
		Веб-архитектура X-Tron
		Надежный массив независимых узлов (RAIN)
		Планировщик легальных информационных потоков

1. Пользователь через сервер приложений отправляет свой запрос на прокси-сервер.
2. Прокси-сервер шифрует полученный запрос.
3. Прокси-сервер отправляет зашифрованный запрос на сервер СУБД.
4. В базе данных выполняются SQL-операции над зашифрованными данными.
5. Сервер СУБД отправляет зашифрованный ответ на прокси-сервер.
6. Прокси-сервер расшифровывает полученный от базы данных ответ.
7. Прокси-сервер отправляет расшифрованный ответ на сервер приложений пользователя.

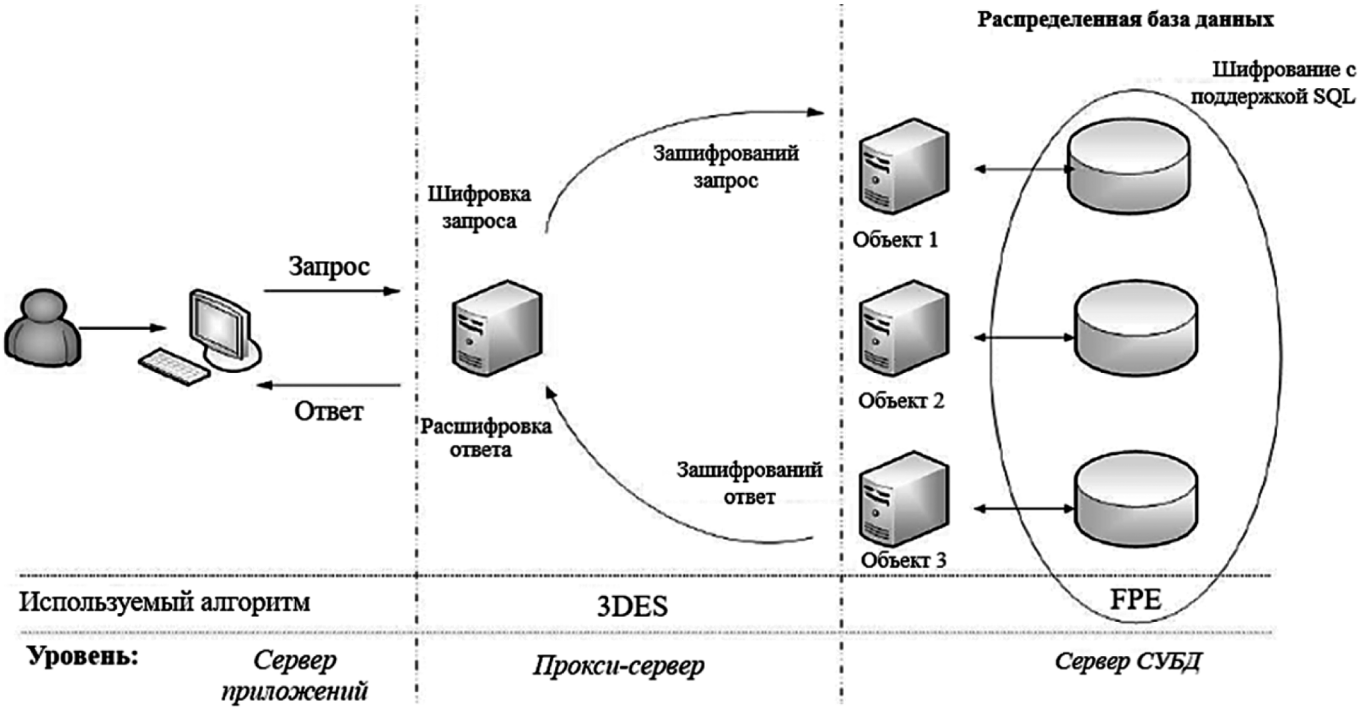

Рис. 2. Базовая схема безопасности для распределенной базы данных

Рис. 3. Алгоритм обеспечения безопасности для распределенной базы данных

На четвертом шаге алгоритма, приведенного на рис. 2, задачи SQL-операций выражаются следующим математическим базисом:

$$M\prod_{i=1}^{n}\prod_{j=1}^{m}\left\|\frac{Delete}{p,t}\right\|\left\|\frac{Insert}{\frac{A_i\,v_i}{t}}\right\|\left\|\frac{Select}{\frac{A_i\,t_j}{p}}\right\|\left\|\frac{Update}{\frac{A_i\,v}{p,t}}\right\|$$

где: M — количество распределяемых объектов, n — количество атрибутов для вставки или обновления, m — количество таблиц для выбора, A_i — атрибут таблицы, t_j — одна из таблиц для выбора, P — предикат или условие, t — таблица для удаления, вставки или обновления, v_i — значение для вставки или обновления, i не зависит от j.

На шестом шаге алгоритма, приведенного на рис. 2 только в том случае, если запрашиваемая операция была SELECT, получается набор данных; каждая возвращенная запись должна быть расшифрована следующим математическим образом:

$$\prod_{i=1}^{n}\sigma_i\ v_i\ =\ \sigma_1 v_1,\sigma_2 v_2,\ldots\sigma_n v_n$$
$$v\ \in\ D$$

где: σ — операция дешифрования, v — строка или вектор, D — набор данных или векторов, n — количество строк набора данных.

Если результат операции отличен от нуля, то получается набор данных в виде D_{nm}.

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

При операциях удаления, вставки или обновления база данных возвращает только количество затронутых строк или ноль.

Итак, в представленной базовой схеме принята трехуровневая архитектура. На уровне прокси-сервера для запросов и ответов к базе данных использован алгоритм 3DES, на уровне сервера СУБД — алгоритмы шифрования на основе SQL и FPE. Предложенные схема и алгоритм не зависят от относительной безопасности, предлагаемой распределенной базой данных. Предполагаемый контроль доступа не зависит от прав пользователя. В процессе реализации схемы и алгоритма следует учитывать эффективность при увеличении объема данных.

Подводя итоги проведенному анализу, можно отметить следующее. Системы распределенных баз данных являются неотъемлемой частью современных вычислений, однако они имеют ряд проблем с обеспечением сохранности и безопасности информации. В статье рассмотрены различные стратегии и подходы, позволяющие защитить распределенные базы данных от злонамеренных действий и вторжений. Также предложена авторская базовая схема и алгоритм обеспечения безопасности распределенной базы данных. Представляется, что развитие технологий искусственного интеллекта, блокчейна, машинного обучения и других прорывных решений позволит повысить эффективность и упростит защиту распределенных баз данных, сделав их более надежными.

ЛИТЕРАТУРА

1. Закирова Э.Ф., Павлов С.В., Трубин В.Д., Христодуло О.И. Детализация пространственной информации для обеспечения защищенности баз данных в распределенных информационных системах // Системная инженерия и информационные технологии. 2022. Т. 4. № 1 (8). С. 20–26.
2. Унгер А.Ю. Формальный язык описания транзакций в реляционных базах данных // Высокопроизводительные вычислительные системы и технологии. 2022. Т. 6. № 1. С. 101–106.
3. Пучков А.Ю., Соколов А.М., Широков С.С., Прокимнов Н.Н. Алгоритм выявления угроз информационной безопасности в распределенных мультисервисных сетях органов государственного управления // Прикладная информатика. 2023. Т. 18. № 2 (104). С. 85–102.
4. Hongliang Tian, Xiaonan Ge Research on distributed blockchain-based privacy-preserving and data security framework in IoT // IET Communications. 2020. Volume 14, Issue 13. P. 56–64.
5. Shitharth Selvarajan, Achyut Shankar A smart decentralized identifiable distributed ledger technology-based blockchain (DIDLT-BC) model for cloud-IoT security // Expert Systems. 2024. №56. P. 45–49.
6. Alireza Chamkoori, Serajdean Katebi Security and storage improvement in distributed cloud data centers by increasing reliability based on particle swarm optimization and artificial immune system algorithms // Concurrency and Computation: Practice and Experience. 2023. Volume 35, Issue 6. P. 129–134.
7. Полтавцева М.А., Зегжда Д.П., Калинин М.О. Многоуровневая концепция безопасности систем управления большими данными // Вопросы кибербезопасности. 2023. № 5 (57). С. 25–36.

Журнал «Современная наука: актуальные проблемы теории и практики»