СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ / SYSTEM ANALYSIS, MANAGEMENT AND PROCESSING OF INFORMATION

DOI: https://doi.org/10.60797/IRJ.2024.148.91

СТРАТЕГИЯ БАЛАНСИРОВКИ НАГРУЗКИ И УПРАВЛЕНИЕ РАЗНОРОДНЫМИ ХРАНИЛИЩАМИ ДАННЫХ С ВЫСОКОЙ СТЕПЕНЬЮ ПАРАЛЛЕЛИЗМА

Научная статья

Багутдинов Р.А.[1, *]
[1] ORCID : 0000-0002-9645-8707;
[1] Институт транспорта и сервиса, Сочи, Российская Федерация

* Корреспондирующий автор (rav379[at]mail.ru)

Аннотация

В современном мире наблюдается экспоненциальный рост объёма данных и высокий уровень параллельного доступа в мощных Интернет вещах (PIoT), что требует эффективных стратегий балансировки нагрузки для равномерного распределения сетевого трафика между несколькими серверами. Это необходимо для повышения общей скорости отклика и доступности системы. Целью данного исследования является разработка стратегии балансировки нагрузки, основанной на четырёх и семи уровнях доступа IoT. Для достижения поставленной цели были использованы следующие материалы и методы: анализ существующих стратегий балансировки нагрузки; изучение протоколов транспортного и прикладного уровней; разработка новой стратегии балансировки нагрузки. В результате исследования была предложена новая стратегия балансировки нагрузки, основанная на сочетании четырёх- и семиуровневых подходов. Это позволит добиться более эффективного и гибкого распределения трафика в IoT, повысив производительность системы и удовлетворив требования к обработке сложных запросов.

Ключевые слова: Интернет вещей, IoT, разнородные, хранилища, источники, большие, управление, данные, метаданные.

LOAD BALANCING STRATEGY AND MANAGEMENT OF HETEROGENEOUS DATA STORAGES WITH A HIGH DEGREE OF PARALLELISM

Research article

Bagutdinov R.A.[1, *]
[1] ORCID : 0000-0002-9645-8707;
[1] Institute of Transport and Service, Sochi, Russian Federation

* Corresponding author (rav379[at]mail.ru)

Abstract

In today's world, there is an exponential growth of data volume and a high level of concurrent access in the powerful Internet of Things (PIoT), which requires effective load balancing strategies to evenly distribute network traffic among multiple servers. This is necessary to improve the overall response rate and availability of the system. The objective of this study is to develop a load balancing strategy based on four and seven levels of IoT access. To achieve this goal, the following materials and methods were used: analysis of existing load balancing strategies; study of transport and application layer protocols; development of a new load balancing strategy. As a result of the study, a new load balancing strategy was proposed based on a combination of four- and seven-level approaches. This will achieve more efficient and flexible traffic distribution in the IoT, increasing system performance and meeting the requirements for processing complex queries.

Keywords: Internet of Things, IoT, heterogeneous, storage, sources, large, management, data, metadata.

Введение

В мощном Интернете вещей (PIoT) используются сотни миллионов терминальных устройств. Даже если данные загружаются пакетами, в часы пик они будут достигать более десяти миллионов уровней параллелизма. Технология доступа с высокой степенью параллелизма основана на балансировке нагрузки, оптимизации трафика и других стратегиях для достижения оптимального выбора сетевых соединений и эффективного, сбалансированного использования облачных ресурсов [1]. Балансировка нагрузки необходима в системах с высокой степенью параллелизма и легкодоступности [2]. Цель состоит в том, чтобы равномерно распределить сетевой трафик между несколькими серверами, чтобы повысить общую скорость отклика и доступность системы. Чтобы справиться с экспоненциальным ростом объема данных и высоким уровнем параллельного доступа, крупные центры обработки данных должны развёртывать модули балансировки нагрузки для обработки больших внешних или внутренних рабочих нагрузок и улучшения использования ресурсов [3].

В настоящее время стратегия балансировки нагрузки платформы управления IoT обычно заключается в развёртывании программного обеспечения для балансировки нагрузки, такого как NGINX (HTTP-сервер и обратный прокси-сервер, почтовый прокси-сервер, а также TCP/UDP прокси-сервер общего назначения. Nginx обслуживает серверы многих высоконагруженных российских сайтов, таких как Яндекс, Mail.Ru, ВКонтакте и Рамблер) [4], на сервере доступа для достижения балансировки нагрузки на нескольких уровнях, а именно на уровне приложений. Этот метод может удовлетворить требования к параллельному доступу от 100000 до одного миллиона уровней; однако трудно поддерживать высокие требования к параллельному доступу более чем на 10 миллионах уровней.

Таким образом, предлагается стратегия балансировки нагрузки, основанная на четырёх и семи уровнях доступа IoT, для достижения высокого уровня параллелизма – более десяти миллионов уровней [5]. Четырёхуровневая балансировка нагрузки основана на протоколах транспортного уровня, таких как TCP и UDP. Он распределяет трафик между различными серверами на основе IP-адреса клиента и номера порта. Его основным преимуществом является высокая скорость обработки, поскольку он фокусируется только на сетевой информации более низкого уровня. Четырёхуровневая балансировка нагрузки подходит для сценариев со строгими требованиями к задержке.

Семиуровневая балансировка нагрузки основана на протоколах прикладного уровня, таких как HTTP и HTTPS. Она может распределять трафик между различными серверами на основе содержимого запроса, такого как URL-адреса, информация заголовка и самих сообщений. Поскольку он фокусируется на сетевой информации более высокого уровня, семиуровневая балансировка нагрузки позволяет реализовать более сложные стратегии распределения. Сочетание четырёх- и семиуровневых стратегий балансировки нагрузки позволяет добиться более эффективного и гибкого распределения трафика в IoT. Например, четырёхуровневый балансировщик нагрузки может обрабатывать несколько запросов с низкой задержкой, тогда как сложные запросы прикладного уровня могут направляться для обработки в семиуровневый балансировщик нагрузки. Это повышает производительность системы и удовлетворяет требованиям к обработке сложных запросов. В частности, коммутатор уровня 4 предоставляет единый IP-адрес доступа для внешнего мира. Терминальному устройству и пограничному IoT-агенту не нужно знать реальный IP-адрес, соответствующий каждому серверу на облачной платформе. Трафик внешних данных, получаемый облачной платформой, должен проходить через коммутатор уровня 4, который отвечает за пересылку запроса терминального устройства и пограничного агента Интернета вещей на сервер, а затем за установление TCP-соединения между терминальным устройством, пограничным агентом интернета вещей и сервером. В режиме NAT (от англ. Network Address Translation – «преобразование сетевых адресов») – это механизм в сетях TCP/IP, позволяющий преобразовывать IP-адреса транзитных пакетов [6], когда уровень 4 обменивается данными и устройство планирует запросы на доступ, сначала преобразуется IP-адрес назначения, а затем запрос на доступ пересылается на каждый сервер внешнего доступа в серверной части. Таким образом, четырёхуровневая балансировка нагрузки позволяет объединить терминал доступа и оборудование пограничного агента Интернета вещей и передавать исходящую информацию об оборудовании программно-определяемому агенту доступа в соответствии с установленной стратегией балансировки нагрузки.

Основные результаты

2.1. Управление разнородными хранилищами данных с несколькими источниками

Объём данных мощных IoT (PIoT) достиг петабайтного масштаба, что означает, что система должна справляться с проблемами, связанными с большим объемом данных и различными форматами данных. Диапазон скоростей генерации данных обширен, включая миллисекундные измерения вектора в широком диапазоне, данные в режиме реального времени, данные мониторинга стационарного состояния второго уровня, микрометеорологические данные минутного уровня [7], данные о циркуляции эксплуатационных характеристик часового уровня и экспериментальные данные оборудования с более длительным циклом [8]. Существует множество источников данных и сложных методов взаимодействия, таких как веб-сервисы, специальные протоколы и специальные форматы файлов [9]. Существует много типов данных, таких как данные в реальном времени, текстовые, мультимедийные, временные ряды, структурированные, полуструктурированные и неструктурированные данные и др. [10], [11]. Отсутствие эффективного управления различными типами данных и трудности с формированием эффективной информации из изолированных данных значительно затрудняют управление, эксплуатацию и принятие быстрых решений [12].

В области хранения больших данных команда разработчиков Hadoop разработала файловую систему Google с открытым исходным кодом (GFS) и внедрила распределенную файловую систему Hadoop с открытым исходным кодом (HDFS) GFS. Программная библиотека Apache Hadoop – это платформа, которая позволяет распределять обработку больших наборов данных между кластерами компьютеров с использованием простых программных моделей [13]. Что касается управления данными, то наиболее известной является технология управления данными Bigtable, предложенная Google. Bigtable – это распределенная система хранения, предназначенная для управления структурированными данными [14], [15]. Эти данные могут быть расширены до очень больших масштабов, таких как петабайты данных на тысячах коммерческих серверов. Основными источниками данных PIoT являются данные управления, данные мониторинга, социальные сети и метеорологические данные, которые характеризуются различными источниками, сложными типами, различными мощностями и высокими уровнями управления безопасностью и контроля [16], [17]. Для управления большими и разнообразными данными и метаданными нужна надёжная техническая архитектура, которая обеспечит контроль над данными, их качеством, структурой, безопасностью [18], а также эффективное управление ими.

2.2. Управление основными данными и метаданными

Основные данные в основном разрабатываются внутри систем и извлекаются из неструктурированных данных, которые были проанализированы или не исследованы. Таким образом, необходимо обеспечить плавную интеграцию метаданных неструктурированного контента с традиционным управлением метаданными. Основное содержание управления основными данными и метаданными проиллюстрировано на рис. 1.

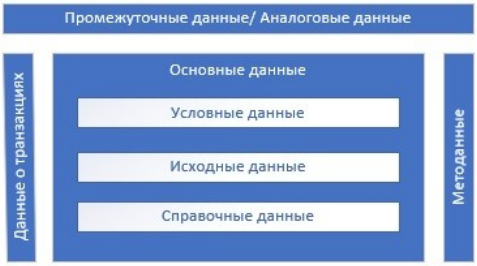

Рисунок 1 - Распределение контента для управления основными данными и метаданными
DOI: https://doi.org/10.60797/IRJ.2024.148.91.1

Архитектуры управления метаданными и основными данными нуждаются в корректировке в связи с внедрением технологии больших данных. Для предлагаемой сводной архитектуры управления основными данными и метаданными в среде больших данных PIoT приложения для сбора, хранения и управления метаданными и основными данными основаны на архитектуре данных, технической архитектуре и архитектуре приложений больших данных. Поток данных представлен на рис. 2.

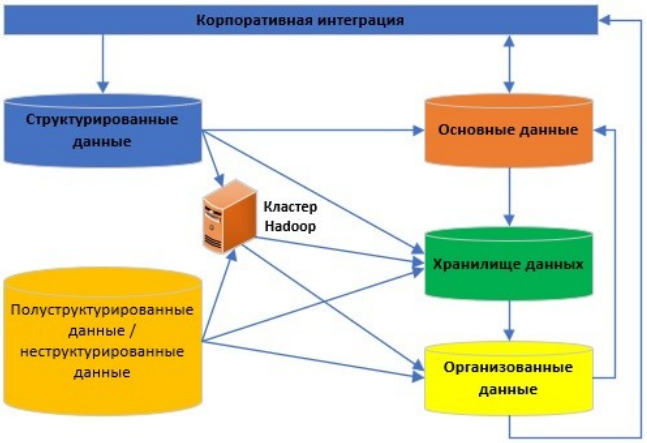

Рисунок 2 - Организация потока данных
DOI: https://doi.org/10.60797/IRJ.2024.148.91.2

Управление жизненным циклом данных относится к управлению информационными активами. Управление охватывает правила, политики, процессы, роли и обязанности, которые используются для руководства общим управлением, данные являются точными, непротиворечивыми, полными, доступными и безопасными. Стратегия и стандарты обработки данных должны быть обновлены в связи с внедрением неструктурированных и разнообразных источников данных. На всех этапах жизненного цикла данных данные должны храниться, защищаться и получать доступ к ним в полном соответствии с бизнес-требованиями. Данные являются активами предприятия, и их жизненные циклы часто длиннее, чем у аппаратного обеспечения и приложений. Управление жизненным циклом данных должно осуществляться с точки зрения контроля затрат, управления информацией и отслеживания, а также безопасности данных.

Качество данных относится к способности данных удовлетворять бизнес, системным и техническим требованиям организации. Качество данных обычно описывается в соответствии с целостностью, своевременностью, точностью, непротиворечивостью и актуальностью. Качество данных в средах больших данных определяется случаями использования. Приложения предъявляют разные требования к качеству данных. Например, анализ потока кликов и обнаружение вторжений требуют разных уровней точности. В этой среде необходимо пересмотреть правила, политики и стандарты очистки качества данных.

Структура данных описывает внутреннюю организацию данных. Она включает в себя множество уровней данных, начиная от междисциплинарных моделей данных и заканчивая независимыми системами. Влияние структуры данных в среде больших данных в первую очередь обусловлено разнообразием данных. Стандарты классификации данных и логические модели данных должны адаптироваться к содержанию неструктурированных данных и процессу извлечения структурированной информации из неструктурированных данных.

Безопасность данных – это процесс и технология, которые гарантируют, что к данным нельзя получить доступ, просмотреть, отредактировать или удалить их без разрешения. При интеграции различных источников данных необходимо проанализировать и стандартизировать соответствующие политики безопасности и требования в соответствии с местными, национальными и международными правилами. Это включает в себя:

1 – безопасность доступа к данным: авторизацию доступа к хост-системе, хранение базы данных на разных уровнях безопасности и разделение пользователей для различных приложений;

2 – безопасность хранения данных: регулярное полное резервное копирование и инкрементное резервное копирование в режиме реального времени, которые могут поддерживать быстрое восстановление данных;

3 – аудит безопасности данных: записи журнала доступа к файлам, сети и другим данным, независимо хранящиеся журналы и регулярно проверяемые операции с конфиденциальными данными.

Управление службами данных обеспечивает унифицированный доступ к различным службам данных через сервисные интерфейсы. Благодаря внедрению стандартизированного протокола доступа к интерфейсу и унифицированного контроля доступа пользователей и данных для обеспечения безопасности данных, сервис может быть расширен и настроен в соответствии с будущими требованиями к доступу к данным.

Информация, хранящаяся в PIoT, является массивной, распределенной, разнообразной, действующей в режиме реального времени, динамичной и интерактивной [19]. Границы сети на стороне терминала станут нечёткими и сложными. Индивидуальный доступ между данными и пользователями в рамках обычной схемы шифрования с открытым ключом не может соответствовать требованиям сложной системы PIoT. Механизм нулевого доверия был впервые предложен Kindervag. Его основная идея заключается в том, что по умолчанию не следует доверять человеку, устройству или системе внутри сети или за её пределами. Она должна основываться на доверительной основе аутентификации и авторизации с жёстким контролем доступа. После многолетней практики Google в 2014 году выпустила BeyondCorp, архитектуру с нулевым доверием, которая постепенно получила признание в отрасли. Нулевое доверие подрывает парадигму контроля доступа и направляет архитектуру безопасности от сети к идентификации, ориентированной на личность. Его основным требованием является управление доступом, ориентированное на личность [20].

Построение PIoT основано на построении унифицированной идентификации с использованием архитектуры сетевой безопасности с нулевым доверием в качестве эталона для проведения унифицированного управления идентификацией и обеспечения аутентификации между устройствами и службами PIoT. Надежное соединение, взаимодействие в области безопасности, интеллектуальная защита, а также динамическое предотвращение и контроль достигаются благодаря защите архитектуры PIoT «облачный интерфейс управления». Архитектура защиты PIoT, основанная на механизме нулевого доверия, проиллюстрирована на рис. 3.

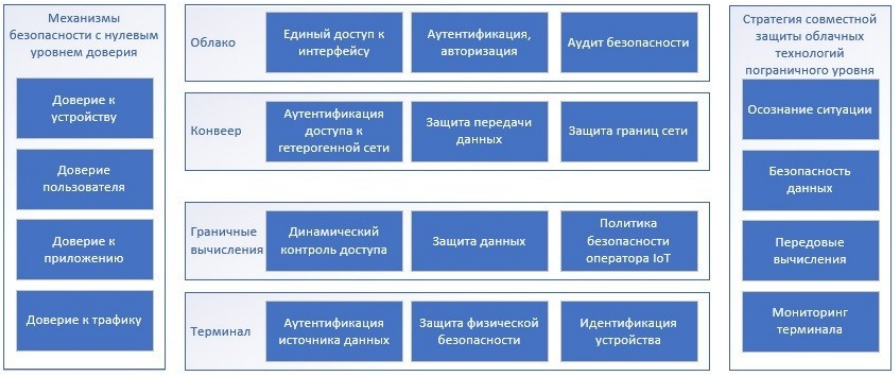

Рисунок 3 - Поток доступа к данным PIoT основан на многоуровневой балансировке нагрузки
DOI: https://doi.org/10.60797/IRJ.2024.148.91.3

Создание единой идентификационной библиотеки PIoT в конце и на стороне обеспечивает базовую идентификацию ключей для бизнеса и обеспечивает стандартизацию субъектов бизнес-системы. После реализации комплексной идентификации пользователь, устройство и приложение полностью идентифицируются для завершения централизованного управления идентификацией. Аутентификация осуществляется непрерывно через службу списка устройств. Благодаря динамическому расчёту риска и доверия доступ к приложениям блокируется и проверяется, состояние терминала PIoT постоянно отслеживается, и корректируется соответствующая стратегия контроля доступа. Доступ приложения к центру политик постоянно контролируется, а выходные данные журнала доступа отправляются на аналитическую платформу для оценки рисков. Анализ больших данных и технологии искусственного интеллекта используются для анализа рисков, поддержки количественной оценки рисков, осуществления мониторинга и аудита доступа приложения к центру стратегии, реализации динамического восприятия и интеллектуального анализа сценариев безопасности PIoT, своевременного реагирования на атаки и обеспечения безопасной и стабильной работы системы PIoT.

Заключение

В заключении можно отметить, что предложенная стратегия балансировки нагрузки, основанная на сочетании четырёх- и семиуровневых подходов, позволяет эффективно и гибко распределять трафик в IoT, повышая производительность системы и удовлетворяя требования к обработке сложных запросов. Это особенно актуально в условиях экспоненциального роста объёма данных и высокого уровня параллельного доступа в мощный Интернет вещей (PIoT), где требуется балансировка нагрузки для равномерного распределения сетевого трафика между несколькими серверами. С увеличением объёма данных и разнообразием их типов, становится критически важным наличие эффективных инструментов и методов для управления этими данными.

Однако, несмотря на существующие технологии, управление данными в PIoT остается сложной задачей из-за разнообразия источников данных, их сложности и высоких требований к безопасности и контролю. Для успешного

управления данными необходима надёжная техническая архитектура, обеспечивающая контроль над данными, их качеством, структурой, безопасностью, а также эффективное управление ими.

Исследование подчёркивает важность разработки и внедрения новых технологий и методов балансировки нагрузки для обеспечения стабильной и эффективной работы систем IoT в условиях растущих требований к объёму обрабатываемых данных и уровню параллелизма.

Конфликт интересов

Не указан.

Рецензия

Сообщество рецензентов Международного научно-исследовательского журнала
DOI: https://doi.org/10.60797/IRJ.2024.148.91.4

Conflict of Interest

None declared.

Review

International Research Journal Reviewers Community
DOI: https://doi.org/10.60797/IRJ.2024.148.91.4

Список литературы / References

1. Gao Z. B. Improved load balancing algorithm based on weighted least-connections / Z.B. Gao, Y.C. Pan, Z. Hua [et al.] // Science Technology and Engineering. — 2016. — Vol. 16. — № 6. — P. 81–85.

2. Zhang C.K. State-of-the-Art survey on software-defined networking (SDN) / C.K. Zhang, Y. Cui, H.H.Tang [et al.] // Ruan Jian Xue Bao Journal of Software. — 2015. — Vol. 26. — №1. — P. 62–81.

3. Irteza S.M. Efficient load balancing over asymmetric datacenter topologies / S.M. Irteza, H.M.Bashir, T. Anwar [et al.] // Computer Communications. — 2018. — Vol. 127. — P. 1–12.

4. Сысоев И.В. HTTP-сервер и обратный прокси-сервер, почтовый прокси-сервер, а также TCP/UDP прокси-сервер общего назначения / И.В. Сысоев // Nginx. — 2022 — URL: https://nginx.org/ru/ (дата обращения: 31.08.2024)

5. Itsekson A. What Are the 7 Layers of IoT Architecture? / A. Itsekson. — 2024 — URL: https://jelvix.com/blog/iot-architecture-layers (accessed: 30.04.2024)

6. What Is Network Address Translation (NAT)? — 2024. — URL: https://www.cisco.com/c/en/us/products/routers/network-address-translation.html (accessed: 31.07.2024)

7. Багутдинов Р.А. Проектирование модульной мультисенсорной системы для задач мониторинга окружающей среды на базе Arduino / Р.А. Багутдинов // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. — 2019. — Т. 46. — № 1. — С. 173–180.

8. Jiang X.T. Research and design on general communication protocol analysis system for power information collecting / X.T. Jiang, J.B. He, N. Li // Power System Protection and Control. — 2012. — Vol. — 40. — № 9. — P. 118–122.

9. Багутдинов Р.А. Разработка единой централизованной системы управления транспортными потоками / Р.А. Багутдинов, Д.В. Бежуашвили // Техник транспорта: образование и практика. — 2021. — Т. 2. — № 1. — С. 71–77.

10. Багутдинов Р.А. Алгоритм обнаружения пожара для мультисенсорной системы / Р.А. Багутдинов, М.Ф. Степанов // Вестник Дагестанского государственного технического университета. Технические науки. — 2021. — Т. 48. — № 3. — С. 59–67.

11. Багутдинов Р.А. Методы интеграции, уменьшение размеров и нормализация обработки разнородных и разномасштабных данных / Р.А. Багутдинов, М.Ф. Степанов // International Journal of Open Information Technologies. — 2021. — Т. 9. — № 2. — С. 39–44.

12. Pan K.J. Research on multi-source heterogeneous data fusion technology for the big data of power / K.J. Pan, X. Wang, F. Yang [et al.] // Machinery & Electronics. — 2017. — Vol. 35. — № 9. — P. 7–11.

13. Apache Hadoop Software Library. — 2024. — URL: https://hadoop.apache.org (accessed: 28.08.2024)

14. Scale your latency-sensitive applications with the NoSQL pioneer. — 2024. — URL: https://cloud.google.com/bigtable (accessed: 20.08.2024)

15. Apache Hadoop Programming Mode MapReduce Tutorial. — 2024. — URL: https://hadoop.apache.org/docs/current/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html (accessed: 20.08.2024)

16. Островский О.А. Вопросы развития компьютерно-технической экспертизы и их взаимосвязь с телекоммуникациями / О.А. Островский // Право и государство: теория и практика. — 2020. — № 1 (181). — С. 312–314.

17. Островский О.А. Меры уголовного правосудия в отношении киберпреступности согласно модели общего права через призму профессиональных компетенций следователя и его навыков интеллектуального анализа криминалистической цифровой информации / О.А. Островский, Ю.П. Гармаев // Транспортное право и безопасность. — 2020. — № 4 (36). — С. 84–92.

18. Островский О.А. Значение цифровых доказательств при расследовании уголовных преступлений / О.А. Островский // Вестник Российского университета дружбы народов. Серия: Политология. — 2019. — № 1. — С. 123.

19. Zhou F. Development ideas of key technologies for intelligent perception of ubiquitous power Internet of things [J] / F. Zhou, H. Zhou, Y. Gao // Proceedings of the Chinese Electrical Engineering. — 2020. — Vol. 40. — № 1. — P. 70–82.

20. Zhang Y.Y. Electric Internet of Things security framework and technologies for energy interconnection / Y.Y. Zhang, B.X. Zhoum H.Y. Pang [et al.] // Telecommunications Science. — 2021. — Vol. 37. — № 2. — P. 115–124.

Список литературы на английском языке / References in English

1. Gao Z. B. Improved load balancing algorithm based on weighted least-connections / Z.B. Gao, Y.C. Pan, Z. Hua [et al.] // Science Technology and Engineering. — 2016. — Vol. 16. — № 6. — P. 81–85.

2. Zhang C.K. State-of-the-Art survey on software-defined networking (SDN) / C.K. Zhang, Y. Cui, H.H.Tang [et al.] // Ruan Jian Xue Bao Journal of Software. — 2015. — Vol. 26. — №1. — P. 62–81.

3. Irteza S.M. Efficient load balancing over asymmetric datacenter topologies / S.M. Irteza, H.M.Bashir, T. Anwar [et al.] // Computer Communications. — 2018. — Vol. 127. — P. 1–12.

4. Sysoev I.V. HTTP-server i obratnyj proksi-server, pochtovyj proksi-server, a takzhe TCP/UDP proksi-server obschego naznachenija [HTTP server and reverse proxy server, mail proxy server, and general purpose TCP/UDP proxy server] / I.V. Sysoev // Nginx. — 2022 — URL: https://nginx.org/ru/ (accessed: 31.08.2024) [in Russian]

5. Itsekson A. What Are the 7 Layers of IoT Architecture? / A. Itsekson. — 2024 — URL: https://jelvix.com/blog/iot-architecture-layers (accessed: 30.04.2024)

6. What Is Network Address Translation (NAT)? — 2024. — URL: https://www.cisco.com/c/en/us/products/routers/network-address-translation.html (accessed: 31.07.2024)

7. Bagutdinov R.A. Proektirovanie modul'noj mul'tisensornoj sistemy dlja zadach monitoringa okruzhajuschej sredy na baze Arduino [Designing a modular multisensory system for environmental monitoring tasks based on Arduino] / R.A. Bagutdinov // Scientific Bulletin of Belgorod State University. Series: Economics. Computer science. — 2019. — Vol. 46 — № 1. — P. 173–180. [in Russian]

8. Jiang X.T. Research and design on general communication protocol analysis system for power information collecting / X.T. Jiang, J.B. He, N. Li // Power System Protection and Control. — 2012. — Vol. — 40. — № 9. — P. 118–122.

9. Bagutdinov R.A. Razrabotka edinoj tsentralizovannoj sistemy upravlenija transportnymi potokami [Development of a unified centralized traffic management system] / R.A. Bagutdinov, D.V. Bezhuashvili // Transport technician: education and practice. — 2021. — Vol. 2. — № 1. — P. 71–77. [in Russian]

10. Bagutdinov R.A. Algoritm obnaruzhenija pozhara dlja mul'tisensornoj sistemy [Fire detection algorithm for a multisensory system] / R.A. Bagutdinov, M.F. Stepanov // Bulletin of Dagestan State Technical University. Technical sciences. — 2021. — Vol. 48. — № 3. — P. 59–67. [in Russian]

11. Bagutdinov R.A. Metody integratsii, umen'shenie razmerov i normalizatsija obrabotki raznorodnyh i raznomasshtabnyh dannyh [Methods of integration, size reduction and normalization of processing of heterogeneous and multi-scale data] / R.A. Bagutdinov, M.F. Stepanov // International Journal of Open Information Technologies. — 2021. — Vol. 9 — № 2. — P. 39–44. [in Russian]

12. Pan K.J. Research on multi-source heterogeneous data fusion technology for the big data of power / K.J. Pan, X. Wang, F. Yang [et al.] // Machinery & Electronics. — 2017. — Vol. 35. — № 9. — P. 7–11.

13. Apache Hadoop Software Library. — 2024. — URL: https://hadoop.apache.org (accessed: 28.08.2024)

14. Scale your latency-sensitive applications with the NoSQL pioneer. — 2024. — URL: https://cloud.google.com/bigtable (accessed: 20.08.2024)

15. Apache Hadoop Programming Mode MapReduce Tutorial. — 2024. — URL: https://hadoop.apache.org/docs/current/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html (accessed: 20.08.2024)

16. Ostrovskij O.A. Voprosy razvitija komp'juterno-tehnicheskoj ekspertizy i ih vzaimosvjaz' s telekommunikatsijami [Issues of development of computer-technical expertise and their relationship with telecommunications] / O.A. Ostrovskij // Law and the state: theory and practice. — 2020. — № 1 (181). — P. 312–314. [in Russian]

17. Ostrovskij O.A. Mery ugolovnogo pravosudija v otnoshenii kiberprestupnosti soglasno modeli obschego prava cherez prizmu professional'nyh kompetentsij sledovatelja i ego navykov intellektual'nogo analiza kriminalisticheskoj tsifrovoj informatsii [Criminal justice measures in relation to cybercrime according to the common law model through the prism of the investigator's professional competencies and his skills in the intellectual analysis of criminalistic digital information] / O.A. Ostrovskij, Ju.P. Garmaev // Transport Law and Security. — 2020. — № 4 (36). — P. 84–92. [in Russian]

18. Ostrovskij O.A. Znachenie tsifrovyh dokazatel'stv pri rassledovanii ugolovnyh prestuplenij [The importance of digital evidence in the investigation of criminal offenses] / O.A. Ostrovskij // Bulletin of the Peoples' Friendship University of Russia. Series: Political Science. — 2019. — № 1. — P. 123. [in Russian]

19. Zhou F. Development ideas of key technologies for intelligent perception of ubiquitous power Internet of things [J] / F. Zhou, H. Zhou, Y. Gao // Proceedings of the Chinese Electrical Engineering. — 2020. — Vol. 40. — № 1. — P. 70–82.

20. Zhang Y.Y. Electric Internet of Things security framework and technologies for energy interconnection / Y.Y. Zhang, B.X. Zhoum H.Y. Pang [et al.] // Telecommunications Science. — 2021. — Vol. 37. — № 2. — P. 115–124.