

УДК 658.5

СОЗДАНИЕ ШАБЛОНА ИНТЕГРАЦИИ И ОБРАБОТКИ БИЗНЕС-СОБЫТИЙ НА ПРОИЗВОДСТВЕ

Розанова Анна Вячеславовна, магистрант, Базовая кафедра «Аналитика больших данных и методы видеоанализа», Уральский федеральный университет имени первого Президента России Б.Н. Ельцина

Достовалов Андрей Андреевич, магистрант, Базовая кафедра «Аналитика больших данных и методы видеоанализа», Уральский федеральный университет имени первого Президента России Б.Н. Ельцина

Капустин Савелий Андреевич, магистрант, Базовая кафедра «Аналитика больших данных и методы видеоанализа», Уральский федеральный университет имени первого Президента России Б.Н. Ельцина

Аннотация

В условиях глобальной цифровизации предприятия все чаще обращаются к современным платформам для интеграции и управления данными. Одним из перспективных решений является комбинация Apache NiFi и DATA-TRACK для создания единой системы управления данными. Статья посвящена исследованию возможностей этих платформ и их совместного применения в производственной среде. Цифровизация производства становится всё более актуальной, позволяя повысить эффективность и качество за счет внедрения информационных технологий, таких как IoT, большие данные и AI. В исследовании были сравнены платформы Apache Kafka, Apache NiFi, Apache Flink, AWS Kinesis и RabbitMQ, и DATA-TRACK (DATA-CI), и в качестве наиболее перспективной выбрана DATA-TRACK благодаря её способности интегрироваться с промышленными системами и обеспечивать обработку данных в реальном времени.

КЛЮЧЕВЫЕ СЛОВА: цифровизация производства, Continuous Intelligence, автоматизация, Apache NiFi, DATA-TRACK, интеграция данных, автоматизация потоков данных, мониторинг данных.

CREATING A TEMPLATE FOR INTEGRATION AND PROCESSING OF BUSINESS EVENTS IN PRODUCTION

Rozanova Anna Vyacheslavovna, Master's student, Department of Big Data Analytics and Video Analysis Methods, Ural Federal University named after the first President of Russia B.N. Yeltsin

Dostovalov Andrey Andreevich, Master's student, Department of Big Data Analytics and Video Analysis Methods, Ural Federal University named after the first President of Russia B.N. Yeltsin

Kapustin Savely Andreevich, Master's student, Department of Big Data Analytics and Video Analysis Methods, Ural Federal University named after the first President of Russia B.N. Yeltsin

Abstract

In the context of global digitalization, enterprises are increasingly turning to modern platforms for data integration and management. One of the promising areas is the integration of Apache NiFi and Data-track, which provide many services to create a unified data management system. This article is devoted to the study of the capabilities of these platforms and their joint application in a production environment. Digitalization of production is becoming increasingly relevant, allowing for increased efficiency and quality through the introduction of information technologies such as the Internet of Things, big data and AI. As a result, powerful platforms Apache Kafka, Apache NiFi, Apache Flink, AWS Kinesis and RabbitMQ were created, and a DATA tracking system (DATA-CI) was selected that allows data integration with other applications. systems and monitor user activity in real time.

KEYWORDS: Data-track, digitalization of production, Continuous Intelligence, automation, Apache NiFi, Data-track, data integration, automation of data flows, data monitoring.

Введение

Цифровизация производства представляет собой актуальную и значимую тему в современном мире, где информационные и коммуникационные технологии играют ключевую роль в трансформации экономических и производственных процессов. Актуальность этой темы обусловлена глобальными изменениями в экономике, повышением производительности, развитием инноваций и новых бизнес-моделей, а также государственной поддержкой стратегий, направленных на развитие цифровой экономики.

Целью исследования является изучение возможности интеграции платформ Apache NiFi и DATA-TRACK для оптимизации обработки данных в производственных процессах. Гипотеза заключается в том, что комбинирование этих платформ позволит значительно повысить эффективность и качество производства за счет автоматизации, централизованного управления потоками данных и улучшенного мониторинга. Новизна исследования состоит в разработке и анализе комплексного решения, объединяющего преимущества обеих платформ, что до настоящего времени не было реализовано.

Цифровизация, как процесс интеграции цифровых технологий в различные сферы деятельности, включает использование информационно-коммуникационных технологий для сбора, обработки, хранения и передачи данных, что позволяет повысить эффективность и качество процессов [1]. Влияние цифровизации на производство проявляется через автоматизацию и роботизацию, использование интернета вещей (IoT), анализ больших данных и аналитику, облачные технологии и искусственный интеллект (AI). Эти технологии способствуют оптимизации производственных процессов, снижению издержек, повышению качества продукции и созданию новых возможностей для инноваций и развития. Цифровизация в деловом мире ведет к технологическому сдвигу за счет оптимизации и повышения точности и эффективности работы. Цифровизация бизнес-процессов — это перевод предприятий на электронные платформы [2].

Основными преимуществами цифровизации являются:

- 1) экономия времени и повышение производительности;
- 2) оптимизация и совершенствование коммуникаций - как внутренних, так и внешних;
- 3) возможности перекрестных продаж - выход на новый уровень обслуживания клиентов и стимулирование их к приобретению большего количества продуктов;
- 4) конкурентные возможности за счет улучшения качества обслуживания клиентов и общей оптимизации рабочего процесса.

Цифровизация производственных процессов стала неотъемлемой частью стратегии развития современных предприятий, обеспечивая повышение эффективности, снижение издержек и улучшение качества продукции. Однако с внедрением цифровых технологий возникает критически важная задача — выбор подходящей платформы для обработки данных. В условиях, когда объемы данных растут экспоненциально, а требования к их обработке становятся всё более сложными, правильный выбор платформы становится ключевым фактором успеха.

В условиях конкурентной среды современные компании всё чаще управляют своими процессами ориентированным образом, чтобы быстро реагировать на изменения рынка. Управление бизнес-процессами объединяет концепции, методы и техники для поддержки проектирования, администрирования, конфигурации, выполнения и анализа бизнес-процессов [3]. Важнейшим элементом BPM (Business Process Management) являются процессные модели, которые явно описывают операции, необходимые для достижения бизнес-целей компании, и используются для документирования, сертификации и выполнения [4]. Для улучшения этих моделей необходимо мониторить и анализировать выполнение процессов, чтобы выявлять слабые места и находить потенциальные улучшения. В условиях современной цифровой трансформации и растущей сложности производственных процессов, предприятия сталкиваются с необходимостью интеграции различных информационных систем для оптимизации работы [5]. Однако этот процесс сопряжен с рядом проблем, таких как несовместимость систем, высокие затраты на внедрение и поддержку, а также необходимость обучения персонала.

Методы

Для объективного и обоснованного принятия решения применяются такие аналитические методы, как сравнительный анализ и SWOT-анализ. Эти методы позволяют оценить сильные и слабые стороны, возможности и угрозы различных решений, что особенно важно при работе с высоконагруженными потоками данных, характерными для производственной среды. В ходе исследования был применён метод сравнительного анализа платформ для интеграции и обработки бизнес-событий на производстве, включая Apache Kafka, Apache NiFi, Data-track, Apache Flink, AWS Kinesis и RabbitMQ. Данный метод позволил объективно оценить функциональные возможности, простоту использования, гибкость интеграции, масштабируемость и поддержку мониторинга у каждой из систем. Его полезность заключалась в том, чтобы выявить наиболее подходящее решение для реализации шаблона

обработки данных в рамках платформы DATA-TRACK с учётом специфики производственной среды.

Основной подход заключается в использовании Apache NiFi для потоковой обработки данных благодаря её визуальному интерфейсу, гибкой настройке процессов, масштабируемости и возможности отслеживания происхождения данных. Это позволяет эффективно интегрировать файлы, API, базы данных и преобразовывать их в нужный формат для последующей передачи в систему DATA-TRACK.

Сильные стороны	Слабые стороны
Повышение эффективности: Новые платформы часто предлагают более современные и эффективные решения для автоматизации и оптимизации производственных процессов. Гибкость и масштабируемость: Современные платформы позволяют легко адаптироваться под изменяющиеся условия и требования бизнеса. Инновации и новые возможности: Переход на новые платформы открывает доступ к передовым технологиям, таким как искусственный интеллект, машинное обучение и анализ больших данных.	Высокие затраты: Внедрение новых платформ может потребовать значительных финансовых вложений. Сложность внедрения: Процесс перехода может быть длительным и требовать значительных усилий со стороны ИТ-специалистов и пользователей. Риск сопротивления изменениям: Персонал может сопротивляться переходу на новые системы из-за необходимости обучения и адаптации.
Возможности	Угрозы
Улучшение конкурентоспособности: Внедрение современных платформ может повысить конкурентоспособность предприятия за счет повышения качества продукции и снижения издержек. Доступ к новым рынкам: Использование передовых технологий может открыть новые возможности для расширения бизнеса и выхода на новые рынки.	Технологические риски: Возможны проблемы с интеграцией и совместимостью новых систем с уже существующими. Санкционные риски: Использование зарубежных платформ может быть связано с рисками, связанными с санкциями и ограничениями.

Рис. 1. SWOT-анализ по внедрению платформ по обработке данных на производстве

Переход на новые платформы для обработки данных на производственных предприятиях позволяет преобразовать бизнес-процессы и повысить конкурентоспособность компании (Рисунок 1). Несмотря на высокие затраты, сложность внедрения и возможное сопротивление со стороны персонала, преимущества такого перехода значительно перевешивают эти риски. Современные платформы предлагают инструменты для автоматизации и оптимизации, что позволяет значительно повысить производительность и качество продукции, а также снизить издержки за счет более эффективного управления ресурсами [6]. Кроме того, новые технологии, такие как искусственный интеллект и анализ больших данных, открывают доступ к инновационным возможностям, которые могут улучшить управленческое принятие решений и адаптивность бизнеса к изменяющимся

условиям рынка. В условиях растущей конкуренции и быстро меняющейся экономической среды, внедрение передовых платформ становится необходимым шагом для укрепления позиций на рынке.

Результаты

SWOT-анализ перехода на новые платформы выявил преимущества и риски, связанные с внедрением современных технологий в производственные процессы. Для более глубокого понимания и выбора оптимального решения необходимо провести сравнительный анализ существующих платформ, таких как Apache Kafka, Apache NiFi, DATA-TRACK, Apache Flink, AWS Kinesis и RabbitMQ. Эти платформы предназначены для интеграции и обработки бизнес-событий в реальном времени, что критически важно для повышения эффективности и конкурентоспособности производственных предприятий.

В ходе исследования использовалась методология многофакторного экспертного сравнения для оценки платформ Apache NiFi, DATA-TRACK и других. Оценка проводилась по шести критериям, выбранным на основании анализа требований к производственным информационным системам, согласно стандартам ISO/IEC 25010. Эти критерии включали централизованную обработку, гибкость интеграции, возможность мониторинга, масштабируемость, настраиваемость и совместимость, а также простоту использования. Данные для анализа собирались из официальной документации платформ, научных статей, кейсов внедрений и интервью с экспертами в области цифровизации производства. Экспертное сравнение позволило объективно оценить сильные и слабые стороны каждой платформы, а также выявить их потенциал для интеграции и совместного использования (Таблица 1).

Таблица 1. Сравнительный анализ платформ для интеграции и обработки бизнес-событий на производстве [7–10]

Критерии оценивания	Apache Kafka	Apache NiFi	DATA-TRACK	Apache Flink	AWS Kinesis	RabbitMQ
Централизованная обработка	+	+	+	+	+	-
Гибкость интеграции	+	+	+	+	+	+
Возможность мониторинга	-	+	+	+	+	-
Масштабируемость	+	+	+	+	+	-
Настраиваемость и совместимость	+	+	+	+	-	+
Простота использования	-	+	+	-	+	+

Сравнительный анализ платформ для интеграции и обработки бизнес-событий на производстве показал, что Apache NiFi в сочетании с DATA-TRACK является оптимальным решением для обеспечения централизованной обработки данных, гибкости интеграции и возможности мониторинга. Apache NiFi выделяется своим визуальным интерфейсом, который упрощает создание и управление потоками данных, что особенно важно для сложных производственных процессов [10]. DATA-TRACK дополняет NiFi, предоставляя мощные инструменты для мониторинга и анализа данных, что позволяет оперативно реагировать на изменения и улучшать качество данных. В совокупности эти платформы обеспечивают высокую масштабируемость и простоту использования, что делает их подходящими для крупных предприятий.

В отличие от Apache Kafka, Flink и AWS Kinesis, которые могут быть сложными в настройке и требуют значительных технических знаний, Apache NiFi и DATA-TRACK предлагают более интуитивно понятный интерфейс. Это позволяет быстрее и легче интегрировать их в существующие системы. Кроме того, Kafka и Flink могут быть избыточными для задач, не требующих сложных вычислений в реальном времени, в то время как NiFi и DATA-TRACK обеспечивают необходимый уровень мониторинга и анализа без излишней сложности. NiFi и DATA-TRACK предоставляют высокую степень настраиваемости и совместимости с различными системами и форматами данных, что позволяет легко адаптировать их под специфические требования. AWS Kinesis ограничен в плане кастомизации и требует работы в рамках экосистемы AWS.

DATA-TRACK известен также как DATA-CI, с другой стороны, представляет собой систему мониторинга и анализа данных, которая позволяет отслеживать изменения и события в данных [11]. Она обеспечивает возможность отслеживания изменений в данных, что позволяет оперативно реагировать на изменения в производственных процессах. Основное преимущество платформы заключается в ее способности интегрироваться с различными промышленными системами и оборудованием через широкий спектр протоколов, таких как OPC DA, Modbus и MQTT, что обеспечивает бесшовное взаимодействие и сбор данных в реальном времени. Это позволяет предприятиям оперативно реагировать на изменения и принимать обоснованные решения.

Платформа DATA-CI включает в себя мощные инструменты для сбора, обработки и анализа больших данных, что делает ее особенно эффективной в условиях современного производства [12]. Использование таких технологий, как Apache Drill, NiFi и Flink, позволяет обрабатывать огромные объемы данных и предоставлять аналитику в реальном времени. Это

особенно важно для крупных производственных предприятий, где оперативность и точность данных играют ключевую роль в оптимизации процессов и повышении эффективности.

Одним из ключевых преимуществ DATA-CI является поддержка концепции Continuous Intelligence (CI), которая автоматизирует фиксацию производственных проблем и предлагает оптимальные решения на основе анализа данных. Это позволяет предприятиям не только выявлять проблемы, но и оперативно реагировать на них, изменяя производственное поведение и повышая общую производительность. Например, в сталеплавильном производстве платформа помогает диспетчерам оптимизировать маршруты ковшей с горячим металлом, что снижает риск задержек и повышает качество продукции [12]. В отличие от многих общедоступных программ, DATA-CI предлагает комплексное решение для управления бизнес-процессами и MES, что позволяет синхронизировать, координировать и оптимизировать производственные операции. Платформа включает в себя инструменты для визуализации данных, такие как BI-системы и интерактивные блокноты, которые позволяют анализировать данные и принимать решения без необходимости глубоких технических знаний. Это делает её доступной для широкого круга пользователей, от инженеров до менеджеров.

После выбора платформы на начальном этапе работы с Apache NiFi была выполнена детальная проработка архитектуры решения и изучены ключевые компоненты платформ, включая процессоры, репозитории и механизмы отслеживания данных. Основное внимание было уделено изучению архитектуры и компонентов Apache NiFi, а также их интеграции с платформой DATA-TRACK. Были рассмотрены различные виды процессоров, такие как Data Ingestion, Data Transformation, Routing and Mediation, Database Access, Attribute Extraction, System Interaction, Data Egress/Sending Data, Splitting and Aggregation, HTTP, и Amazon Web Services. Также были изучены механизмы отслеживания изменений FlowFile, такие как Data Provenance, Nodes Status History и Bulletin Board.

На следующем этапе были настроены базовые процессы обработки данных, включая создание сущности "Партия заготовок" со всеми необходимыми параметрами. Были использованы процессоры GenerateFlowFile, SplitJson, InvokeHTTP и LogAttribute для генерации запросов, разбиения JSON-параметров, передачи данных на внешние API и логирования операций соответственно. Были настроены эндпоинты для создания типа сущности, параметров, экземпляров сущностей и получения информации о них.

Далее работа была сфокусирована на реализации автоматизированной цепочки обработки данных между внешними системами и платформой DATA-TRACK. Это

потребовало внедрения дополнительных процессоров для преобразования данных, таких как EvaluateJsonPath, UpdateAttribute и AttributesToJSON. Были настроены процессоры для обработки полученной информации от API, переработки JSON данных в атрибуты FlowFile и перевода типа данных атрибутов FlowFile. Были реализованы запросы для получения и изменения данных в базе данных DATA-TRACK.

На заключительном этапе была завершена интеграция с внешней средой и обеспечен механизм предотвращения дублирования данных через реализацию булевых параметров контроля состояния экземпляров. Были добавлены дополнительные процессоры для изменения данных в таблице коллег и настройки виртуального сигнала. Были реализованы запросы для изменения состояния экземпляров в базе данных DATA-TRACK и получения информации об экземплярах типа сущности.

В ходе последнего этапа был успешно реализован шаблон, который с периодичностью в 15 секунд изменяет значение виртуального сигнала «температура в печи» на случайное значение от 1.11 до 100.11. Были использованы процессоры GenerateFlowFile и InvokeHTTP для генерации случайных чисел и записи их в значение виртуального сигнала.

Обсуждение

В ходе выполнения работы был выявлен ряд проблем и задач, которые потребовали оперативного решения и взаимодействия с техническими специалистами предприятия. Основные проблемы включали некорректные пути подключения к сервису «DATA-TRACK», различия в параметрах процессоров между версиями Apache NiFi, а также необходимость поиска качественной документации по взаимодействию со средой Apache NiFi.

Благодаря совместным усилиям и анализу актуальной документации, проблемы были успешно решены и поставленные цели достигнуты. Были реализованы и протестированы различные процессоры Apache NiFi, настроены ключевые маршруты взаимодействия с внешними эндпоинтами платформы, а также разработаны и внедрены логики обработки для создания новой сущности «Партия заготовок» с заданным набором параметров и соответствующей маршрутизацией данных.

Заключение

В заключение можно сказать, что переход к цифровым форматам взаимодействия и обработки информации меняет подход к управлению и организации производственных процессов, делая их более прозрачными, гибкими и эффективными. Сравнительный анализ платформ, таких как Apache Kafka, Apache NiFi, DATA-TRACK, Apache Flink, AWS Kinesis и RabbitMQ, показал, что Apache NiFi в сочетании с DATA-TRACK является оптимальным

решением для централизованной обработки данных, обеспечивая гибкость интеграции и возможность мониторинга. Эти платформы предлагают высокую масштабируемость и простоту использования, что делает их подходящими для крупных предприятий. В результате работы была успешно реализована система интеграции с внешней средой и обработки её данных с использованием платформы DATA-TRACK и инструментария Apache NiFi. Разработанное решение является масштабируемым и готовым к дальнейшему развитию и внедрению в реальных производственных условиях.

Литература

1. Ли Цзюнь, Юй Гуанюань Актуальность внедрения процесса цифровизации в деятельность предприятий // Universum: экономика и юриспруденция. 2021. №11 (86). URL: <https://cyberleninka.ru/article/n/aktualnost-vnedreniya-protssessa-tsifrovizatsii-v-deyatelnost-predpriyatiy> (дата обращения: 07.04.2025).
2. Акулова Елена. Что такое цифровизация и в каких областях она применяется // NEPPE 23.03.2022//URL:<https://neppo.ru/news/что-такое-цифровизация-и-в-каких-сферах-она-применяется/>
3. M. Weske, Business Process Management: Concepts, Languages, Architectures. Second Edition. Springer, 2012.
4. Herzberg N., Meyer A., Weske M. An event processing platform for business process management //2013 17th IEEE International Enterprise Distributed Object Computing Conference. – IEEE, 2013. – С. 107-116.
5. Зоиров Ш. К. Методика интеграции информационных систем для оптимизации бизнес-процессов: разработка и анализ методов и рекомендаций для деревообрабатывающих предприятий: магистерская диссертация / Ш. К. Зоиров; Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, Институт радиоэлектроники и информационных технологий-РТФ, Базовая кафедра «Аналитика больших данных и методы видеоанализа». — Екатеринбург, 2024. — 77 с. — Библиогр.: с. 68-77 (51 назв.).
6. Seiger R. et al. Integrating process management and event processing in smart factories: A systems architecture and use cases //Journal of Manufacturing systems. – 2022. – Т. 63. – С. 575-592.
7. RabbitMQ против Kafka: два разных подхода к обмену сообщениями // HABR URL: <https://habr.com/ru/companies/itsumma/articles/416629/> (дата обращения: 21.03.2025).

8. Amazon Kinesis vs Apache Flink: Which Tool is Better for Your Next Project? // ProjectPro
URL: <https://www.projectpro.io/compare/amazon-kinesis-vs-apache-flink> (дата обращения: 21.03.2025).
9. Что такое Apache NiFi // gitverse blog URL:
<https://gitverse.ru/blog/articles/architecture/248-что-такое-apache-nifi> (дата обращения: 21.03.2025).
10. Что такое NiFi // Habr URL: <https://habr.com/ru/companies/tbank/articles/743886/> (дата обращения: 19.03.25).
11. Продуктовая линейка платформы DATA-TRACK // Дата-центр автоматика URL:
https://dc.ru/delivery_of_products/ (дата обращения: 21.03.2025).
12. Спринт в будущее — код успеха. «ДАТА-ЦЕНТР Автоматика» заявила о выпуске единого решения в области Continuous Intelligence // ДАТА-ЦЕНТР Автоматика URL:
<https://dc.ru/news/2024/11/02/159/> (дата обращения: 07.04.2025).