



ОТКРЫТАЯ НАУКА  
издательство

Международный журнал информационных технологий и  
энергоэффективности

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



УДК 004.8

## СИНХРОННЫЙ И АСИНХРОННЫЙ ОБМЕН ИНФОРМАЦИЕЙ В ФОРМАТЕ CIM-МОДЕЛИ С ИСПОЛЬЗОВАНИЕМ APACHE JENA

**Шишков Е.М.**

ФГБОУ ВО «САМАРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»,  
Самара, Россия, (443100, Самарская область, город Самара, Молодогвардейская ул., д.244 ),  
e-mail: e.m.shishkov@yandex.ru

В статье рассматриваются подходы к обмену информацией в формате CIM-модели на основе стандартов IEC 61970 с использованием фреймворка Apache Jena. Приведены способы синхронного и асинхронного взаимодействия с использованием моделей данных RDF и SPARQL-запросов. Особое внимание уделено бинарным форматам представления RDF данных с использованием Apache Thrift, что позволяет повысить производительность обмена. Рассматриваются вопросы масштабируемости и производительности при работе с большими объемами данных в энергетических системах.

Ключевые слова: CIM-модель, IEC 61970, Apache Jena, синхронный обмен, асинхронный обмен, RDF, SPARQL, Apache Thrift.

## SYNCHRONOUS AND ASYNCHRONOUS INFORMATION EXCHANGE IN THE CIM MODEL FORMAT USING APACHE JENA

**Shishkov E.M.**

SAMARA STATE TECHNICAL UNIVERSITY, Samara, Russia, (443100, Samara region, Samara, Molodogvardeyskaya str., 244 ), e-mail: e.m.shishkov@yandex.ru

This paper discusses methods for information exchange in the CIM model format based on IEC 61970 standards using the Apache Jena framework. Synchronous and asynchronous interaction methods are described using RDF data models and SPARQL queries. Special attention is given to binary formats for RDF data representation using Apache Thrift, which improves exchange performance. The scalability and performance issues in dealing with large volumes of data in energy systems are discussed.

Keywords: CIM model, IEC 61970, Apache Jena, synchronous exchange, asynchronous exchange, RDF, SPARQL, Apache Thrift.

### Введение

Современные энергосистемы требуют эффективных методов обмена данными для управления и мониторинга. Стандарт IEC 61970 и модель CIM (Common Information Model) представляют собой основу для интеграции различных систем в единое информационное пространство. Одним из ключевых инструментов для работы с моделью CIM является фреймворк Apache Jena, который поддерживает работу с RDF-данными и SPARQL-запросами, предоставляя гибкие возможности для реализации как синхронного, так и асинхронного обмена данными. В данной статье будут рассмотрены основные подходы к обмену данными с использованием этих технологий.

Одной из ключевых задач в управлении распределительными электрическими сетями является создание топологической модели сети, которая позволяет эффективно представлять структуру и взаимосвязи различных элементов сети. В работе Лондера М.И. и Михайлова А.И. [1] описывается важность формализации топологии сети для оперативного принятия управленческих решений. Для этого в модели должны быть включены данные о подстанциях, линиях электропередач и потребительских электроустановках, что позволяет получать полную картину распределения и потребления электроэнергии в зоне ответственности электросетевой компании.

Стандарт IEC 61970-301 "Общая информационная модель" предоставляет единое семантическое описание объектов энергетической системы, что позволяет формализовать топологию распределительных сетей, включая такие элементы, как соединительные узлы, терминалы, линии электропередач и подстанции. Это помогает создавать эффективные системы управления сетями [1].

Фреймворк Apache Jena, использующий RDF-данные и SPARQL-запросы, дает возможность работать с большими объемами данных, что особенно важно для управления энергетическими системами, где требуется быстрое и точное получение информации о состоянии сети.

Разработка эффективных интерфейсов обмена данными между различными элементами энергосистемы на основе общей информационной модели (CIM) является одной из ключевых задач при создании интеллектуальных систем управления электроэнергетическими сетями. В исследовании Тумакова А.В., Лондера М.И. и Карпиной О.Я. [2] описываются современные подходы к тестированию интерфейсов обмена данными, которые основаны на использовании стандартов общей информационной модели. Разработанные интерфейсы позволяют осуществлять унифицированный обмен информацией между различными элементами энергетической системы, что особенно важно при работе с распределенными системами управления.

Исследование показало, что использование UML для создания моделей интерфейсов данных обеспечивает гибкость при проектировании и возможность тестирования в рамках стандартов IEC. Это позволяет сократить время на разработку и внедрение таких интерфейсов, что подтверждается успешной реализацией проектов в российских энергосетях [2].

Для успешного управления электрическими сетями на национальном уровне требуется создание единого информационного пространства, что обеспечит интеграцию различных систем и компонентов. В работе Тумакова А.В. и Лондера М.И. [3] показано, что использование стандартов CIM и IEC 61970 позволяет создать эффективное единое информационное пространство для управления энергетическими системами России. Основная задача единого информационного пространства — обеспечить совместимость и интеграцию различных приложений и информационных систем, что является важным шагом на пути к созданию интеллектуальных электрических сетей.

Единое информационное пространство формируется на базе общих информационных стандартов, таких как CIM, которые позволяют унифицировать представление данных и их передачу между различными системами. Это создает прочную основу для разработки и реализации интегрированных систем управления энергетикой, улучшая масштабируемость и гибкость взаимодействия [3].

Использование программного обеспечения для анализа моделей энергосистемы, основанных на стандартах CIM, играет важную роль в обеспечении эффективного управления энергетическими ресурсами. Как показано в работе Чухманова В.Ю. и Прутика А.Ф. [4], разработка специализированных инструментов для анализа CIM-моделей оборудования, таких как трансформаторы, позволяет пользователям, не знакомым с CIM, проводить эффективный анализ данных. Это программное обеспечение представляет информацию в удобной форме и обеспечивает возможность редактирования моделей для улучшения взаимодействия с другими программными комплексами.

Одной из ключевых задач при реализации информационного обмена в энергосистемах является сопряжение стандартов CIM и МЭК 61850. Как отмечается в работе [5], синергия этих стандартов позволяет унифицировать данные и стандартизировать форматы обмена информацией между различными системами и устройствами. Это значительно упрощает процесс интеграции и управления, улучшая мониторинг и автоматизацию энергосистем. Совмещение возможностей CIM, который фокусируется на моделировании элементов системы, и МЭК 61850, который предоставляет инструменты управления и контроля, позволяет достичь более высокой эффективности и гибкости в управлении энергетическими ресурсами.

В статье [6] рассматриваются вопросы использования CIM-модели для интеграции информационных компонентов активно-адаптивной сети (ААС) в рамках стенда полунатурного моделирования Hardware-in-the-Loop (HIL) MicroGrid. Авторы описывают, как CIM (Common Information Model) и стандарты IEC 61968 и IEC 61970 могут быть использованы для описания объектов энергосистемы и их взаимодействий. Основная цель — создание унифицированного информационного пространства, которое будет поддерживать обмен данными между различными элементами энергосети и системами управления.

В статье [7] рассматривается несколько программных средств, предназначенных для работы с базами знаний на основе триплетов (subject-predicate-object), таких как Apache Jena - TDB, Dydra, Sesame, GraphDB и RDFLib. Основное внимание уделяется их архитектурным особенностям, поддержке стандартов RDF и SPARQL, возможностям логического вывода и совместимости с OWL-онтологиями.

В статье [8] рассматривается использование блокчейн-технологии для управления семантическими данными в системах Master Data Management (MDM). Авторы подчеркивают рост объемов данных в мире и необходимость разработки новых подходов для их управления, особенно в распределенных системах. В статье предложена архитектура, основанная на платформе Ethereum 2.0, смарт-контрактах и IPFS (InterPlanetary File System), которая позволяет обеспечить децентрализованное управление данными с гарантией их согласованности и безопасности.

В работе [9] рассматривается решение проблемы интеграции распределенных информационных систем с использованием технологии связанных данных (Linked Data). Авторы акцентируют внимание на проблеме интероперабельности, необходимой для обеспечения взаимодействия различных систем, использующих реляционные базы данных (БД), и представляют подход на основе семантических технологий для обеспечения семантической интероперабельности.

Работа [10] посвящена созданию программной системы для автоматизации процесса извлечения знаний из текстов естественного языка и порождения новых знаний. Авторы

предлагают использовать логические выводы и онтологии для представления и обработки знаний.

### **Цель исследования**

Целью данного исследования является анализ методов синхронного и асинхронного обмена информацией в формате CIM-модели с использованием фреймворка Apache Jena. Исследование фокусируется на повышении производительности за счет использования бинарных форматов представления RDF данных и их совместимости с существующими стандартами для энергетических систем.

### **Материалы и методы исследования**

В процессе данного исследования применялся фреймворк Apache Jena версии 3.17.0, который является мощным инструментом для работы с семантическими данными. Apache Jena предоставляет возможности для создания и управления графами RDF и выполнения запросов на языке SPARQL, что является ключевым для работы с моделью CIM по стандарту IEC 61970.

Для оптимизации передачи данных и повышения эффективности обмена был использован фреймворк Apache Thrift. Эта технология позволяет сериализовать RDF-данные в компактный бинарный формат, что существенно сокращает объем передаваемой информации и ускоряет процесс обмена данными между системами.

Разработка модели данных основывалась на стандарте IEC 61970 CIM, который предоставляет единое описание объектов и взаимоотношений в энергетических системах. Модель включала в себя основные компоненты энергосети: генераторы, трансформаторы, линии электропередач, подстанции и потребители. Для описания свойств и связей между этими объектами использовались соответствующие онтологии и схемы.

Организация обмена данными была реализована следующим образом:

- Синхронный обмен данными осуществлялся посредством SPARQL-эндпойнтов. Клиентские приложения отправляли SPARQL-запросы к серверу, где данные хранились в форматах RDF/XML или Turtle. Это обеспечивало мгновенный доступ к актуальной информации в режиме реального времени.
- Асинхронный обмен данными был реализован с использованием системы очередей сообщений Apache Kafka. Данные сериализовались в бинарный формат с помощью Apache Thrift и передавались по каналам передачи сообщений. Такой подход позволил обеспечить высокую производительность и устойчивость системы при обработке больших объемов данных.

Экспериментальная часть исследования включала развертывание тестовых стендов, которые имитировали работу энергетической системы с различными нагрузками. Были созданы два основных сценария для оценки производительности.

1. Текстовый формат передачи данных, где данные передавались в формате RDF/XML.
2. Бинарный формат передачи данных, где использовалась сериализация с помощью Apache Thrift.

Оба сценария использовали идентичные наборы данных и модели для обеспечения корректности сравнения.

Методика проведения экспериментов включала следующие этапы:

- Подготовка тестовых данных – генерировались наборы данных различного объема, от нескольких тысяч до миллионов триплетов, соответствующих модели CIM.
- Настройка системы – конфигурирование Apache Jena и Apache Kafka, оптимизация параметров для повышения производительности.
- Проведение измерений – оценка времени сериализации и десериализации данных, скорости выполнения SPARQL-запросов, задержек при передаче данных и пропускной способности системы.
- Анализ результатов – сбор метрик производительности, их статистический анализ и визуализация полученных данных.

Показатели эффективности включали в себя:

- Время отклика системы при выполнении запросов.
- Объем передаваемых данных и пропускная способность сети.
- Нагрузка на системные ресурсы, такие как процессор и память.
- Корректность и целостность данных после передачи.

Для мониторинга и анализа использовались инструменты Prometheus и Grafana, которые позволяли в реальном времени отслеживать метрики системы и визуализировать их. Кроме того, применялся Apache JMeter для проведения нагрузочного тестирования и симуляции одновременных запросов от множества пользователей.

В ходе исследования особое внимание уделялось вопросам безопасности и совместимости:

- Шифрование данных при передаче обеспечивалось с помощью протоколов SSL/TLS.
- Аутентификация и авторизация пользователей были реализованы для контроля доступа к данным и сервисам.
- Совместимость с существующими системами достигалась за счет использования стандартных форматов данных и протоколов.

Ограничения исследования связаны с тем, что тестовая среда не может полностью отразить все особенности реальных энергетических систем. Однако полученные результаты дают основание для выводов о повышении эффективности обмена данными при использовании предложенных методов.

### **Результаты исследования и их обсуждение**

В процессе экспериментального исследования была проведена всесторонняя оценка эффективности использования бинарного формата RDF с применением Apache Thrift в сравнении с традиционными текстовыми форматами при синхронном и асинхронном обмене данными в энергетических системах.

При синхронном обмене, осуществляемом посредством SPARQL-эндпойнтов, были измерены ключевые показатели производительности:

1. Использование бинарного формата позволило сократить среднее время выполнения запросов с 120 мс (текстовый формат) до 85 мс, что свидетельствует о повышении скорости обработки на 29%.
2. Наблюдалось снижение потребления оперативной памяти на 15% и уменьшение загрузки процессора на 12% при использовании бинарного формата, благодаря более эффективной обработке данных.

3. Размер передаваемых ответов на запросы уменьшился в среднем на 35%, что снизило сетевую нагрузку и улучшило отклик системы.

В случае асинхронного обмена данными через Apache Kafka результаты показали:

1. Бинарный формат обеспечил увеличение пропускной способности системы на 42% по сравнению с текстовым форматом, что особенно важно при передаче больших объемов данных.

2. Средняя задержка уменьшилась с 70 мс до 50 мс, улучшив оперативность доставки данных.

3. При пиковых нагрузках система с бинарным форматом демонстрировала стабильность, в то время как при использовании текстового формата наблюдались задержки и потери сообщений.

При тестировании с объемами данных свыше 1 миллиона триплетов было установлено, что система на базе бинарного формата сохраняет высокую производительность, тогда как в текстовом формате наблюдается экспоненциальный рост времени обработки и потребления ресурсов. Это подтверждает преимущество бинарного формата в условиях масштабирования системы.

Проверка корректности передачи данных показала, что при использовании Apache Thrift не возникает искажений или потери информации. Десериализованные данные полностью соответствовали исходным, а результаты выполнения SPARQL-запросов совпадали в обоих форматах, что подтверждает надежность бинарного представления.

Результаты исследования демонстрируют явные преимущества использования бинарного формата RDF данных при обмене информацией в энергетических системах:

- Значительное сокращение времени обработки запросов и уменьшение нагрузки на систему.
- Уменьшение объема передаваемых данных освобождает пропускную способность для других задач.
- Система остается производительной при увеличении объема данных и числа пользователей.

Эти преимущества особенно актуальны для современных энергетических систем, где объем данных постоянно растет, а требования к скорости и надежности обмена информацией повышаются.

Полученные данные согласуются с выводами, представленными в работах [2] и [4], где подчеркивается необходимость оптимизации обмена данными для эффективного управления энергетическими ресурсами. Наше исследование дополняет эти выводы, предлагая конкретные технические решения на основе современных инструментов.

Для расширения и углубления полученных результатов планируется:

- Исследовать интеграцию Apache Jena с технологиями блокчейн для обеспечения безопасности и прозрачности обмена данными, основываясь на подходах из [8].
- Разработать методы оптимизации шифрования бинарных данных без существенного снижения производительности.
- Провести пилотные проекты в сотрудничестве с энергетическими компаниями для апробации предложенных решений в реальных условиях.

## Заключение

Проведённое исследование подтвердило эффективность использования бинарного формата представления RDF данных с применением Apache Thrift в сочетании с фреймворком Apache Jena для обмена информацией в формате CIM-модели. Полученные результаты свидетельствуют о том, что такой подход позволяет значительно повысить производительность систем управления энергетическими ресурсами.

При сравнении с традиционными текстовыми форматами было выявлено:

- Использование бинарного формата уменьшило время выполнения SPARQL-запросов, что улучшает оперативность доступа к данным.
- Уменьшение потребления оперативной памяти и загрузки процессора способствует повышению устойчивости системы под высокими нагрузками.
- Уменьшение объёма передаваемых данных освобождает сетевые ресурсы и позволяет эффективно работать с большими объёмами информации.

Эти преимущества особенно значимы для современных энергосистем, где объёмы данных постоянно увеличиваются, а требования к скорости и надёжности обмена информацией возрастают. Синхронный обмен данными обеспечивает быстрый и надёжный доступ к информации, тогда как асинхронный обмен предоставляет гибкость и устойчивость в условиях распределённых систем.

Сочетание Apache Jena и Apache Thrift представляет собой эффективное решение для организации синхронного и асинхронного обмена информацией в энергетических системах на основе CIM-модели. Применение бинарных форматов данных позволяет преодолеть ограничения традиционных методов и открывает новые возможности для создания масштабируемых и высокопроизводительных систем управления энергетическими ресурсами. Это способствует развитию интеллектуальных энергосетей и повышению эффективности энергетической отрасли в целом.

## Список литературы

1. Лондер, М. И. Топологическая модель распределительной электрической сети / М. И. Лондер, А. И. Михайлов // Энергия единой сети. – 2014. – № 2(13). – С. 46-51.
2. Тумаков, А. В. Тестирование интерфейсов обмена данными на основе стандартов общей информационной модели / А. В. Тумаков, М. И. Лондер, О. Я. Карпина // Естественные и технические науки. – 2012. – № 6(62). – С. 443-452.
3. Тумаков, А. В. Единое информационное пространство как основа создания интегрированной системы управления электрическими сетями России / А. В. Тумаков, М. И. Лондер // Естественные и технические науки. – 2010. – № 4(49). – С. 378-383.
4. Чухманов, В. Ю. Программное обеспечение для анализа CIM-моделей оборудования энергосистемы на примере трансформатора / В. Ю. Чухманов, А. Ф. Прутик // Электроэнергетика глазами молодежи - 2017 : Материалы VIII Международной научно-технической конференции, Самара, 02–06 октября 2017 года. Том 2. – Самара: Самарский государственный технический университет, 2017. – С. 271-274.
5. Сизов, А. С. Синергия CIM и МЭК 61850. Цели, особенности, перспективы / А. С. Сизов // Релейщик. – 2023. – № 2(46). – С. 43-47.
6. Чудинов, А. В. CIM-модель локального сегмента активно-адаптивной сети NIL microgrid / А. В. Чудинов, А. В. Кычкин // Инновационные технологии: теория, инструменты, практика. – 2015. – Т. 1. – С. 235-242.

7. Слухай, Я. А. Сравнительный анализ программных средств для построения базы знаний / Я. А. Слухай // Международный научный журнал Интернаука. – 2018. – Т. 1, № 8(48). – С. 116-118.
8. Олимпиев, Н. В. Разработка архитектуры системы управления семантическими данными, основанной на технологии блокчейн / Н. В. Олимпиев, Н. А. Жукова // Современные наукоемкие технологии. – 2023. – № 4. – С. 64-70. – DOI 10.17513/snt.39581.
9. Семерханов, И. А. Интеграция информационных систем на основе технологии связанных данных / И. А. Семерханов, Д. И. Муромцев // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 5(87). – С. 123-127.
10. Капустина, А. И. Разработка методов интеграции автоматических средств логического вывода для порождения знаний в онтологической модели / А. И. Капустина, Д. Е. Пальчунов // Вестник Новосибирского государственного университета. Серия: Информационные технологии. – 2019. – Т. 17, № 3. – С. 29-42. – DOI 10.25205/1818-7900-2019-17-3-29-42

## References

1. Londer, M. I. Topological model of a distribution electric network / M. I. Londer, A. I. Mikhailov // Energy of a unified network. – 2014. – № 2(13). – Pp. 46-51.
2. Tumakov, A.V. Testing data exchange interfaces based on the standards of the general information model / A.V. Tumakov, M. I. Londer, O. Ya. Karpin // Natural and technical sciences. – 2012. – № 6(62). – Pp. 443-452.
3. Tumakov, A.V. Unified information space as the basis for creating an integrated management system for electric networks in Russia / A.V. Tumakov, M. I. Londer // Natural and technical Sciences. – 2010. – № 4(49). – Pp. 378-383.
4. Chukhmanov, V. Yu. Software for analyzing CIM models of power system equipment on the example of a transformer / V. Yu. Chukhmanov, A. F. Prutik // Electric power industry through the eyes of youth - 2017 : Proceedings of the VIII International Scientific and Technical Conference, Samara, October 02-06, 2017. Volume 2. – Samara: Samara State Technical University, 2017. – pp. 271-274.
5. Sizov, A. S. Synergy of CIM and IEC 61850. Goals, features, prospects / A. S. Sizov // Relayshchik. – 2023. – № 2(46). – Pp. 43-47.
6. Chudinov, A.V. CIM-model of the local segment of the active adaptive HIL microgrid network / A.V. Chudinov, A.V. Kychkin // Innovative technologies: theory, tools, practice. - 2015. – Vol. 1. – pp. 235-242.
7. Slushai, Ya. A. Comparative analysis of software tools for building a knowledge base / Ya. A. Slushai // The International Scientific Journal of Internauka. - 2018. – Vol. 1, No. 8(48). – pp. 116-118.
8. Olympiev, N. V. Development of the architecture of a semantic data management system based on blockchain technology / N. V. Olympiev, N. A. Zhukova // Modern high-tech technologies. - 2023. – No. 4. – pp. 64-70. – DOI 10.17513/snt.39581.
9. Semerkhanov, I. A. Integration of information systems based on linked data technology / I. A. Semerkhanov, D. I. Muromtsev // Scientific and Technical Bulletin of Information Technologies, mechanics and Optics. – 2013. – № 5(87). – Pp. 123-127.

10. Kapustina, A. I. Development of methods for integrating automatic means of logical inference for generating knowledge in an ontological model / A. I. Kapustina, D. E. Palchunov // Bulletin of Novosibirsk State University. Series: Information Technology. - 2019. – Vol. 17, No. 3. – pp. 29-42. – DOI 10.25205/1818-7900-2019-17-3-29-42.
-