Описание функциональных характеристик

Универсальная программная платформа для корпоративных Информационных систем ПАО «РусГидро»

2025 г.

Оглавление

1 Наименование системы 3
2 Описание функциональных характеристик 3
2.1 Цели, назначение и область использования Системы 3
2.2 Решения по структуре и функционированию ЕИП 4
2.3 Решения по структуре и функционированию Озера данных 6
3 Интеграция с использованием Универсального сервиса 8
3.1 Типовые варианты интеграционного взаимодействия 8
3.1.1 Односторонняя асинхронная интеграция 8
3.1.2 Двусторонняя асинхронная интеграция 8
3.1.3 Синхронная интеграция 9
4 Интеграция с холодным слоем Озера данных 10
4.1 Получение сведений о данных, хранящихся в Озере данных 10
4.2 Запись данных в Озеро данных 11
4.2.1 Неявная запись данных 11
4.2.2 Явная запись данных 12
4.3 Чтение данных из Озера данных 12
5 Интеграция с горячим слоем Озера данных 13
5.1 Получение сведений о данных, хранящихся в Озере данных 13
5.2 Запись данных в Озеро данных 14
5.2.1 Прямая запись в базу данных ClickHouse 14
5.2.2 Запись через WEB-сервис 15
5.3 Чтение данных из Озера данных 15
5.3.1 Прямое чтение из базы данных ClickHouse 15
5.3.2 Чтение через WEB-сервис 16
6 Требования к техническим средствам 17
6.1 Техническое обеспечение серверной части 17
6.1.1 Решения по техническим средствам клиентской части Системы 17
6.1.2 Решения по техническим средствам серверной части Системы 17
6.1.3 Решения по размещению комплекса технических средств 20

1 Наименование системы

Полное наименование системы: Универсальная программная платформа для корпоративных Информационных систем ПАО «РусГидро».

Условное обозначение системы: Единая интеграционная платформа для обмена данными между корпоративными ИС (ЕИП, Озеро данных).

В настоящем документе для обозначения Озеро данных также используется термин Система.

2 Описание функциональных характеристик

2.1 Цели, назначение и область использования Системы

Универсальная программная платформа для корпоративных Информационных систем ПАО «РусГидро» по виду автоматизированной деятельности относится к системам обработки и передачи информации.

Система является связующим программным обеспечением для централизованного обмена сообщениями между различными ИС с применением сервис-ориентированной архитектуры.

Средствами ЕИП выполняется интеграция ИС ПАО «РусГидро» с использованием разработанного универсального сервиса.

Основные цели использования ЕИП:

- Повышение надежности информационного обмена с помощью консолидации передачи потоков бизнес-сущностей в контуре информационных систем ПАО «РусГидро» за счет переноса обработки и маршрутизации запросов между ИС на сторону шины данных.
- Увеличение скорости ввода в эксплуатацию интеграций за счет применения новой методологии разработки (универсальный формат).
- Обеспечение возможности обмена персональными данными с использованием Системы в соответствии с ФЗ-152 «О персональных данных».

С помощью ЕИП решаются следующие задачи:

- Интеграция разнородных информационных систем в унифицированном формате, эксплуатируемых в настоящее время, а также внедряемых или планируемых к вводу в эксплуатацию в будущем.
- Защита при обмене персональными данными.
- Гарантия доставки отправленных сообщений в интегрируемые системы.
- Применение асинхронных механизмов обмена данными;
- Сбор информации об обмене сообщениями между системами.

На базе ЕИП разработана аналитическая платформа, спроектировано Озеро данных.

Основные цели использования Озера данных:

- Сбор и хранение данных, передаваемых через ЕИП, в едином хранилище.
- Обеспечение возможности загрузки, классификации и обновления структуры данных.
- Предоставление инструментов для трансформации, обогащения и мониторинга данных.
- Автоматическое создание графиков для упрощения анализа и повышения информативности данных.

С помощью Озера данных решаются следующие задачи:

- Выгрузка "сырых" данных из различных информационных систем в Озеро данных через ЕИП.
- Загрузка классифицированных потоков данных.
- Возможность обновления структуры потока данных.
- Описание атрибутивного состава данных в потоках.
- Трансформация (обогащение) данных.
- Мониторинг исторической динамики и содержания потоков данных.
- Автоматическое построение графиков на основе полученных данных.

2.2 Решения по структуре и функционированию ЕИП

ЕИП разработана как иерархическая централизованная система со следующими уровнями иерархии:

1) Бизнес-процессы.
2) Комплексная интеграция.
3) Обмен сообщениями.
4) Интеграция и трансформация данных.

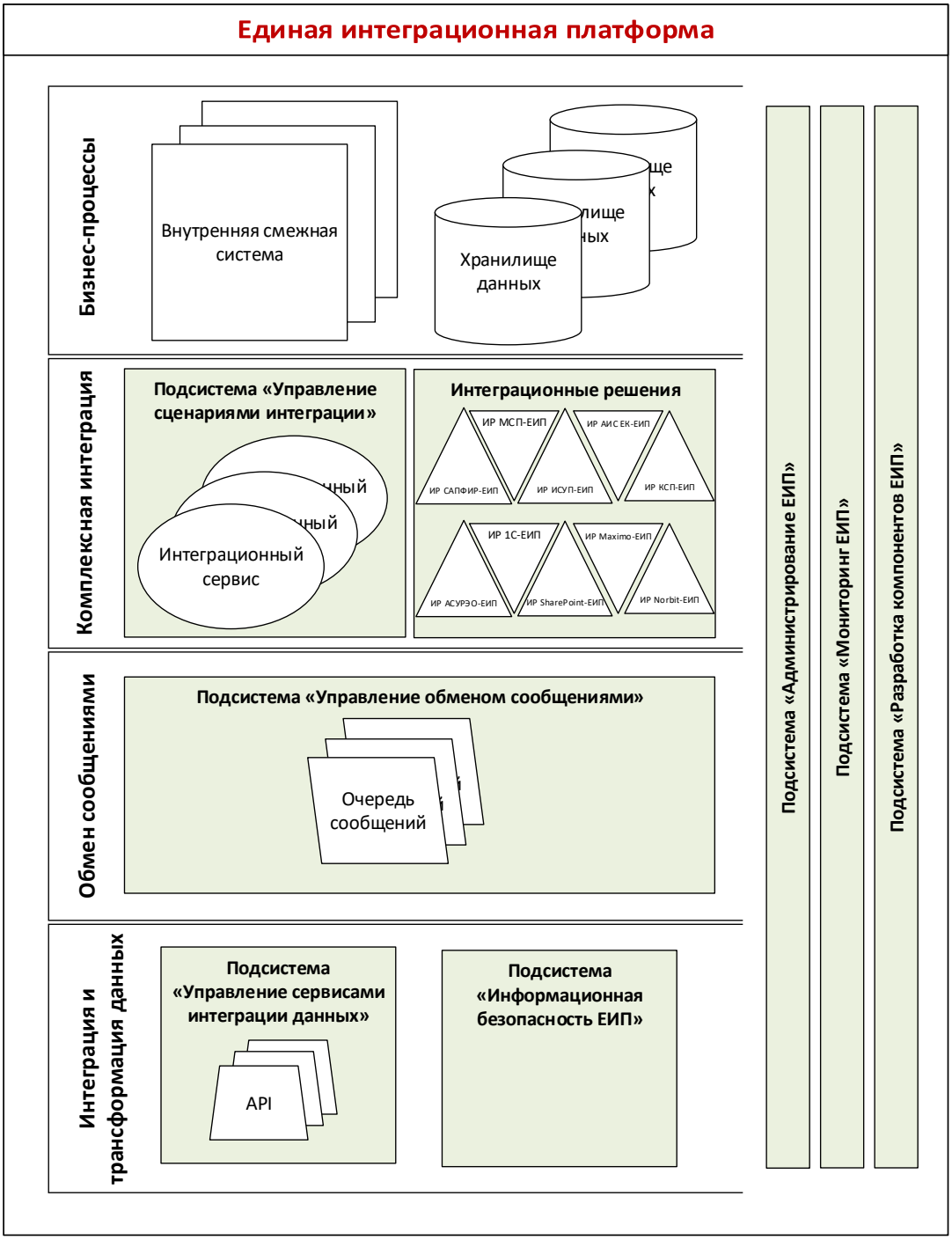


Рисунок 1. Структурная схема Системы

Уровень Бизнес-процессы применяется для управления сквозными бизнес-процессами ПАО «РусГидро». Необходимость использования функциональных подсистем и компонентов данного уровня, включающих средства разработки бизнес-процессов, определяется при разработке интеграционных решений.

Уровень Комплексная интеграция применяется для создания и управления сценариями интеграции данных. Уровень предназначен для решения комплексных задач интеграции информационных потоков между взаимодействующими ИС.

Уровень обмена сообщениями является центральным звеном ЕИП. Компоненты данного уровня строятся на основе интеграционной шины. На данном уровне обеспечиваются основные функции для интеграции сервисов. Интеграционная шина служит посредником между компонентами интеграции и трансформации данных и компонентами создания сценариев интеграции.

Уровень интеграции и трансформации данных применяется для маршрутизации сообщений и обеспечения доступа к данным, хранимым в различных ИС. Уровень предназначен для решения следующих задач:

- трансформация и дополнение информации, с использованием различных разнородных источников данных;
- хранение метаданных обо всех логических сервисах данных и физических источниках данных.

ЕИП является программным слоем (на базе WSO2) для управления потоками данных между интегрированными приложениями. ЕИП выполняет задачи:

- обеспечение интеграции корпоративных ИС;
- поддержка синхронных/асинхронных режимов обмена сообщениями;
- предоставление возможности обмена данными в унифицированном формате;
- повторная доставка сообщений.

ЕИП базируется на сервис-ориентированной архитектуре (SOA). В основе ЕИП имеет «шинный» принцип обмена информацией (ESB — Enterprise Service Bus). Шина ESB обеспечивает централизованное управление, мониторинг, и защиту потоков информации.

2.3 Решения по структуре и функционированию Озера данных

На базе ЕИП разработана аналитическая платформа, спроектировано Озеро данных, содержащее в себе холодный, теплый и горячий слои (рисунок 2).

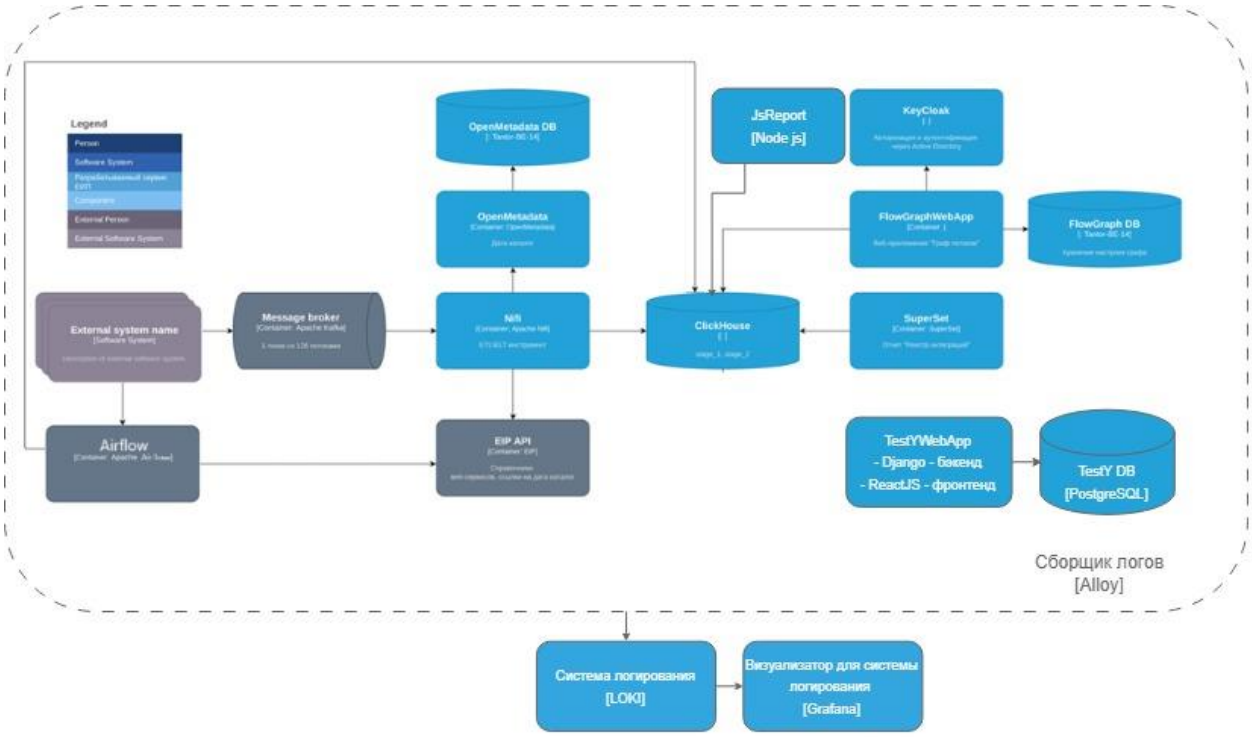


Рисунок 2. Схема компонентов Озера данных

В составе ЕИП разработан модуль «Граф потоков» (Flow Graph) предназначенный для визуализации интеграционного взаимодействия систем. Модуль предоставляет пользователю возможность просмотреть граф в веб-приложении. Используемые технологии: Java, Spring Boot, JavaScript, React.



Рисунок 3. Граф потоков (Flow Graph)

3 Интеграция с использованием Универсального сервиса

3.1 Типовые варианты интеграционного взаимодействия

Описанные в данном разделе типовые варианты интеграции относятся к использованию разработанного универсального сервиса для обмена данными между информационными системами с использованием прикладного протокола взаимодействия.

3.1.1 Односторонняя асинхронная интеграция

При односторонней асинхронной интеграции на стороне «Системы 1» принимающий универсальный сервис не разрабатывается.

Схема односторонней асинхронной интеграции представлена на Рисунке 4. Зеленым цветом на схеме обозначаются процессы, в которых задействован универсальный сервис. Реализованный универсальный сервис на стороне «Системы 2» должен по формату совпадать с универсальным интерфейсом на стороне ЕИП.

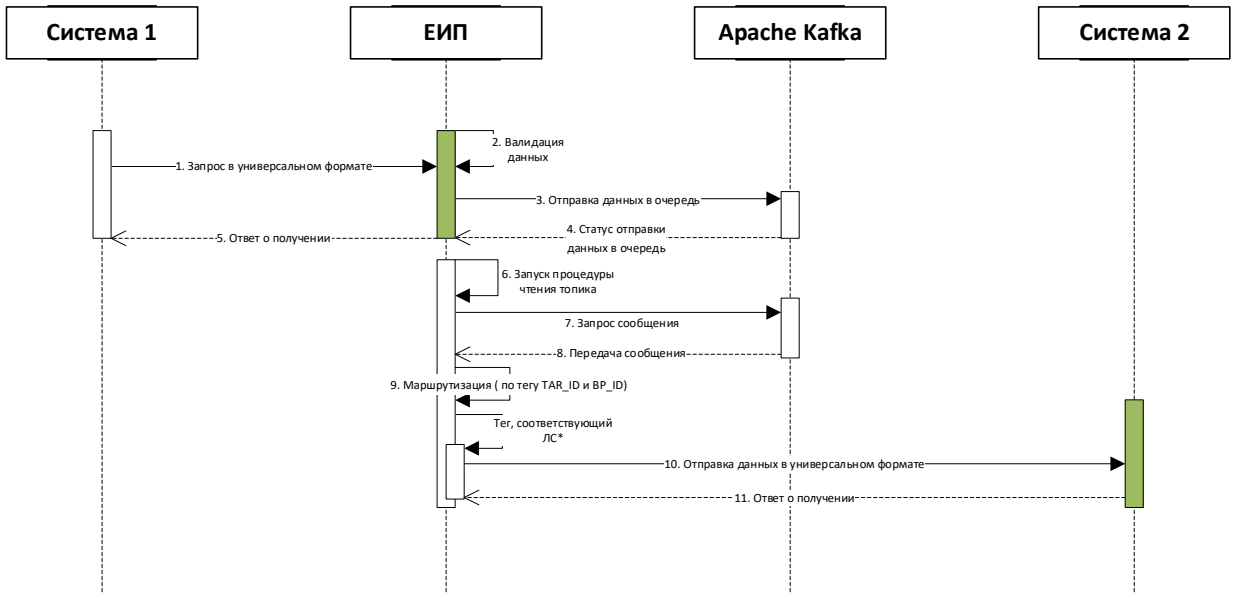


Рисунок 4. Схема взаимодействия между системами при односторонней асинхронной интеграции

3.1.2 Двусторонняя асинхронная интеграция

При двусторонней асинхронной интеграции на стороне «Системы 1» и «Системы 2» необходимо разработать универсальный сервис.

Схема двусторонней асинхронной интеграции представлена на Рисунке 5. Реализованный универсальный сервис на стороне «Системы 1» и «Системы 2» должен по формату совпадать с универсальным интерфейсом на стороне ЕИП.

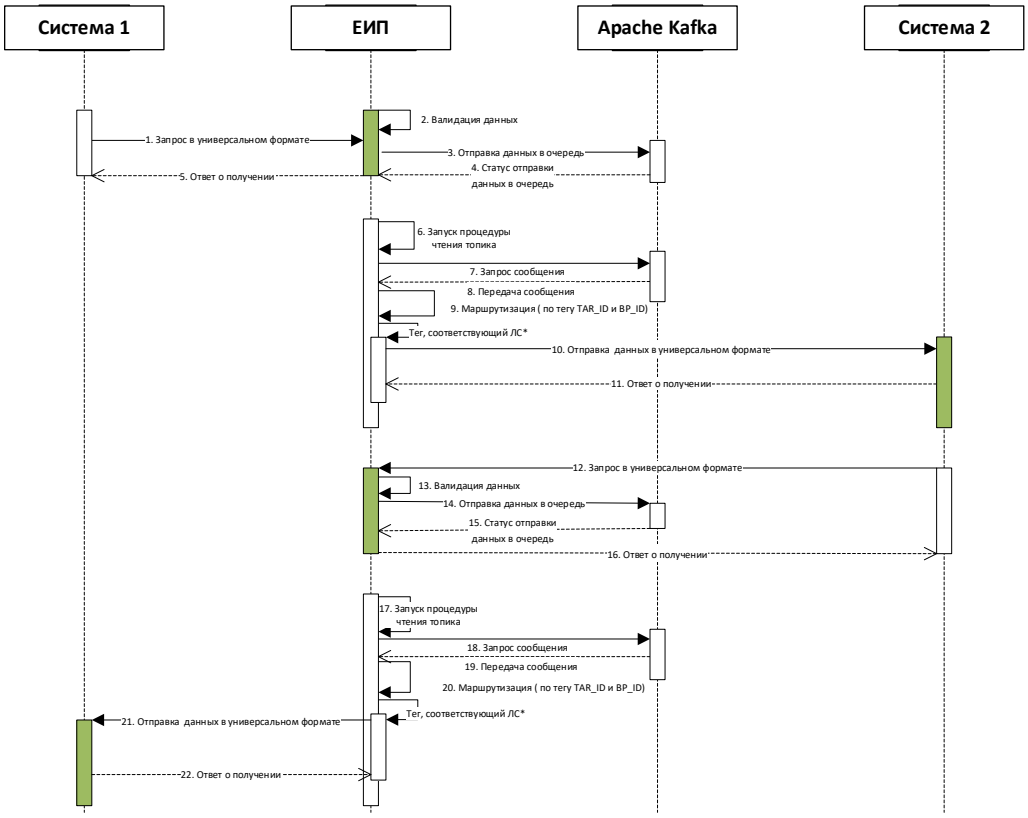


Рисунок 5. Схема взаимодействия между системами при двусторонней асинхронной интеграции

3.1.3 Синхронная интеграция

Синхронная интеграция разрабатывается только при выполнении следующих условий:

1) отсутствует необходимость в гарантированной доставке;

2) ответ от целевой системы нужен мгновенно, а целевая система технически может сформировать ответ за короткий промежуток времени.

Схема синхронной интеграции представлена на Рисунке 6. Реализованный универсальный сервис на стороне «Система 2» должен по формату совпадать с универсальным интерфейсом на стороне ЕИП.

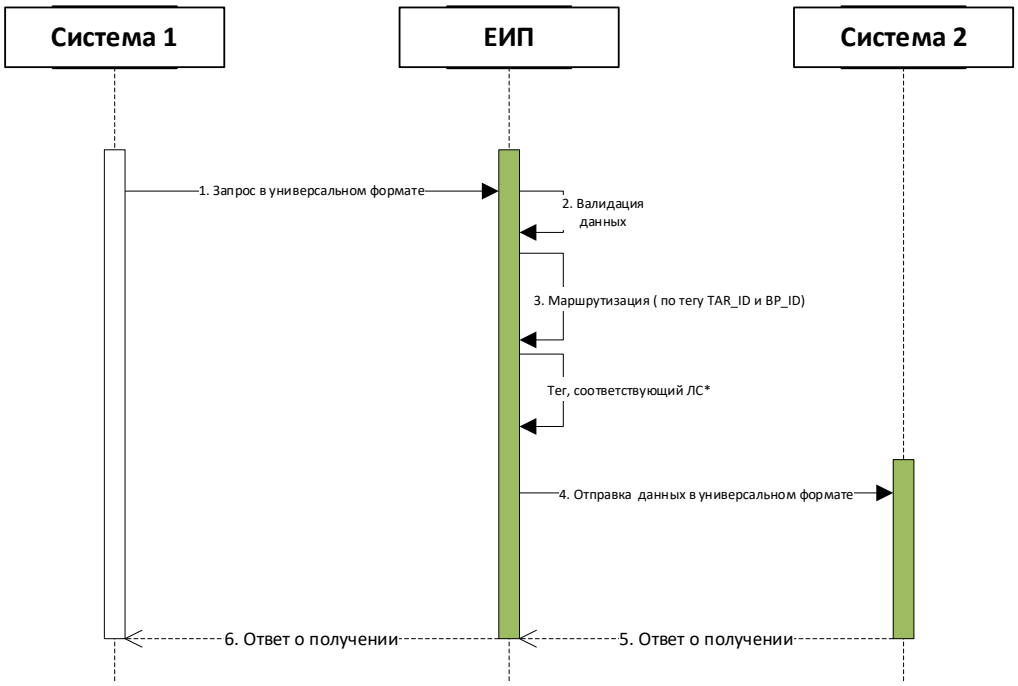


Рисунок 6. Схема взаимодействия между системами при синхронной интеграции

При наличии на стороне интегрируемых систем уже готовых интерфейсов по формату, не соответствующих универсальному сервису, но обеспечивающих обмен данными для решения поставленной задачи, техническое решение прорабатывается индивидуально.

4 Интеграция с холодным слоем Озера данных

4.1 Получение сведений о данных, хранящихся в Озере данных

Для того, чтобы ознакомиться с хранящимися в холодном слое Озера данных потоками интеграции между ИС, можно воспользоваться следующими сервисами:

- Реестр потоков — предоставляет в табличном виде список всех проходящих через ЕИП и загруженных в Озеро данных потоков интеграции. Технически сервис реализован в виде отчета системы Apache Superset.
- Граф потоков — также предоставляет информацию обо всех проходящих через ЕИП и загруженных в Озеро данных потоков интеграции, но в графическом виде. Технически сервис реализован в виде WEB-приложения Flow Graph.

- Каталог потоков — предоставляет информацию об атрибутах потоков интеграции, проходящих через ЕИП и загруженных в холодный слой Озера данных. Позволяет просматривать в иерархическом виде структуры XML и JSON сообщений с комментариями к потоку и каждому атрибуту. Технически сервис реализован на базе системы OpenMetadata.

Для получения доступа к сервисам можно обратиться через КСА СПП в отдел разработки прикладного программного обеспечения.

4.2 Запись данных в Озеро данных

4.2.1 Неявная запись данных

Все информационные потоки, проходящие через ЕИП, автоматически отправляются в Озеро данных и сохраняются в холодном слое. Это относится и к потокам в универсальном формате, и к проксирующим ЕИП потокам. Описание взаимодействия с ЕИП для обоих типов потоков данных приведено в документе «Требования к интеграции с ЕИП». Схема взаимодействия с Озером данных при отправке сообщений информационными системами - источниками в информационные системы - приемники отображена на рисунке 7.

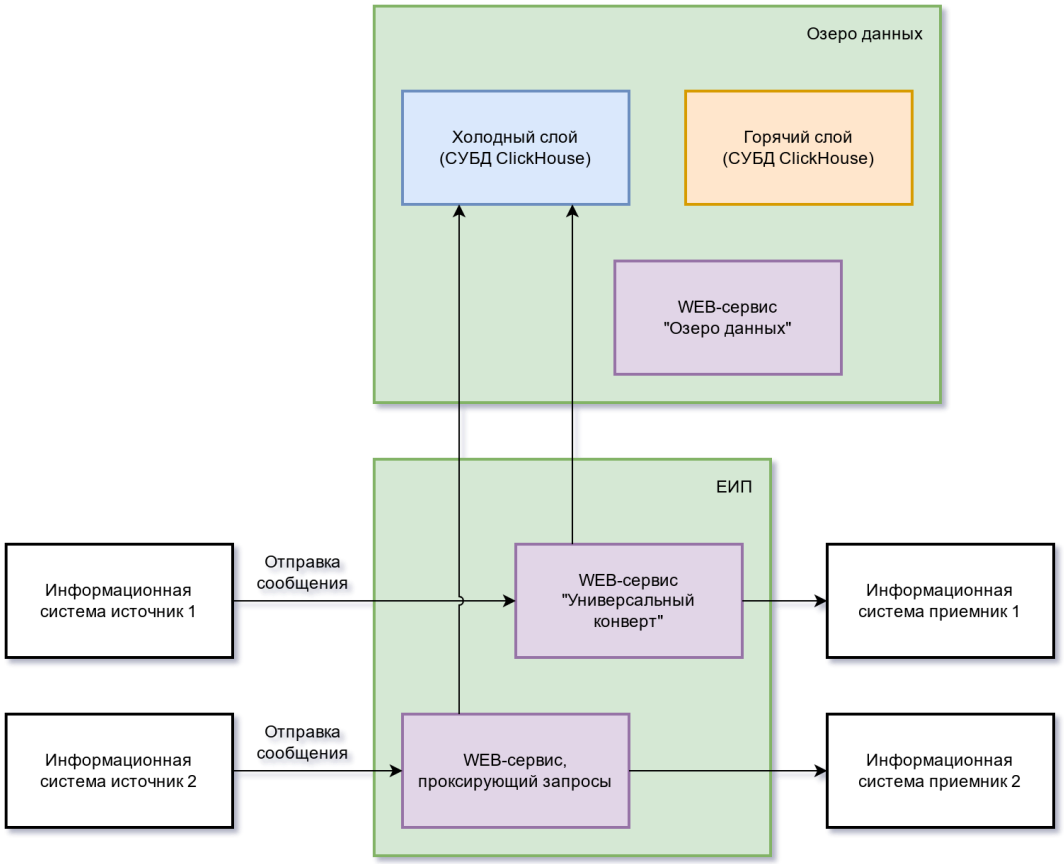


Рисунок 7. Схема взаимодействия с Озером данных при отправке сообщений информационными системами - источниками в информационные системы - приемники

4.2.2 Явная запись данных

Данные можно записать также только в Озеро данных без отправки в другую информационную систему. Для этого нужно отправить запрос в WEB-сервис ЕИП, принимающий данные в универсальном формате, указав Озеро данных в качестве информационной системы - приемника в атрибуте тела запроса (TAR_ID = dl). Описание процесса отправки сообщения в ЕИП приведено в документе «Требования к интеграции с ЕИП». Схема взаимодействия с Озером данных при отправке сообщений информационными системами - источниками только в Озеро данных отображена на рисунке 8.

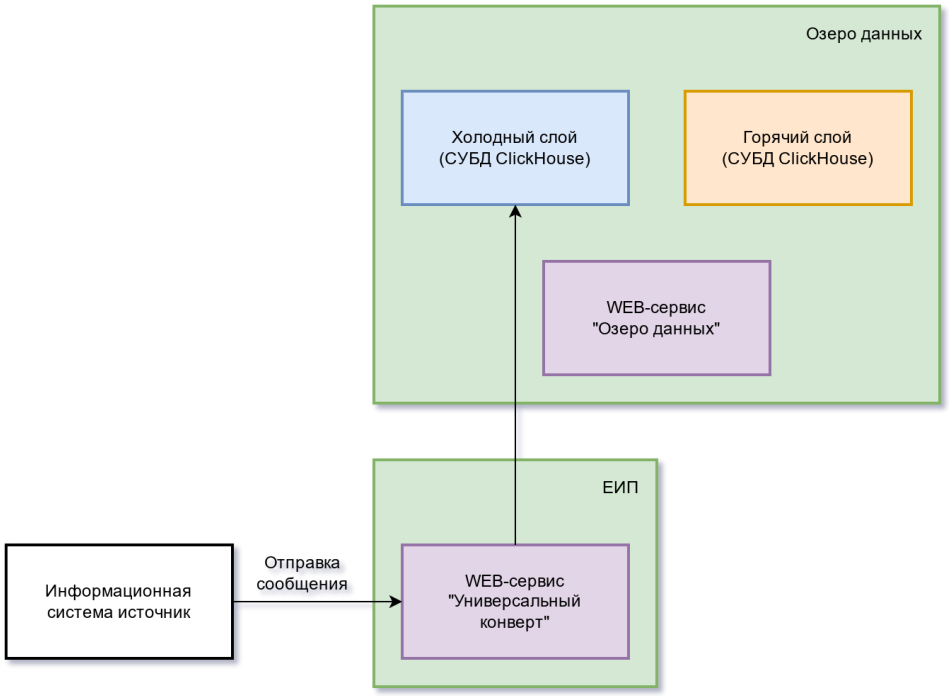


Рисунок 8. Схема взаимодействия с Озером данных при отправке сообщений информационными системами - источниками только в Озеро данных

4.3 Чтение данных из Озера данных

Чтение данных из холодного слоя Озера данных выполняется исключительно через WEB-сервис, развернутый в составе Озера данных. Сервис работает по протоколу HTTPS. Запрос выполняется в JSON-формате с указанием обязательных параметров:

- Дата и время начала периода поступления сообщений в Озеро данных
- Дата и время окончания периода поступления сообщений в Озеро данных
- Код бизнес-процесса

- Код системы-источника или системы-приемника

Коды бизнес-процесса, системы-источника и системы-приемника совпадают с кодами, приведенными в документе «Требования к интеграции с ЕИП».

Для оптимизации нагрузки на Озеро данных период запрашиваемых данных ограничивается 24 часами. При необходимости получения данных за больший период можно выполнить несколько запросов к WEB-сервису.

Ответ WEB-сервиса может быть в формате XML, JSON или plain-текст в зависимости от того, в каком формате сообщения первоначально поступили в холодный слой Озеро данных через систему ЕИП.

При получении данных через WEB-сервис выполняется Basic-авторизация по доменной учетной записи.

Схема взаимодействия с Озером данных при получении информационными системами сообщений из холодного слоя отображена на рисунке 9.

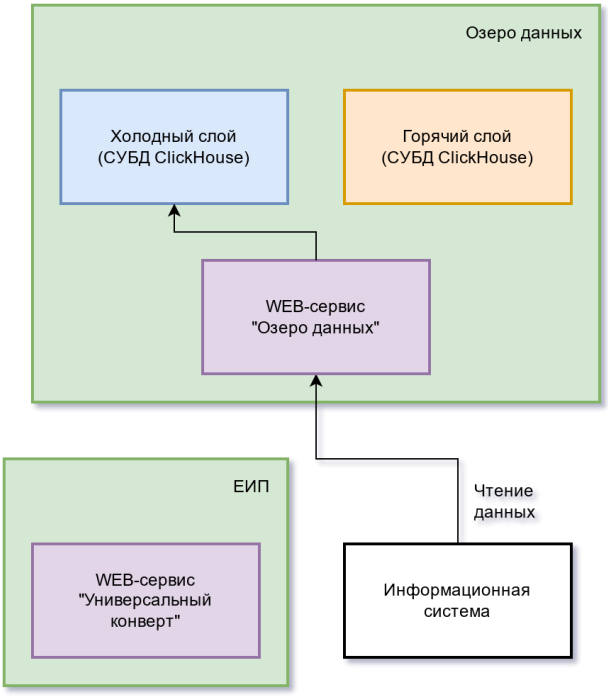


Рисунок 9. Схема взаимодействия с Озером данных при получении информационными системами сообщений из холодного слоя

5 Интеграция с горячим слоем Озера данных

5.1 Получение сведений о данных, хранящихся в Озере данных

Для того, чтобы ознакомиться с массивами данными, хранящимися в горячем слое Озера данных, можно воспользоваться следующим сервисом:

- Каталог данных — предоставляет информацию о составе и атрибутах наборов данных, хранящихся в горячем слое. Сервис позволяет просматривать список баз данных, таблиц, представлений, их полей с типами данных, а также комментарии к каждой таблице, представлению и их атрибутам. Технически сервис реализован на базе системы OpenMetadata.

Для получения доступа к сервису можно обратиться через КСА СПП в отдел разработки прикладного программного обеспечения.

5.2 Запись данных в Озеро данных

5.2.1 Прямая запись в базу данных ClickHouse

Запись информации в горячий слой Озера данных возможна с помощью прямого обращения к базе данных ClickHouse. Для такого взаимодействии создается учетная запись владельца конкретной базы данных и обращение в СУБД выполняется через нее. При необходимости учетной записи могут быть выданы права на несколько без данных, в основную — права владельца, а в остальные базы — на чтение. Схема взаимодействия с Озером данных при записи информации напрямую в базу данных СУБД ClickHouse отображена на рисунке 10.

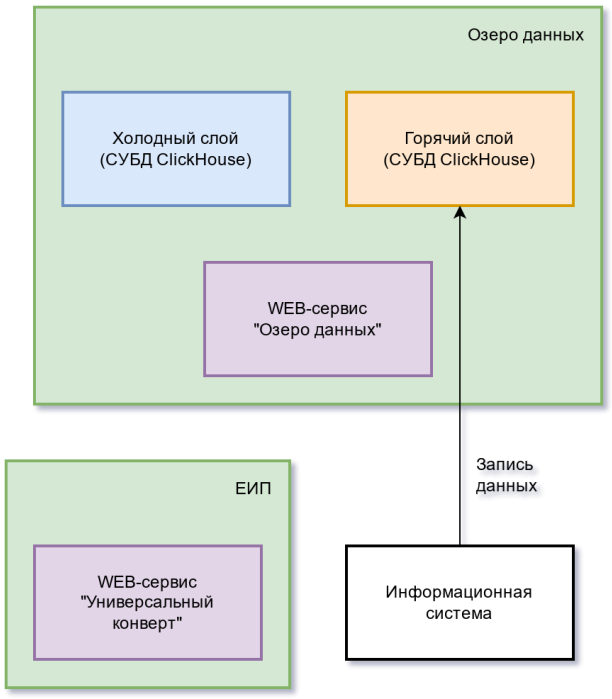


Рисунок 10. Схема взаимодействия с Озером данных при записи информации напрямую в базу данных СУБД ClickHouse

5.2.2 Запись через WEB-сервис

Запись информации в горячий слой Озера данных также возможна через WEB-сервис, развернутый в составе Озера данных. Сервис работает по протоколу HTTPS. Запрос выполняется в JSON-формате с указанием обязательных параметров:

- База данных горячего слоя, в которую нужно записать данные
- Таблица базы данных горячего слоя, в которую нужно записать данные

Строки таблицы, которые нужно записать, передаются в теле запроса.

При отправке данных через WEB-сервис выполняется Basic-авторизация по доменной учетной записи.

Схема взаимодействия с Озером данных при записи информационными системами данных в горячий слой через WEB-сервис отображена на рисунке 11.

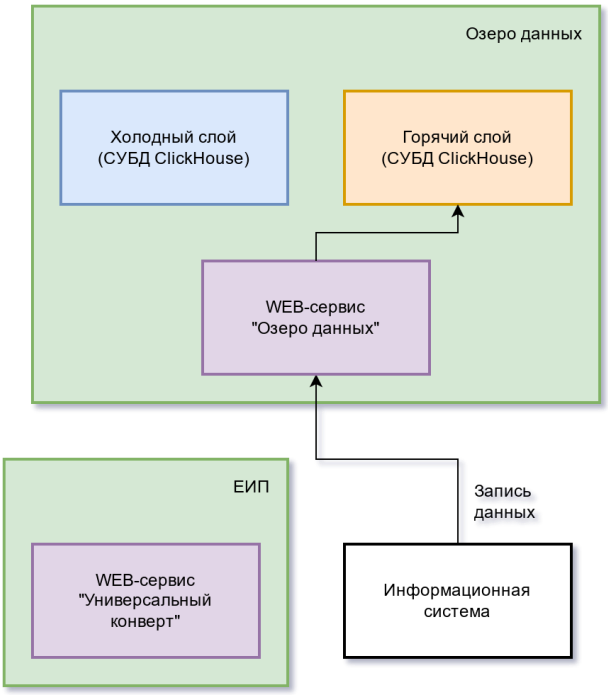


Рисунок 11. Схема взаимодействия с Озером данных при записи информационными системами данных в горячий слой через WEB-сервис

5.3 Чтение данных из Озера данных

5.3.1 Прямое чтение из базы данных ClickHouse

Чтение информации из горячего слоя Озера данных возможна с помощью прямого обращения к базе данных ClickHouse. Для такого взаимодействии создается учетная запись с правами только на чтение конкретной базы данных и обращение информационными системами в СУБД выполняется через нее. При необходимости учетной записи могут быть выданы права на чтение сразу для нескольких баз данных. Схема взаимодействия с Озером данных при чтении информации горячего слоя напрямую из базы данных СУБД ClickHouse отображена на рисунке 12.

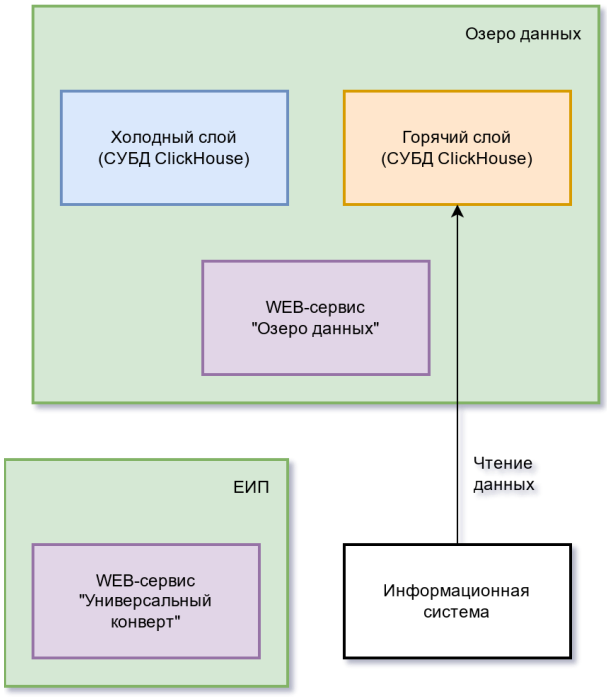


Рисунок 12. Схема взаимодействия с Озером данных при чтении информации горячего слоя напрямую из базы данных СУБД ClickHouse

5.3.2 Чтение через WEB-сервис

Чтение информации из горячего слоя Озера данных также возможна через WEB-сервис, развернутый в составе Озера данных. Сервис работает по протоколу HTTPS. Запрос выполняется в JSON-формате с указанием обязательных параметров:

- База данных горячего слоя, из которой нужно получить данные
- Таблица базы данных горячего слоя, из которой нужно получить данные
- Диапазон идентификаторов строк таблицы для получения

В ответе WEB-сервиса возвращаются запрошенные строки таблицы в формате JSON.

При запросе данных через WEB-сервис выполняется Basic-авторизация по доменной учетной записи. Схема взаимодействия с Озером данных при запросе информационными системами данных из горячего слоя через WEB-сервис отображена на рисунке 13.

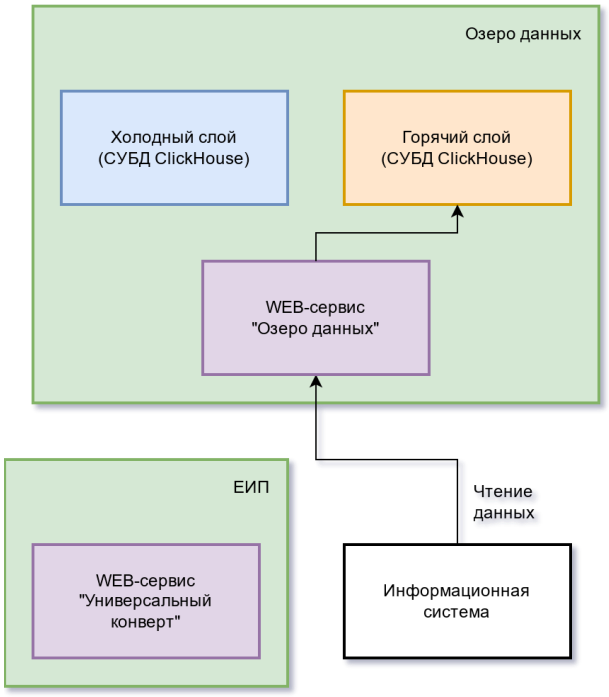


Рисунок 13. Схема взаимодействия с Озером данных при запросе информационными системами данных из горячего слоя через WEB-сервис

6 Требования к техническим средствам

6.1 Техническое обеспечение серверной части

Техническое обеспечение серверной части Системы состоит из технических средств вычислительного комплекса, системы хранения данных, системы мониторинга и системы резервного копирования.

6.1.1 Решения по техническим средствам клиентской части Системы

Специальных требований не предъявляется.

6.1.2 Решения по техническим средствам серверной части Системы

Техническое обеспечение вычислительного комплекса серверной части должно отвечать требованиям, приведенным в следующей таблице (см. таблицу 1).

Таблица 1 — Требования к техническим средствам серверов вычислительного комплекса

Назначение	**CPU cores**	**RAM, Гб**	**SSD, Гб**	**HDD2, Гб**	**ОС**	**СУБД и приложения**
Prod - WSO2 Enterprise Integrator	8	24	-	45	2 сервера	ssh, EI
Prod - Брокер сообщений	8	8	500	-	3 сервера	ssh, Kafka Message Broker
Prod - СУБД	4	16	-	500	1 сервер	ssh, Tantor 14
Prod - API менеджер	2	8	-	64	2 сервера	ssh, APIM
Prod - ClickHouse 01	8	128	100	500	AlterOS 7.5	ssh, yum-utils
Prod - ClickHouse 02	8	128	100	500	AlterOS 7.5	ssh, yum-utils
Prod - БД для сервисов FlowGraph, OpenMetadata, Superset	4	16	200		AlterOS 7.5	ssh, Tantor 14
Prod - Apache Nifi (ETL инструмент)	8	16	200		AlterOS 7.5	ssh, Java 8 или Java 11
Prod - Flow Graph (Веб приложение "Граф интеграций")	2	4		100	AlterOS 7.5	ssh, Docker 20.10, Nginx
Prod - KeyCloak (Авторизация\аутентификация для Flow Graph)	2	4		100	AlterOS 7.5	ssh, Java 8 или Java 11
Prod - Apache Superset (BI система)	8	16	200		AlterOS 7.5	ssh, Docker 20.10, Nginx

Prod - OpenMetadata (Каталог данных)	8	16	200		AlterOS 7.5	ssh, Docker 20.10, Nginx
Prod - Apache Airflow (ETL инструмент)	10	40	50		AlterOS 7.5	ssh, Docker 20.10
Prod - Grafana Loki	4	8	400		AlterOS 7.5	ssh, Docker 20.10
Prod - TestY (TMS инструмент)	4	8	50		AlterOS 7.5	ssh, Docker 20.10
Prod - JSReports	4	8	50		AlterOS 7.5	ssh, Docker 20.10
QA - WSO2 Enterprise Integrator	4	8	-	45	2 сервера	ssh, EI
QA - Брокер сообщений	8	8	-	500	3 сервера	ssh, Kafka Message Broker
QA - СУБД	4	8	-	50	1 сервер	ssh, Tantor 14
QA - WSO2 API Manager	4	8	-	45	2 сервера	ssh, APIM
QA - ClickHouse 01	4	64	100	500	AlterOS 7.5	ssh, yum-utils
QA - ClickHouse 02	4	64	100	500	AlterOS 7.5	ssh, yum-utils
QA - БД для сервисов FlowGraph, OpenMetadata, Superset	4	16	200		AlterOS 7.5	ssh, Tantor 14
QA - Apache Nifi (ETL инструмент)	8	16	200		AlterOS 7.5	ssh, Java 8 или Java 11

QA - Flow Graph (Веб приложение "Граф интеграций")	2	4		100	AlterOS 7.5	ssh, Docker 20.10, Nginx
QA - KeyCloak (Авторизация\аутентификация для Flow Graph)	2	4		100	AlterOS 7.5	ssh, Java 8 или Java 11
QA - Apache Superset (BI система)	8	16		150	AlterOS 7.5	ssh, Docker 20.10, Nginx
QA - OpenMetadata (Каталог данных)	8	16		150	AlterOS 7.5	ssh, Docker 20.10, Nginx
QA - Apache Airflow (ETL инструмент)	4	8	50		AlterOS 7.5	ssh, Docker 20.10
QA - Grafana Loki	2	4	50		AlterOS 7.5	ssh, Docker 20.10
QA - TestY (TMS инструмент)	2	4	50		AlterOS 7.5	ssh, Docker 20.10
QA - JSReports	2	4	50		AlterOS 7.5	ssh, Docker 20.10

Требования к техническим средствам системы хранения данных и системы резервного копирования определяются в соответствии с положениями нормативно-технической и организационно-распорядительной документации.

Серверы приложений Системы объединены в кластер с балансировкой нагрузки.

Серверы Системы должны быть объединены локальной вычислительной сетью с пропускной способностью не менее 100 Мбит/сек.

6.1.3 Решения по размещению комплекса технических средств

Система подготовлена и развернута на комплексе технических средств в соответствии со следующими основными требованиями:

- Функционирование и эксплуатация Системы осуществляется на следующих экземплярах Системы:
- Тестовый стенд — предназначен для проведения испытаний Системы.

- Промышленный стенд — предназначен для постоянной (промышленной) эксплуатации Системы.
- Вычислительные средства и мощности экземпляров Системы предоставлены в виде виртуальных машин.
- Технические средства базируются на применении серийно выпускаемого оборудования и изделий, обеспечивающих исправное функционирование Системы при заданных показателях надежности.
- Средства виртуализации доступны на российском рынке.

Таблица 2 — Перечень внешних компонентов Системы

Система	**Описание**
Kafka Message Broker	Брокер сообщений
EIP API	API сервис для предоставления справочных данных ЕИП
Zabbix	Система мониторинга сервисов
MS Active Directory	Система для авторизации и аутентификации на основе протокола LDAP

Таблица 3 — Перечень внутренних компонентов Системы

Компонент	**Описание**	**Рекомендуемая версия**
Apache Nifi	ETL\ELT инструмент для обработки событий (сообщений), предоставляющий возможности управления потоками данных из разнообразных источников в режиме реального времени с использованием графического интерфейса.	1.22.0 (12.06.2023)
Yandex ClickHouse	Колоночная аналитическая СУБД, позволяющая выполнять аналитические запросы в режиме реального времени на структурированных больших данных.	23.3.3.52-lts (12.06.2023)

Компонент	**Описание**	**Рекомендуемая версия**
Open Metadata	Платформа для управления метаданными. Предоставляет набор функций для импорта, хранения и управления метаданными, включая глоссарий, схемы данных, таблицы, столбцы и связи между ними.	1.0.4 (08.06.2023)
Apache Superset	BI система, для исследования и визуализации данных, ориентированная на большие данные.	5.0.0 (23.06.2025)
Flow Graph	Веб-приложение для визуализации Графа интеграций. Для визуализации графа используется open source библиотека vis.js	vis-network 9.1.6(23.03.2023) React JS 18.2.0(14.06.2022) react-router-dom 6.8.1(6.02.2023) Ant design 4.23.6(17.10.22) React Redux 8.0.2(22.05.2022)
Tantor	СУБД, разработанная на основе PostgreSQL	Tantor-BE-14 (postgresql 14.6)