

# Эксплуатация Arenadata Streaming (Kafka, NiFi)



# Обзор Arenadata Enterprise Data Platform и Arenadata Streaming

# Agenda

- Экосистема EDP: Arenadata Streaming, Arenadata DB, Arenadata QuickMarts, Arenadata Postgres, Arenadata Cluster Manager
- Arenadata Streaming: основные компоненты и назначение (Kafka Broker, Kafka Connect Worker, Kafka-Manager, Kafka REST Proxy, ksqlDB Client, ksqlDB, Schema-Registry, NiFi Registry, NiFi Server, MiNiFi, ZooKeeper)
- Arenadata Streaming Control
- Arenadata Platform Security

Экосистема EDP:  
Arenadata Streaming,  
Arenadata Hadoop,  
Arenadata DB, Arenadata  
QuickMarts, Arenadata  
Postgres, Arenadata Cluster  
Manager

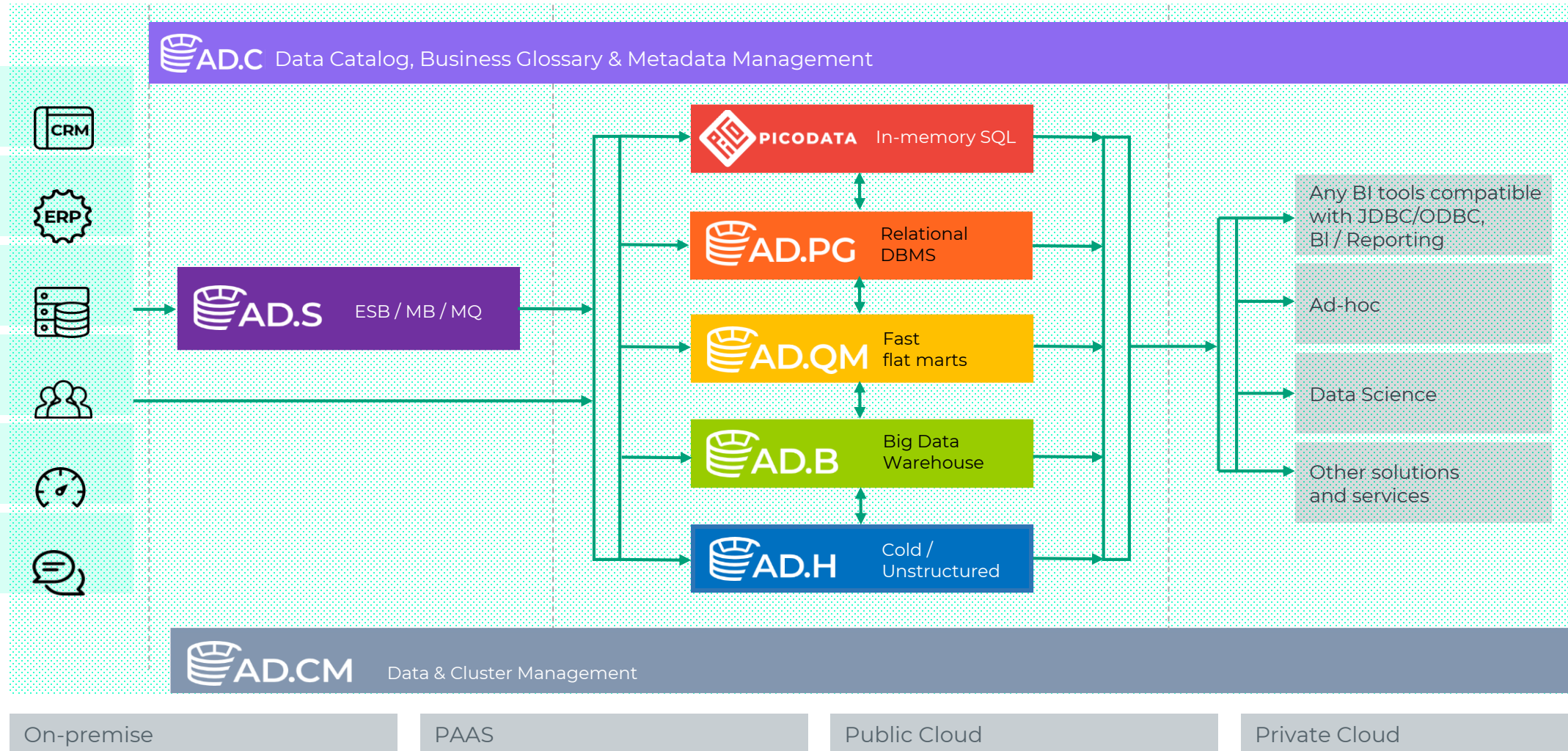
# Arenadata Enterprise Data Platform

ИСТОЧНИКИ

ТРАНСПОРТ

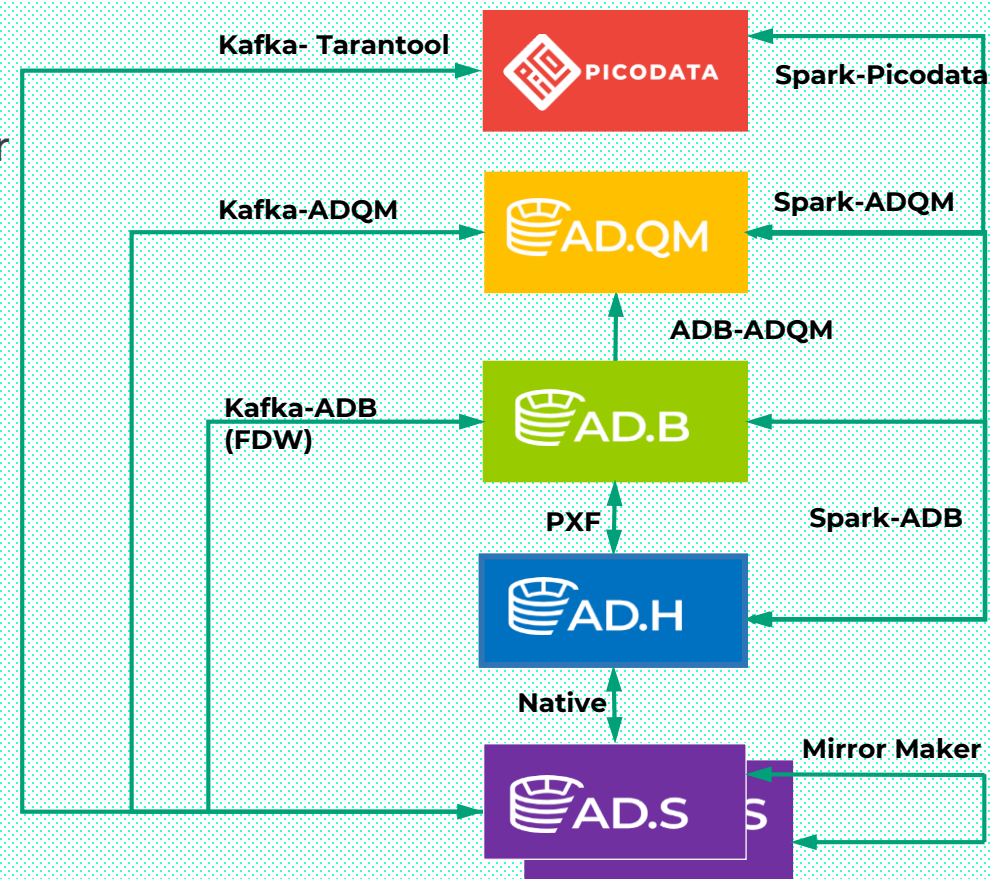
ХРАНЕНИЕ И ПРЕДСТАВЛЕНИЕ ДАННЫХ

ИСПОЛЬЗОВАНИЕ И ВИЗУАЛИЗАЦИЯ



# Коннекторы в платформе

- Kafka-Arenadata DB Connector FDW
- Nifi PutElasticSearch processor
- Arenadata DB – Arenadata QuickMarts Thkemali Connector
- Kafka-Tarantool (Picodata) Connector
- Spark-Tarantool (Picodata) Connector
- Spark-Arenadata DB Connector
- Spark-Arenadata QuickMarts Connector
- Поддержка Kerberos для PXF
- Поддержка Informatica DEI 10.4 для ADH 2.X
- Mirror Maker 2



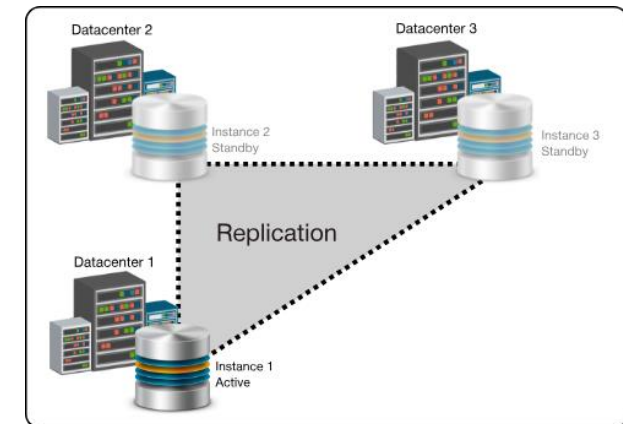
# In-memory data grid



Picodata (PD) — система хранения и обработки информации для задач реального времени в составе единой корпоративной платформы данных Arenadata. Используется для высоконагруженных приложений и интенсивной работы с данными (свыше 10 000 транзакций в секунду)

Ключевые возможности PD:

- использование непосредственно как БД с хранением данных в оперативной памяти (IMDB); распределенное хранилище объектов в памяти с доступом по объектным ключам (IMDG);
- производительность от 10 000 запросов в секунду;
- совместимость с любыми инструментами и средами развёртывания экземпляров (Ansible, K8s, Podman и др.);
- горизонтальное масштабирование кластера;
- объем обрабатываемых данных: 2-100 Тб;
- гарантия сохранности и консистентности данных.



# Транзакционная СУБД



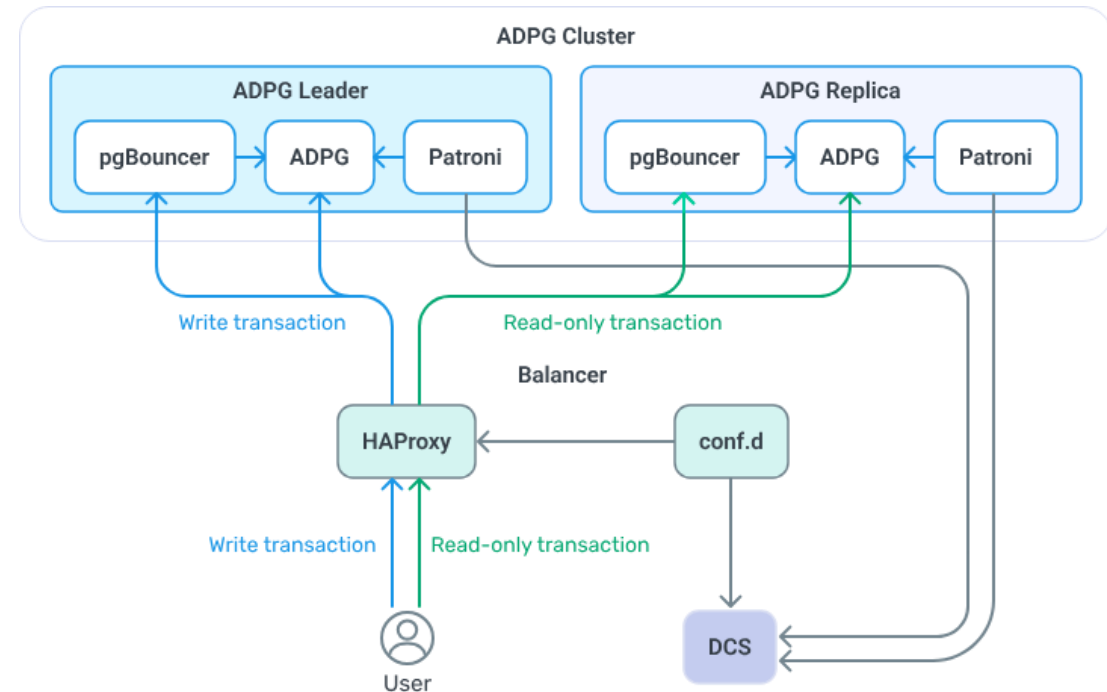
OLTP  
DataBase



Arenadata Postgres — транзакционная российская СУБД на базе объектно-реляционной базы PostgreSQL

Ключевые возможности ADPG:

- сложные запросы, обрабатывающие большие объемы данных и включающие сложные аналитические функции (в том числе аналитику поведения пользователей);
- разнообразные виды отчетности (оперативная, управленческая, обязательная нормативная);
- глубокий анализ AD-HOC;
- эффективное объединение больших таблиц;
- интеграция с системами маркетинговых компаний и системами лояльности;
- полнотекстовый поиск;
- прогнозная аналитика (спрос, остатки).





# Витрины данных



Arenadata QuickMarts (ADQM) — кластерная колоночная система управления базами данных, созданная на основе Yandex ClickHouse.

Может в режиме реального времени генерировать аналитические отчёты разного плана, используя большие объёмы информации, хранящейся в плоских витринах. ADQM многократно быстрее традиционных СУБД.

## Основные возможности Arenadata QuickMarts:

- В связке с Arenadata DB / Greenplum
- Аналитика над структурированными логами и событиями
- Быстрые витрины
- Распределённые вычисления для бизнеса
- Анализ серверных логов

# Хранилище данных (МРР-СУБД)



Arenadata DB (ADB) — аналитическая, распределённая СУБД с открытым исходным кодом, использующая концепцию МРР (massively parallel processing), построенная на базе СУБД Greenplum.

Предназначена для хранения и обработки больших объёмов данных — до десятков петабайт.

## Основные возможности Arenadata DB:

- Row/Column-хранение данных, с возможностью сжатия
- Сложные запросы с большими объёмами данных
- ETL/ELT
- Эффективное соединение больших таблиц
- Data Science
- Аналитические функции на процедурных языках
- Ad-hoc аналитика

# Хранение холодных и слабо структурированных данных



- Arenadata Hadoop (ADH) — это готовый к использованию дистрибутив, собранный на базе множества Open Source — проектов экосистемы Apache Hadoop с предустановленным оркестратором ИТ-ландшафта.
- ADH предназначен:
  - для работы с любыми типами и форматами данных и включает в себя все необходимые для работы компоненты: управление, доступ, анализ, интеграция, безопасность и администрирование;
  - для построения распределённой платформы хранения на базе Apache Hadoop, адаптированный для корпоративного использования для холодных и слабо структурированных данных, которые хранить на MPP-СУБД или проприетарных СУБД было бы слишком дорого.

# Streaming (ESB / MB / MQ)



Arenadata Streaming (ADS) — масштабируемая отказоустойчивая система для потоковой обработки данных в режиме реального времени, адаптированная для корпоративного использования и построенная на базе Apache Kafka и Apache Nifi.

ADS Control — предоставляет возможность создавать, редактировать и удалять коннекторы Kafka.

## Основные возможности Arenadata Streaming:

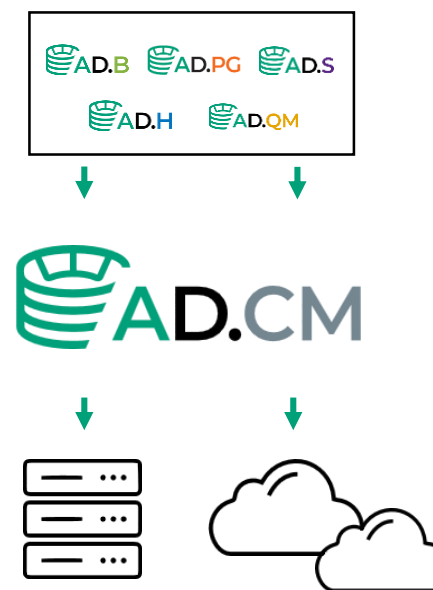
- Корпоративная шина обмена данными для всех приложений
- Среда консолидации больших потоков данных и управления ими в режиме реального времени
- Инструмент разграничения прав доступа к потокам данных
- Фреймворк (NiFi) для разработки потоковых аналитических приложений за считанные минуты в режиме реального времени и без единой строчки кода.

# Arenadata Cluster Manager



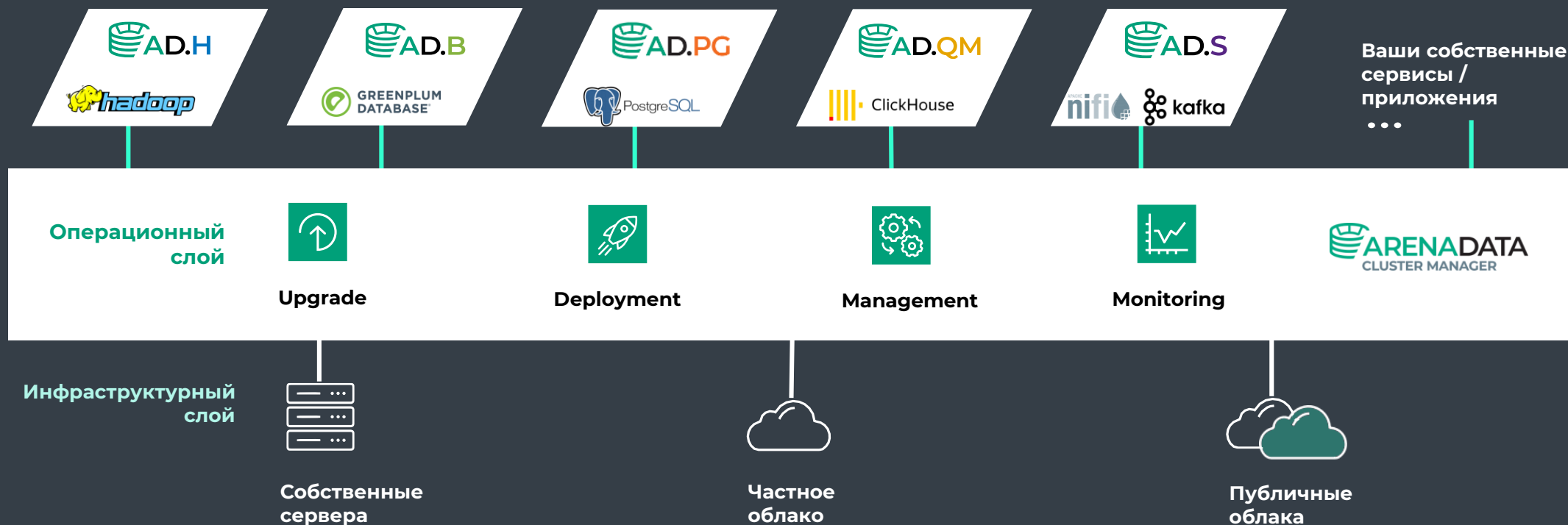
**Arenadata Cluster Manager (ADCM)** – универсальный оркестратор гибридного ландшафта. Позволяет быстро устанавливать, настраивать все data-сервисы и управлять ими независимо от инфраструктуры.

- В Arenadata Cluster Manager установка, настройка и обновление кластеров в оркестраторе производятся по нажатию кнопки в графическом интерфейсе или по запросу в API.
- Конфигурирование ОС, сервисов, сети
- Настройка прав безопасности в ADCM
- Открытый ADCM API
- Открытый формат для создания новых бандлов – добавьте ваши собственные сервисы/инфраструктуру



# Гибридный корпоративный ландшафт

Слой компонент данных



# Arenadata Streaming

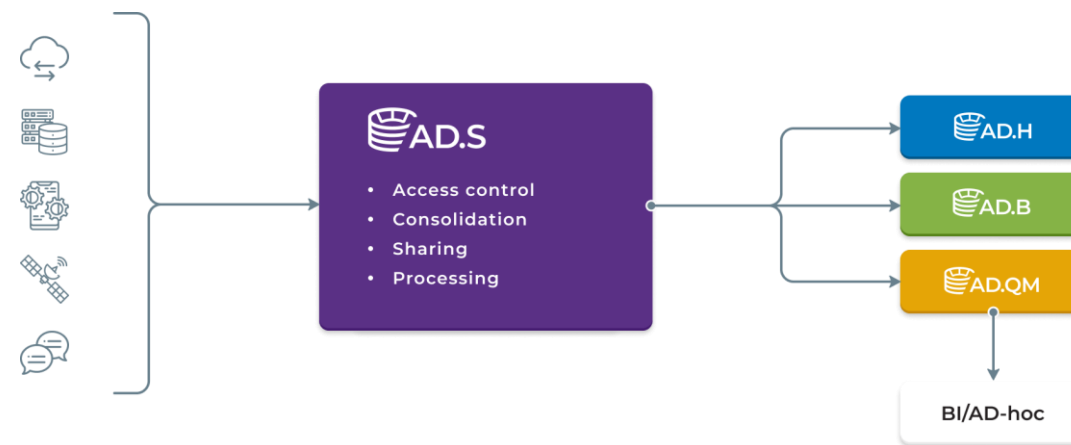


**Arenadata Streaming (ADS) — масштабируемая отказоустойчивая система для потоковой обработки данных в режиме реального времени, адаптированная для корпоративного использования и построенная на базе Apache Kafka и Apache Nifi.**

**ADS является универсальным средством для решения задач, связанных с потоковой обработкой данных в режиме реального времени.**

## Используется как:

- Корпоративная шина обмена данными для всех приложений
- Среда консолидации больших потоков данных и управления ими в режиме реального времени
- Инструмент разграничения прав доступа к потокам данных
- Фреймворк для разработки потоковых аналитических приложений за считанные минуты в режиме реального времени и без единой строчки кода



# Arenadata Streaming: основные компоненты и назначение (Kafka, NiFi, ZooKeeper)

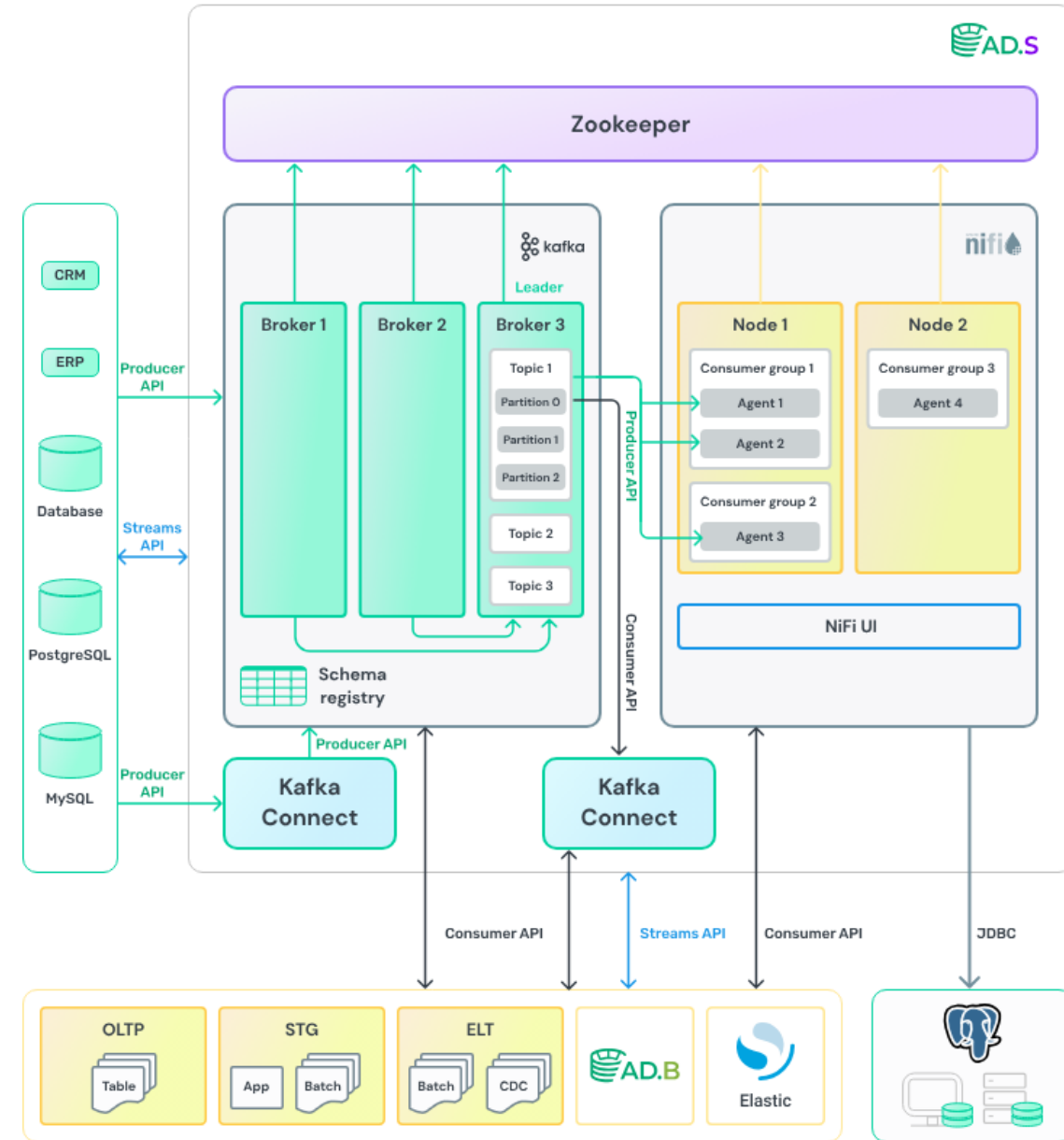


# Arenadata Streaming



## ADS позволяет:

- Строить потоковые конвейеры данных в реальном времени, которые надежно передают данные между системами или приложениями.
- Строить потоковые приложения в реальном времени, которые преобразуют или реагируют на потоки данных.
- Хранить потоки записей отказоустойчивым долговечным способом.



# Компоненты Arenadata Streaming



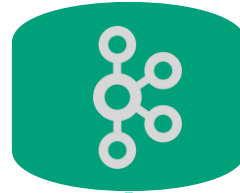
**Apache ZooKeeper** – это служба распределенной координации, используемая Arenadata Streaming для управления конфигурацией и координацией своих кластеров. Помогает обеспечить высокую доступность и отказоустойчивость в кластерах потоковой передачи Arenadata.



- NiFi Server 1.2X.X
- NiFi Registry 1.2X.X
- MiNiFi Agent 1.2X.X
- MiNiFi C2 Server 1.2X.X

## Security

- NiFi LDAP\LDAPS аутентификация
- SSL\TLS поддержка
- Авторизация и аудит в Ranger



- Kafka 3.X.X
- ~~Kafka Manager 3.0.0.5~~
- Kafka REST Proxy 7.X.X
- KSQL 7.X.X
- Schema-Registry 7.X.X
- Kafka Connect 3.X.X

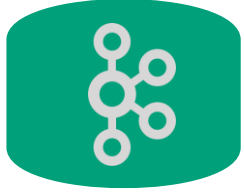
## Security

- Kerberos MIT KDC, AD, FreeIPA
- SSL\TLS поддержка
- Авторизация и аудит в Ranger

## ADSC

UI Managing Kafka Connectors  
Mirror Maker 2

# Компоненты Arenadata Streaming. Kafka

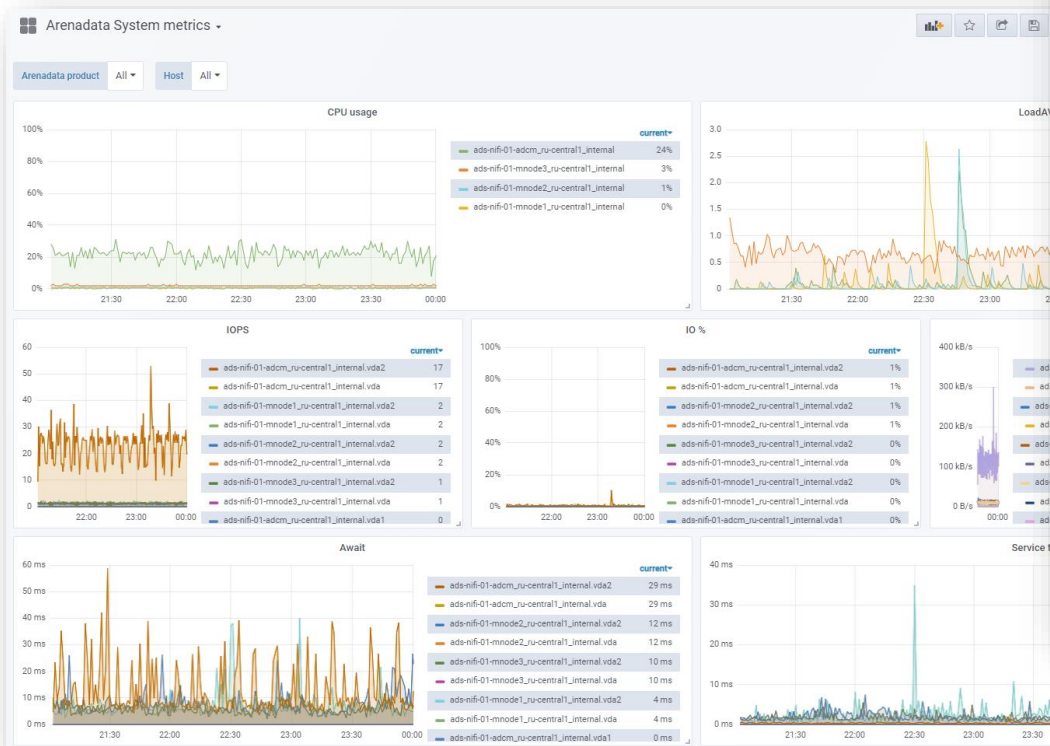


- **Apache Kafka** – это распределенная платформа потоковой передачи, используемая Arenadata Streaming для управления приемом, обработкой и анализом потоков данных в реальном времени.
- **Schema Registry** – это централизованный репозиторий, используемый Arenadata Streaming для хранения и управления схемами данных, создаваемых и потребляемых Apache Kafka.
- **KSQL** – это потоковый механизм SQL, используемый Arenadata Streaming для обработки потоков данных в реальном времени, который позволяет пользователям писать SQL-запросы для преобразования, агрегирования и анализа данных в режиме реального времени.
- **Kafka Connect** – это платформа интеграции данных, используемая Arenadata Streaming для перемещения данных между Apache Kafka и другими системами.
- **Kafka REST Proxy** – это инструмент, используемый Arenadata Streaming для предоставления функциональности Apache Kafka в виде RESTful API.

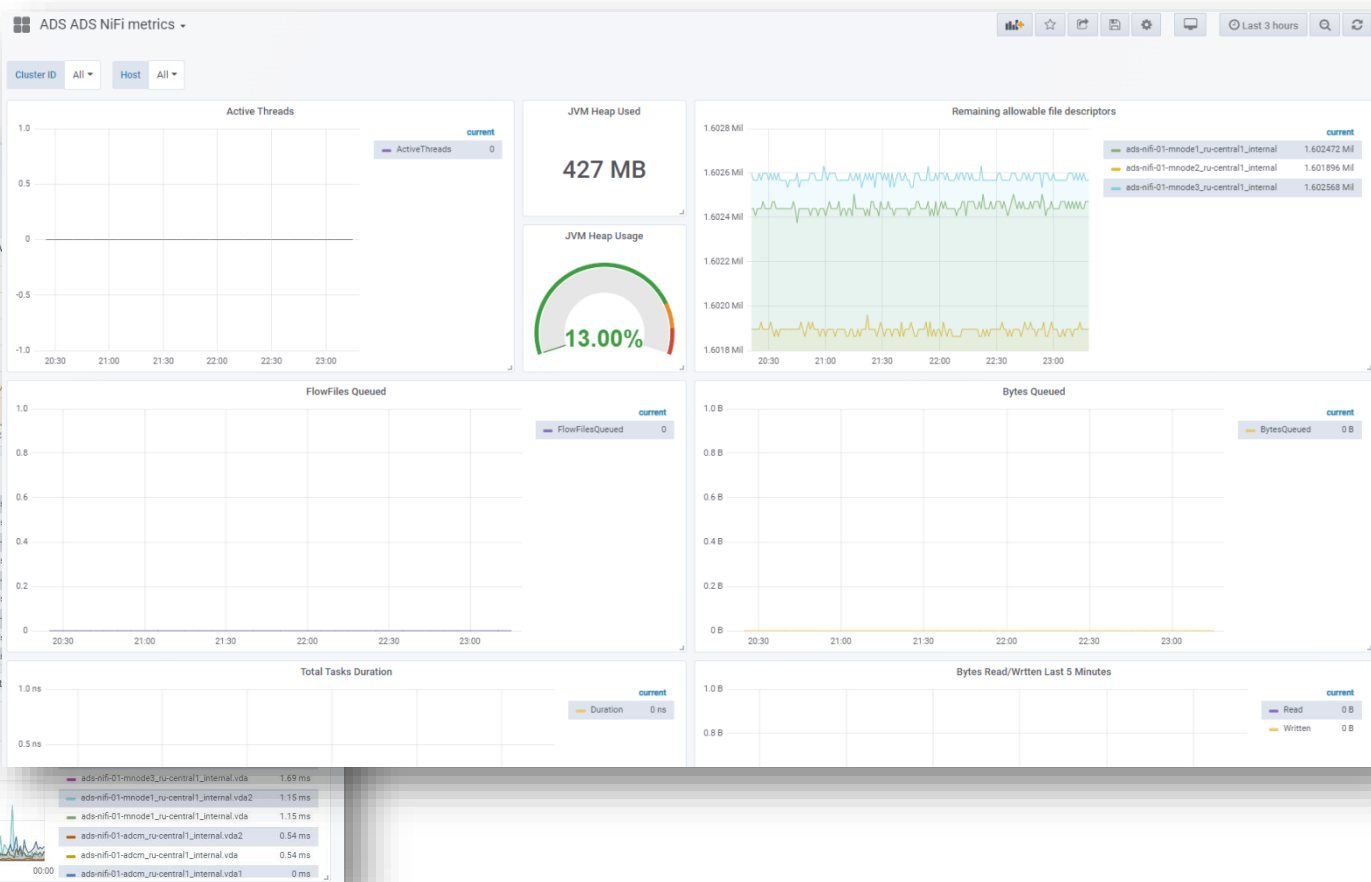
# Cluster Monitoring

Мониторинг в ADS реализован в два уровня:

1. Первый выполняет функции по сбору и хранению системных метрик на уровне хостов.



2. Второй выполняет функции по сбору и хранению метрик от Kafka и NiFi и отображению их в разработанных панелях мониторинга.



+ native-мониторинг на уровне приложений NiFi и Kafka

# ADS: CMAK (Cluster Manager for Apache Kafka)

- Управление кластерами Apache Kafka

CMAK Cluster

Clusters

Clusters

Active

ADS

click

Operations

Modify

Disable

Topics

Show 10 entries

Search:

Topic	# Partitions	# Brokers	Brokers Spread %	Brokers Skew %	Brokers Leader Skew %	# Replicas	Under Replicated %	Producer Message/Sec	Summed Recent Offsets
<a href="#">__consumer_offsets</a>	50	3	100	0	0	1	0	0.00	38
<a href="#">new-topic</a>	1	1	33	0	0	1	0	0.00	0
<a href="#">new-topic-1</a>	3	3	100	0	0	3	0	0.00	0

Showing 1 to 3 of 3 entries

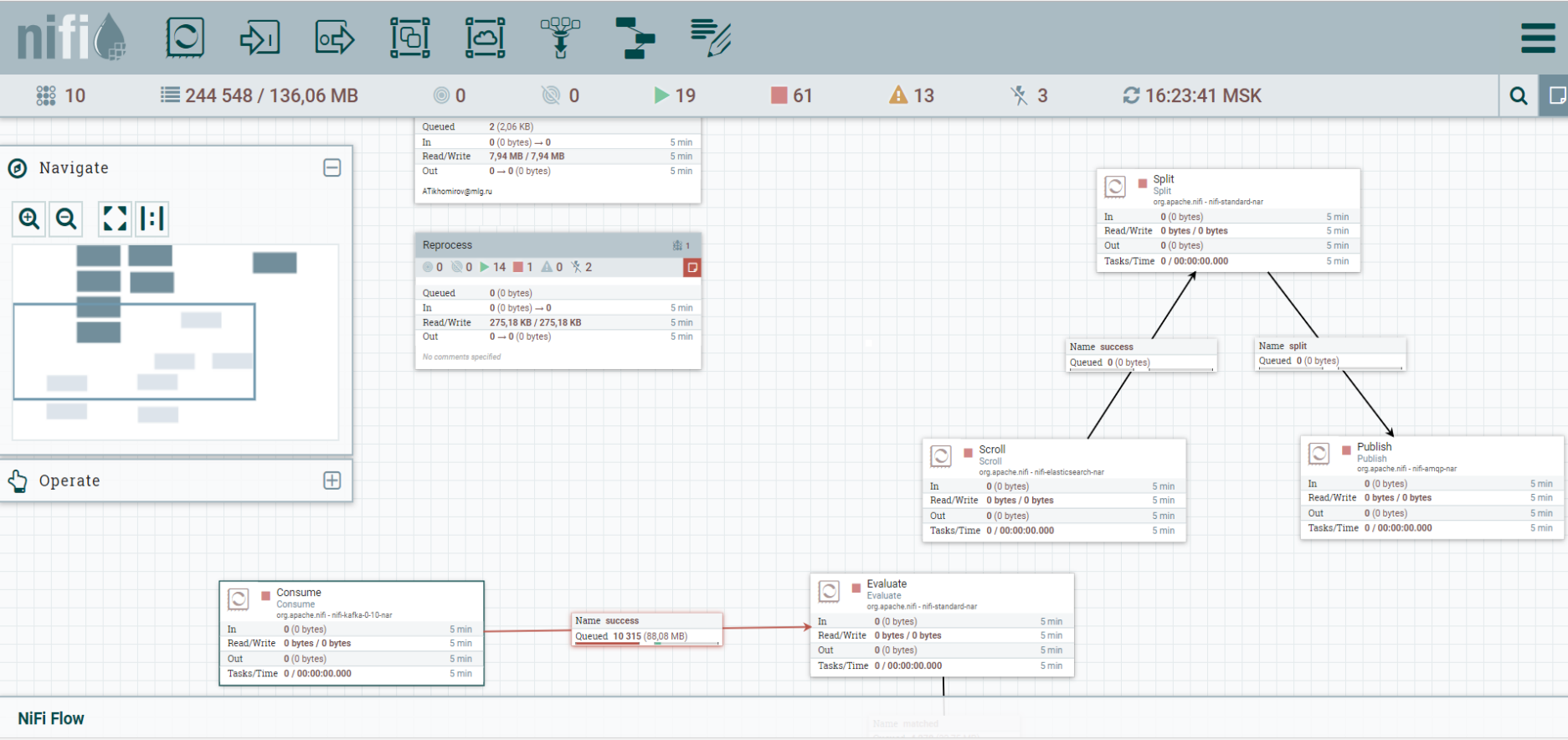
Previous1Next

# Компоненты Arenadata Streaming. NiFi



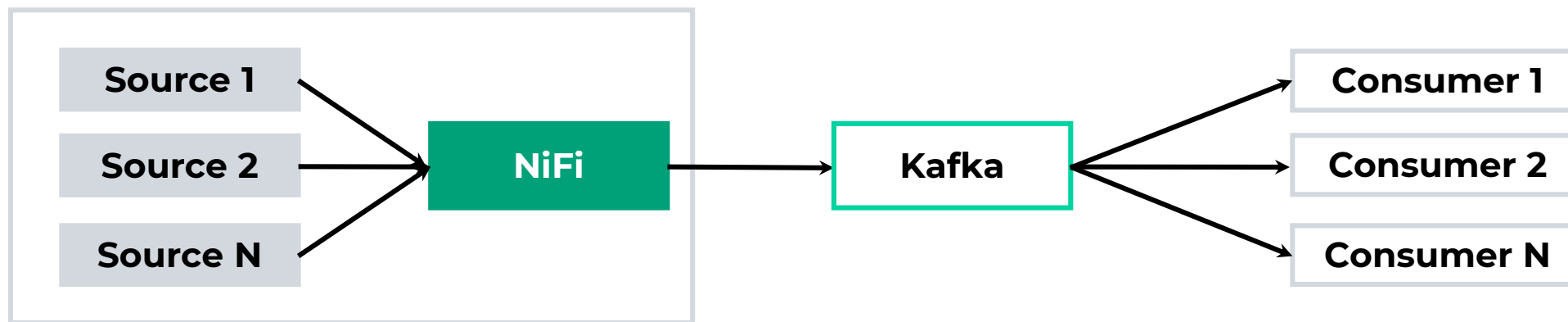
- **Apache NiFi** – это инструмент интеграции данных с открытым исходным кодом, используемый Arenadata Streaming для автоматизации потока данных между различными системами и технологиями.
- **Apache NiFi Registry** – подпроект Apache NiFi, представляющий инструмент для хранения Flow и управления версиями.
- **Apache MiNiFi** – подпроект Apache NiFi, представляющий легковесного агента, использующий процессоры NiFi без графического интерфейса для конфигурации flow.
- **Apache MiNiFi C2 Server** – подпроект Apache NiFi, являющийся центральной точкой конфигурации MiNiFi агентов и предоставления конфигурации через REST API.

# NiFi Server

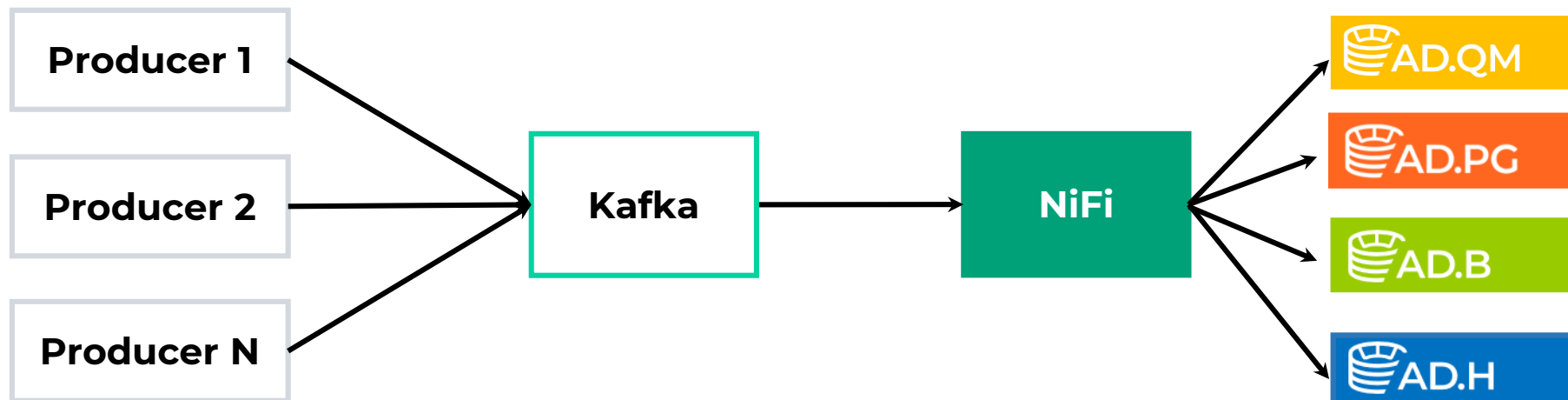


# NiFi Producer Consumer

- Collector&Producer



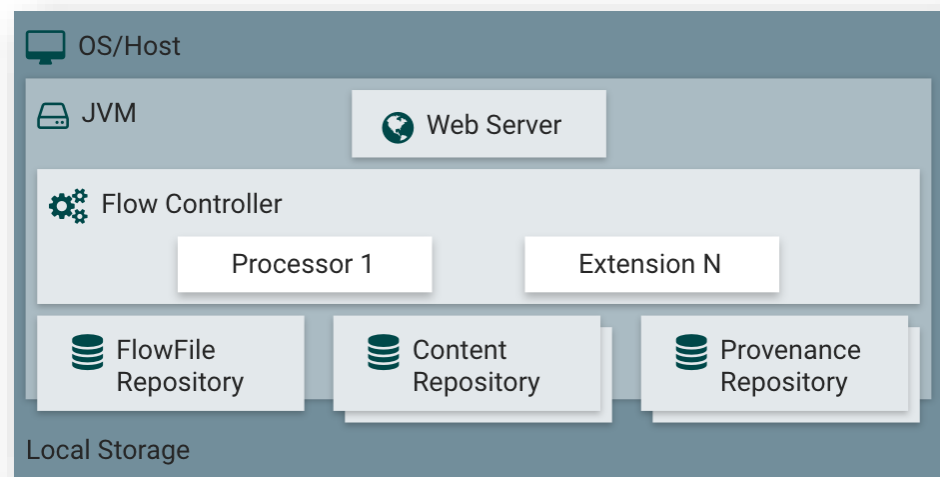
- Consumer





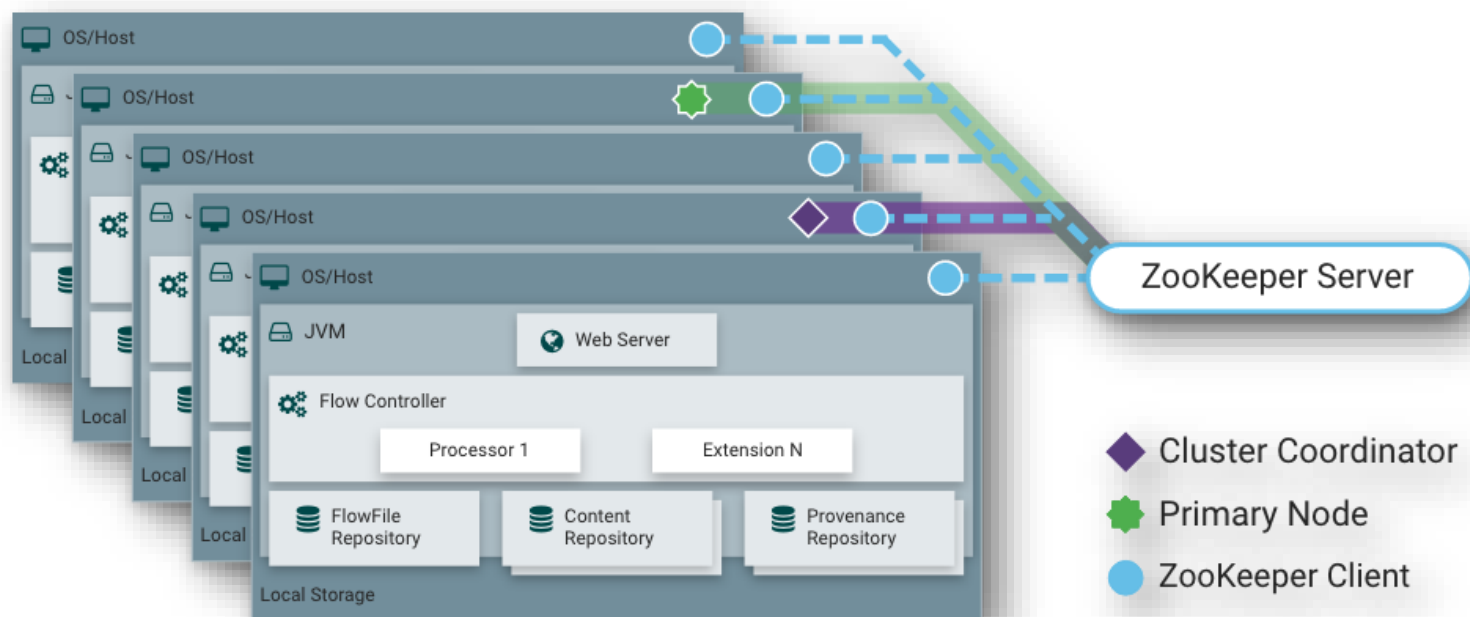
# NiFi Architecture

- **Web Server** — предоставляет веб-интерфейс и REST API.
- **Flow Controller** — предоставляет потоки для запуска расширений и управляет ресурсами для **Extensions** (Processors, Controller Services, Reporting Tasks, ...)
- **FlowFile Repository** — это то место, в котором NiFi хранит всю известную ему информацию о каждом существующем в данный момент FlowFile в системе. По умолчанию используется постоянный журнал упреждающей записи (Write-Ahead Log), расположенный в указанном разделе диска.
- **Content Repository** — репозиторий, в котором находится содержимое всех FlowFile, т.е. сами передаваемые данные и представляет собой механизм, который хранит блоки данных в файловой системе. Можно указать несколько мест хранения файловой системы, чтобы задействовать разные физические разделы.
- **Provenance Repository** — содержит историю(lineage) о каждом FlowFile. Каждый раз, когда с FlowFile происходит какое-либо событие (создание, изменение и т.д.), соответствующая информация заносится в этот репозиторий.
- **FlowFile** — сущность, представляющая собой объект: метаданные (атрибуты) + данные (контент может быть и пустым).



# NiFi Cluster Architecture

- Каждый узел в кластере NiFi выполняет одни и те же задачи с данными, но каждый работает с разным набором данных.
- Apache ZooKeeper выбирает один узел в качестве координатора кластера (**Cluster Coordinator**), а аварийное переключение выполняет автоматически ZooKeeper.
- Все узлы кластера передают Heartbeats координатору кластера. Координатор кластера отвечает за отключение и подключение узлов.
- В каждом кластере есть один первичный узел (**Primary Node**), также выбираемый ZooKeeper. На этом узле можно запускать изолированные процессоры, которые не должны запускаться на всех нодах одновременно (например, GetSFTP).
- Изменения через пользовательский интерфейс (UI) любого узла, реплицируются на все узлы кластера, что позволяет использовать несколько точек входа.
- Apache ZooKeeper сам выбирает **Cluster Coordinator** и **Primary Node**, в случае недоступности узлов.
- Есть возможность горизонтально масштабировать нагрузку.

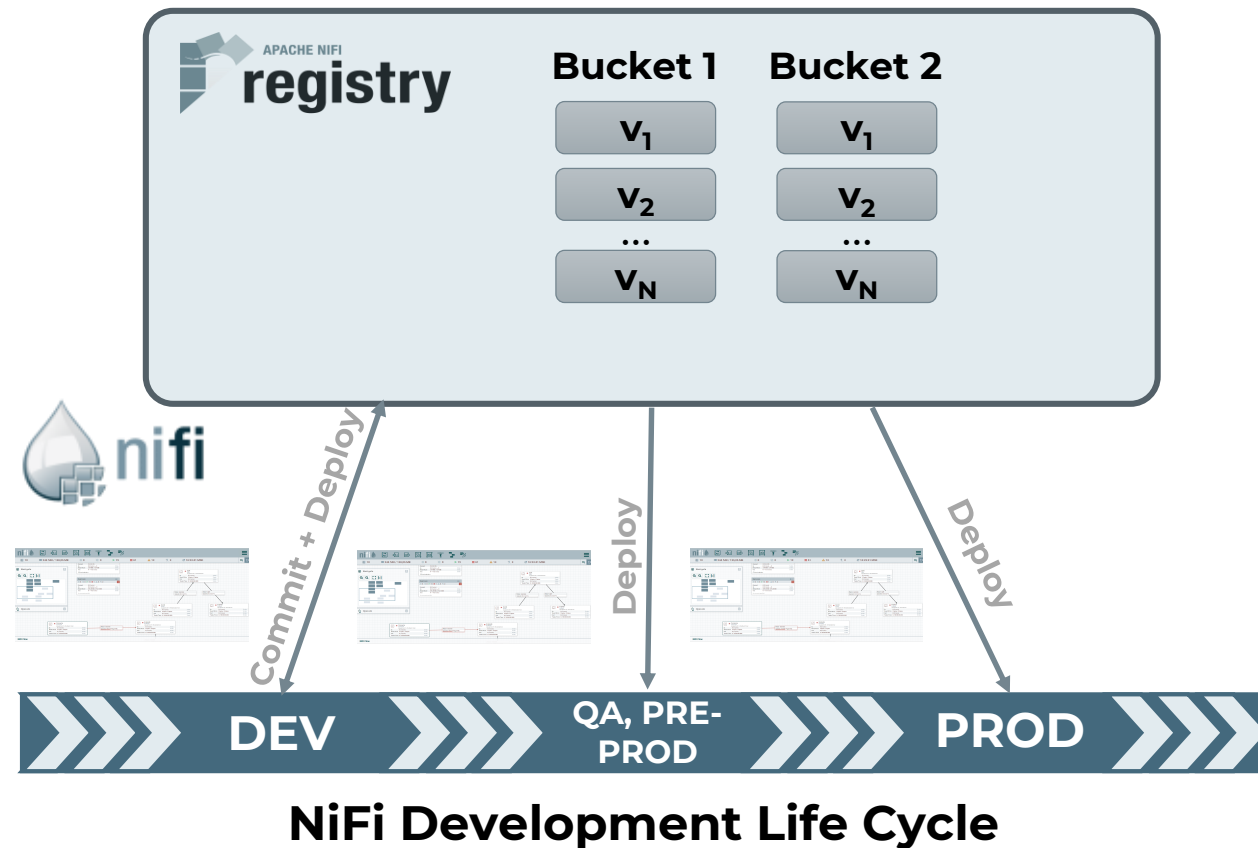


# NiFi Registry

Apache NiFi Registry – это система контроля версий и управления, используемая Arenadata Streaming для управления потоками данных (Flow)

Основные функции:

- Реализация Flow Registry для хранения и управления версиями.
- Интеграция с NiFi, позволяющая хранить, извлекать и обновлять версионные Flows из Flow Registry.
- Администрирование реестра для пользователей, групп и политик.



# Apache MiNiFi

Сервис MiNiFi состоит из компонент:

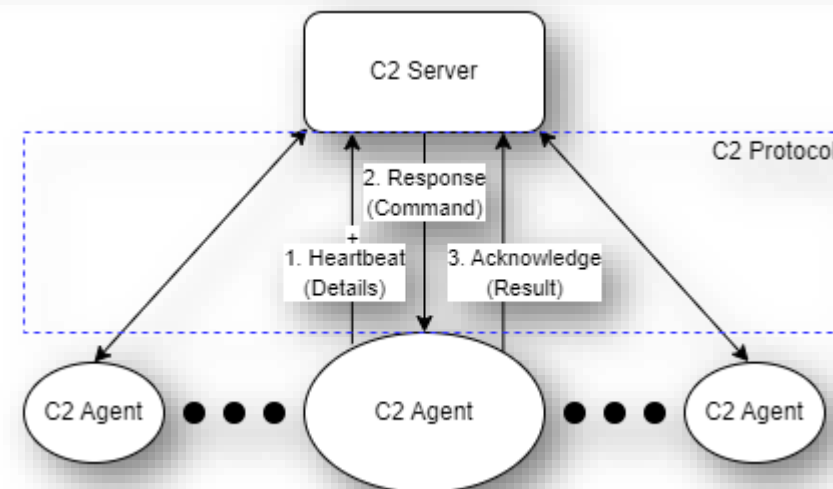
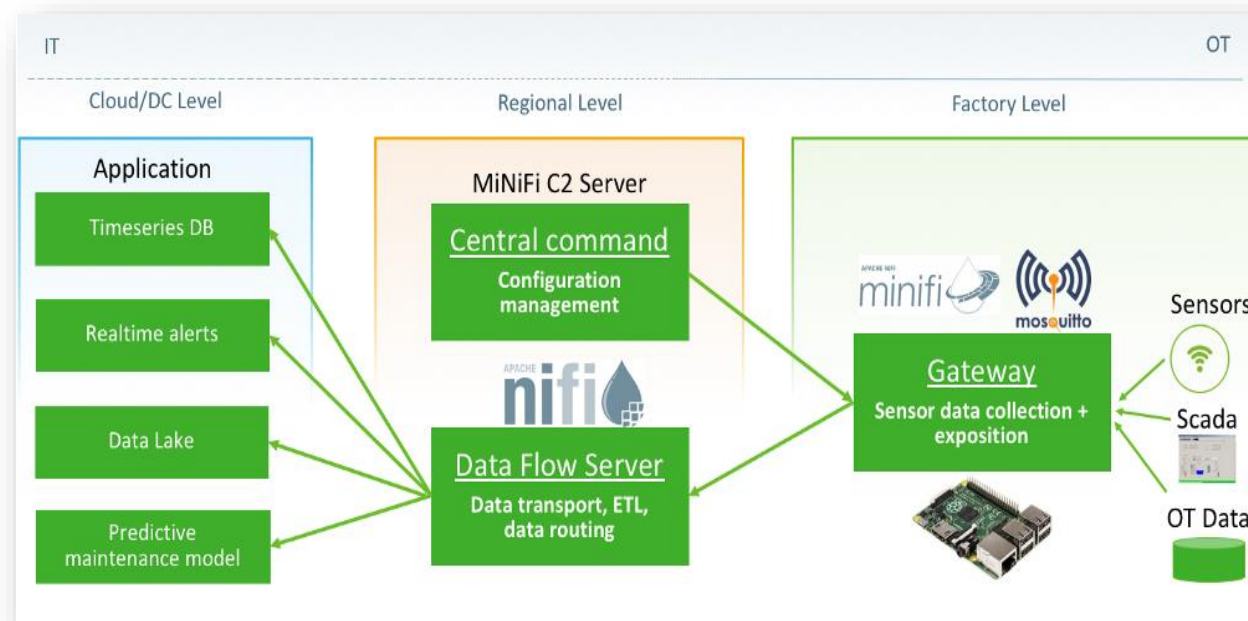
- MiNiFi Agent.
- MiNiFi C2 (Command and Control) Server.
- Агенты MiNiFi (может быть много) взаимодействуют с C2 Server.
- Протоколы взаимодействия Agents и C2 Server через HTTP/HTTPS RESTful

Agent → C2 Server

- HEARTBEAT – предоставляет информацию о состоянии Agent.
- ACKNOWLEDGE – используется агентами MiNiFi для подтверждения получения и выполнения запрошенной операции C2 Server.

C2 Server → Agent

- Flow (json-file) публикуется на стороне сервера.
- Агент получает операцию UPDATE.
- Агент загружает Flow из «location»
- Агент применяет новый Flow
- Агент подтверждает операцию UPDATE на сервере с успешным результатом.



# Компоненты Arenadata Streaming. Версии

Сервис	Компонент	Версия
Kafka	Kafka Broker	<a href="#">3.7.2_arenadata1</a>
Kafka Connect	Kafka Connect Worker	<a href="#">3.7.2_arenadata1</a> Debezium Connector <a href="#">2.7.0</a> Iceberg Sink Connector <a href="#">1.6.1_arenadata1</a>
Kafka REST Proxy	Kafka REST Proxy	<a href="#">7.7.2</a>
ksqlDB	ksqlDB Client	<a href="#">7.7.2</a>
	ksqlDB Server	<a href="#">7.7.2</a>
Schema-Registry	Schema-Registry	<a href="#">7.7.2</a>
NiFi	NiFi Registry	<a href="#">1.26.0_arenadata1</a>
	NiFi Server	<a href="#">1.26.0_arenadata1</a>
MiNiFi	MiNiFi Agent	<a href="#">1.26.0_arenadata1</a>
	MiNiFi C2 Server	<a href="#">1.26.0_arenadata1</a>
ZooKeeper	ZooKeeper Server	<a href="#">3.8.4_arenadata1</a>

<https://docs.arenadata.io/ru/ADStreaming/current/release-notes/support.html>

# Состав компонентов текущей и будущих версий

	ADS 3.6.2.1	ADS 3.6.2.2	ADS 3.7.2.1	ADS 3.9.0.1	ADS 3.9.1.1	
ADCM ⓘ	latest	latest	latest	latest	latest	End of Support (Dec 2023) ADS 2.8.x
Kafka ⓘ	3.6.2	3.6.2	3.7.2	3.9.0	3.9.1	Released
Kafka Connect ⓘ	3.6.2	3.6.2	3.7.2	3.9.0	3.9.1	ADS 3.6.2.1
CMAK ⓘ	3.1.2	3.1.2	—	—	—	ADS 3.6.2.2
Schema Registry ⓘ	7.6.2	7.6.2	7.7.2	7.9.0	7.9.1	ADS 3.7.2.1
KsqIDB ⓘ	7.6.2	7.6.2	7.7.2	7.9.0	7.9.1	In Development
Kafka REST ⓘ	7.6.2	7.6.2	7.7.2	7.9.0	7.9.1	ADS 3.9.0.1
Zookeeper ⓘ	3.8.4	3.8.4	3.8.4	3.8.4	3.8.4	Scheduled
NIFI ⓘ	1.26.0	1.26.0	1.26.0	1.26.0	1.27.0	ADS 3.9.1.1
NIFI Registry ⓘ	1.26.0	1.26.0	1.26.0	1.26.0	1.27.0	Deprecated
MINIFI ⓘ	1.26.0	1.26.0	1.26.0	1.26.0	1.27.0	—
C2 Server ⓘ	1.26.0	1.26.0	1.26.0	1.26.0	1.27.0	
Kerberos ⓘ	latest	latest	latest	latest	latest	

# ADS. Enterprise vs Community

Состав	ADS EE	ADS CE
Kafka, Kafka Manager, Kafka REST Proxy, KSQL, Schema Registry	+	+
NiFi Registry, NiFi Server, MiNiFi	+	+
Гибкая настройка из UI	+	+
Возможность делать expand, shrink из UI	+	+
SSL	+	-
Kerberos	+	-
Ranger plugins for authorization	+	-
LDAP / LDAPS authentication	+	-
ADS Control	+	-

# Arenadata Platform Security



# Arenadata Platform Security (ADPS)

ADPS – это сервис централизованного управления групповыми политиками безопасности кластера ADH, ADS (и др.) в составе единой корпоративной платформы сбора и хранения данных (EDP)

Основные возможности:

- централизованное управление безопасностью данных при работе с множеством сервисов и кластеров ADH, ADS;
- комплексный подход к организации безопасности с помощью защищенного периметра, аутентификации и авторизации пользователей, аудита пользовательских действий и защиты данных;
- единая система обеспечения безопасности одновременно для нескольких инсталляций, кластеров и гетерогенных инфраструктур;
- снижение эксплуатационных издержек на администрирование инфраструктуры;
- интеграция с решениями Arenadata и полностью автоматизированное развертывание (установка/удаление и настройка конфигураций) с помощью оркестратора Arenadata Cluster Manager.



# Arenadata Platform Security (ADPS)

**ADPS** поддерживает следующие модели авторизации:

- Управление доступом для данных, хранящихся в HDFS;
- Контроль доступа на уровне ресурсов для YARN;
- Контроль доступа на уровне семейства таблиц и столбцов для данных HBase;
- Контроль доступа к коллекциям Solr;
- Контроль доступа на уровне баз данных/таблиц/столбцов для данных Apache Hive;
- Контроль доступа Kafka Topics;
- Контроль доступа ресурсов NiFi и NiFi Registry;
- ...

**Особенности:**

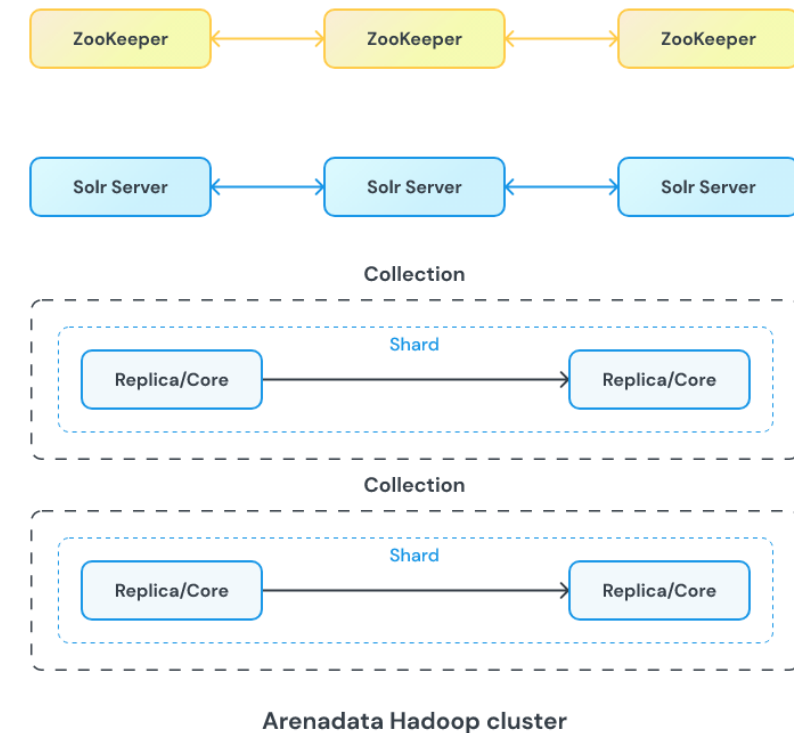
- Аутентификация MIT/LDAP/AD/FreeIPA/local UNIX system;
- Отчетность (аудит и мониторинг безопасности доступа/обработки данных);
- SSL для компонентов ADPS;
- Зоны безопасности в Apache Ranger;
- Поддержка доверенного прокси.

# Arenadata Platform Security (ADPS)

Безопасность периметра	Аутентификация	Авторизация (контроль доступа)	Аудит и мониторинг	Защита данных
<ul style="list-style-type: none"><li>• Apache Knox Gateway;</li><li>• Клиенты шлюза.</li></ul>	<ul style="list-style-type: none"><li>• Kerberos;</li><li>• LDAP/AD;</li><li>• SSO (Apache Knox Gateway).</li></ul>	<ul style="list-style-type: none"><li>• Контроль доступа Kafka Topics</li><li>• Контроль доступа ресурсов NiFi и NiFi Registry</li><li>• Контроль доступа Knox</li><li>• ...</li></ul>	<ul style="list-style-type: none"><li>• Запросы доступа;</li><li>• Операции обработки данных;</li><li>• Изменение данных.</li></ul>	<ul style="list-style-type: none"><li>• SSL для компонентов</li><li>• Шифрование</li></ul>

# ADPS. Компоненты

- Apache ZooKeeper
- Solr – платформа полнотекстового поиска с открытым исходным кодом, основанная на проекте Apache Lucene.
- Apache Ranger использует для хранения журналов аудита и поиска по ним в UI Ranger
- Основная функциональность:
  - Расширенные возможности полнотекстового поиска.
  - Форматы XML, JSON, CSV.
  - Высокая масштабируемость и отказоустойчивость.
  - Постраничный поиск и фильтрация.
  - Поддерживает основные языки.



# Kerberos

- Сетевой протокол аутентификации, позволяющий передавать данные через незащищённые сети для безопасной идентификации.

**Kerberos Key Distribution Center (Центр распределения ключей, KDC)** состоит из трех частей:

- База данных пользователей и сервисов (известных как принципалы), о которых он знает, и соответствующие пароли Kerberos;
- Сервер аутентификации (AS), который выполняет первоначальную проверку подлинности и выдает Ticket Granting Ticket (TGT);
- Ticket Granting Server (TGS) – сервер, который оформляет последующие билеты на основе начального TGT.

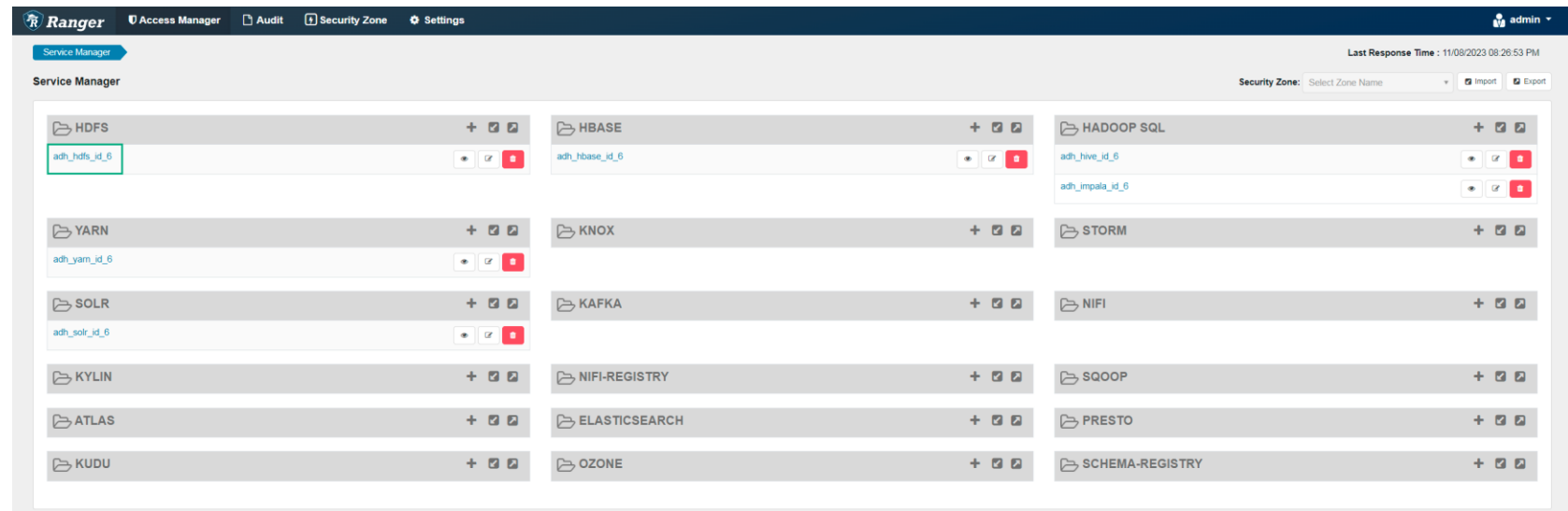
# ADPS. Компоненты



- Apache Ranger обеспечивает комплексный подход к безопасности кластера Hadoop, основанный на централизованном администрировании политик безопасности (авторизацию, аудит, защита данных).
- Приложения интегрируются с Ranger через стандартную модель плагинов (HDFS, YARN, HBase, Hive, Knox, Solr, Kafka, ...).

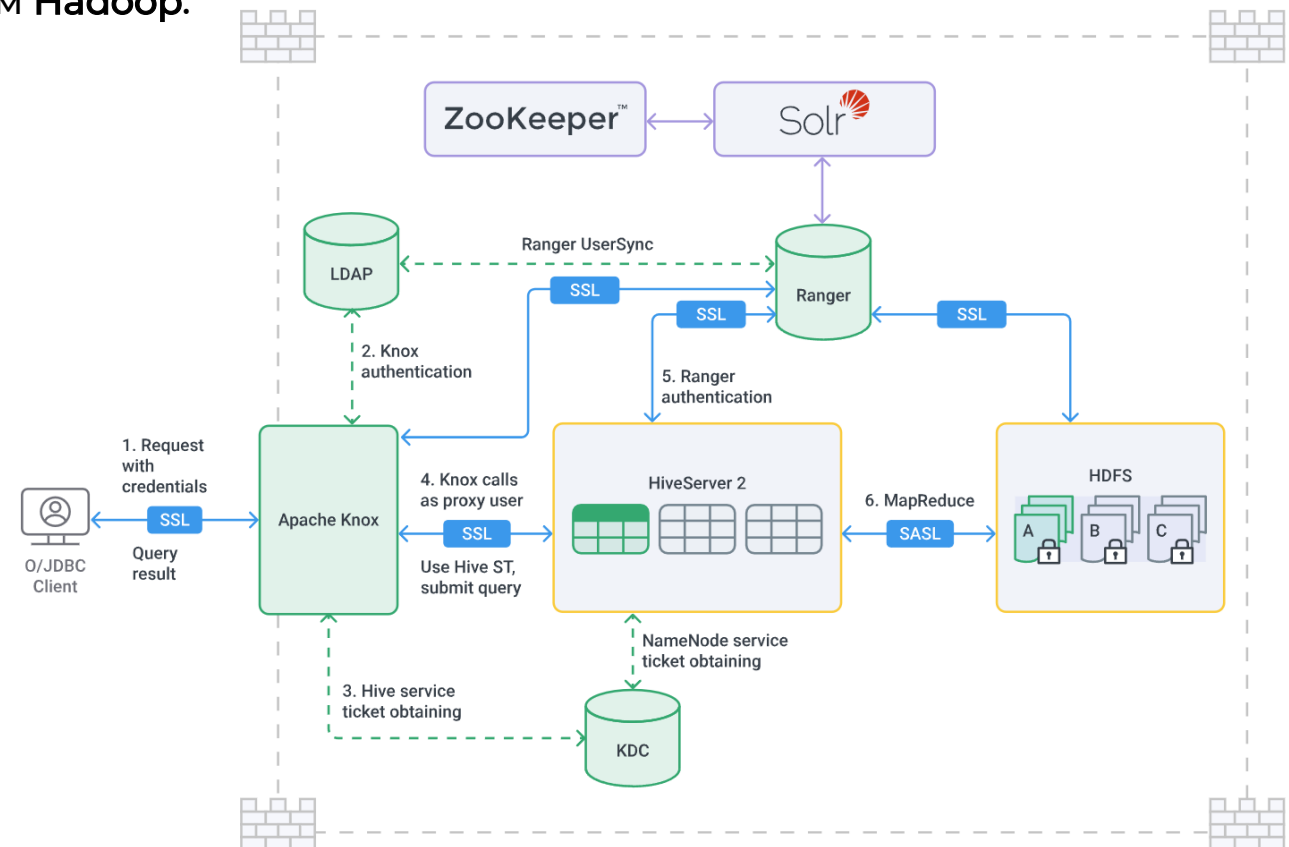
Ranger включает в себя следующие компоненты:

- Ranger Admin
- Ranger UserSync
- Ranger Key Management Service (Key management, Access control policies, Audit)



# Apache Knox

- **Knox Gateway (Knox)** – отвечает за безопасность периметра, что позволяет предприятиям расширять доступ к **Hadoop** новым пользователям, а также поддерживать соответствие политик безопасности предприятия.
- **Knox** интегрируется с системами управления идентификацией и SSO (Single Sign-On), позволяет использовать идентификацию систем для доступа к кластерам **Hadoop**.



# Компоненты Arenadata Platform Security. Версии

Сервис	Компонент	Версия
Apache Knox	Knox Gateway	<a href="#">2.0.0</a> _arenadata1
Apache Ranger	Ranger Admin	<a href="#">2.5.0</a> _arenadata2
	Ranger KMS	<a href="#">2.5.0</a> _arenadata2
	Ranger UserSync	<a href="#">2.5.0</a> _arenadata2
Apache Solr	Server	<a href="#">8.11.3</a> _arenadata1
Apache ZooKeeper	Server	<a href="#">3.8.4</a> _arenadata1
Arenadata Monitoring	Monitoring	3.2
MariaDB	MariaDB	<a href="#">5.5.68</a>

<https://docs.arenadata.io/ru/ADPS/current/release-notes/supported.html>



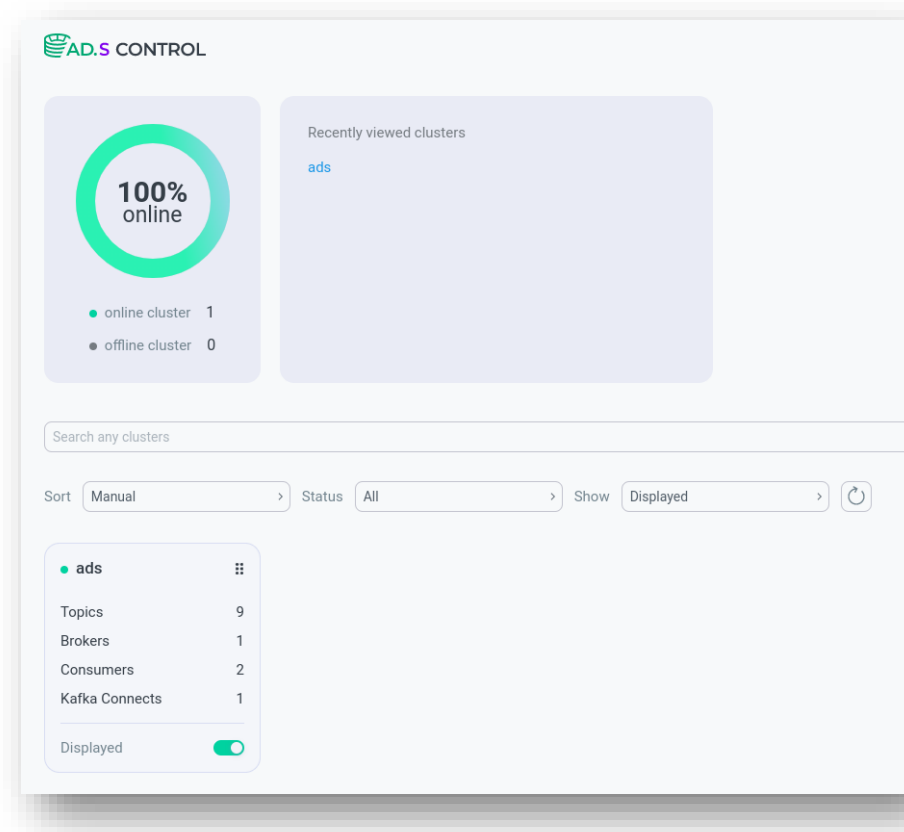
# Arenadata Streaming Control

# Arenadata Streaming Control

- **Arenadata Streaming Control** – это графический веб-интерфейс (GUI) для управления и мониторинга кластеров потоковой передачи Arenadata, предоставляющий управление экземплярами Kafka Connect.
- Предоставляет возможности мониторинга, которые позволяют администраторам просматривать состояние своих кластеров.

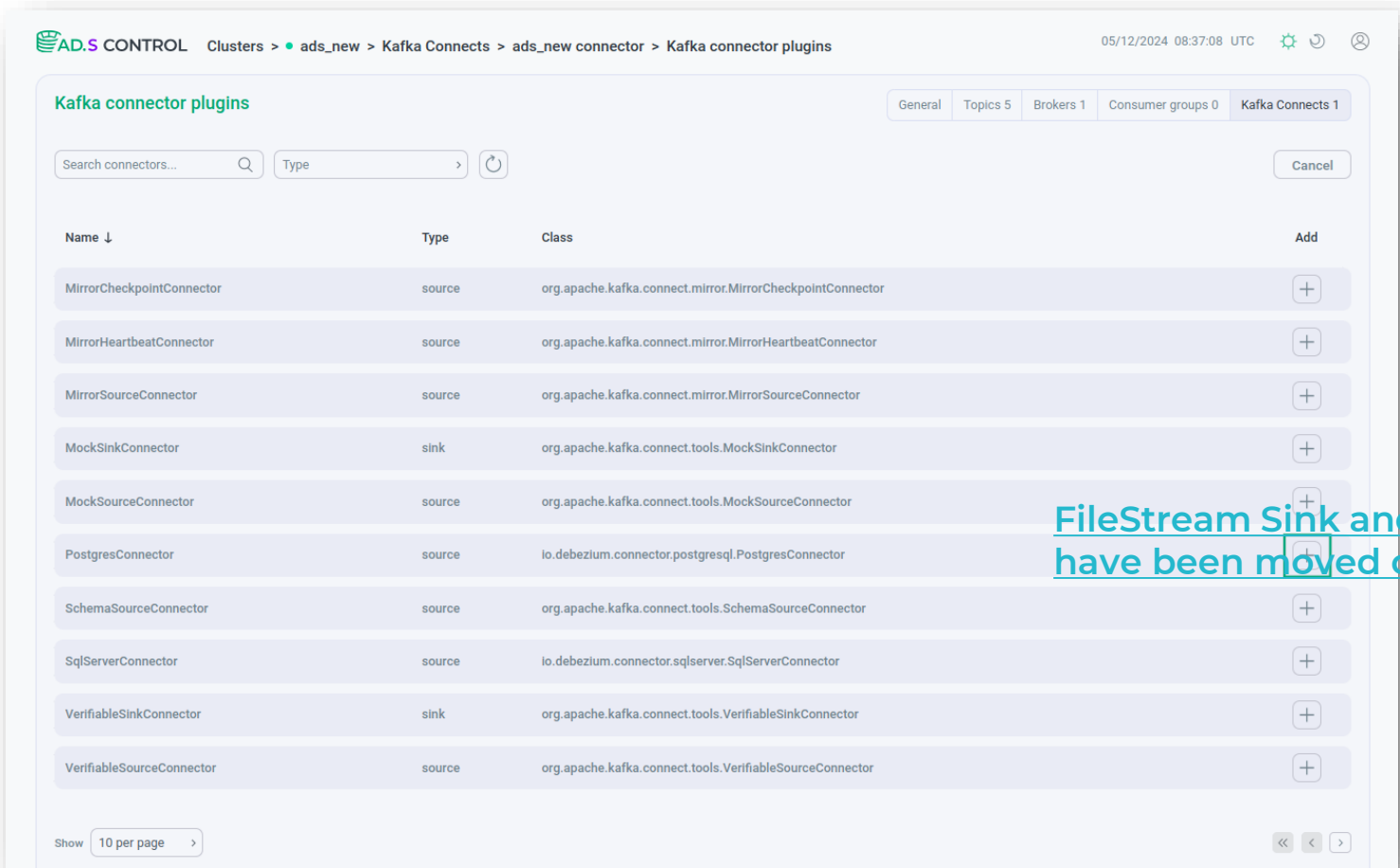
Доступны к использованию следующие коннекторы:

- FileStreamSinkConnector — коннектор, предназначенный для считывания данных из Kafka и вывода их в локальный файл.
- FileStreamSourceConnector — коннектор, предназначенный для считывания данных из файла и отправки их в Kafka.
- Подробнее о работе с коннекторами FileStream описано в статье [FileStream в ADS Control](#).
- MirrorCheckpointConnector, MirrorHeartbeatConnector, MirrorSourceConnector — коннекторы, предназначенные для репликации топиков с использованием механизма [Mirror Maker 2](#). Подробнее о работе с этими коннекторами описано в статье [Mirror Maker 2 в ADS Control](#).
- ...



# Доступные коннекторы

## SOURCE/SINK CONNECTORS



Дополнительные коннекторы устанавливаются при помощи копирования jar-архива с коннектором на kafka connect cluster и настройки ADCM.

# ADS Control



**Arenadata Streaming Control (ADS Control) — это графический пользовательский веб-интерфейс (GUI) для управления и мониторинга кластеров потоковой передачи Arenadata**

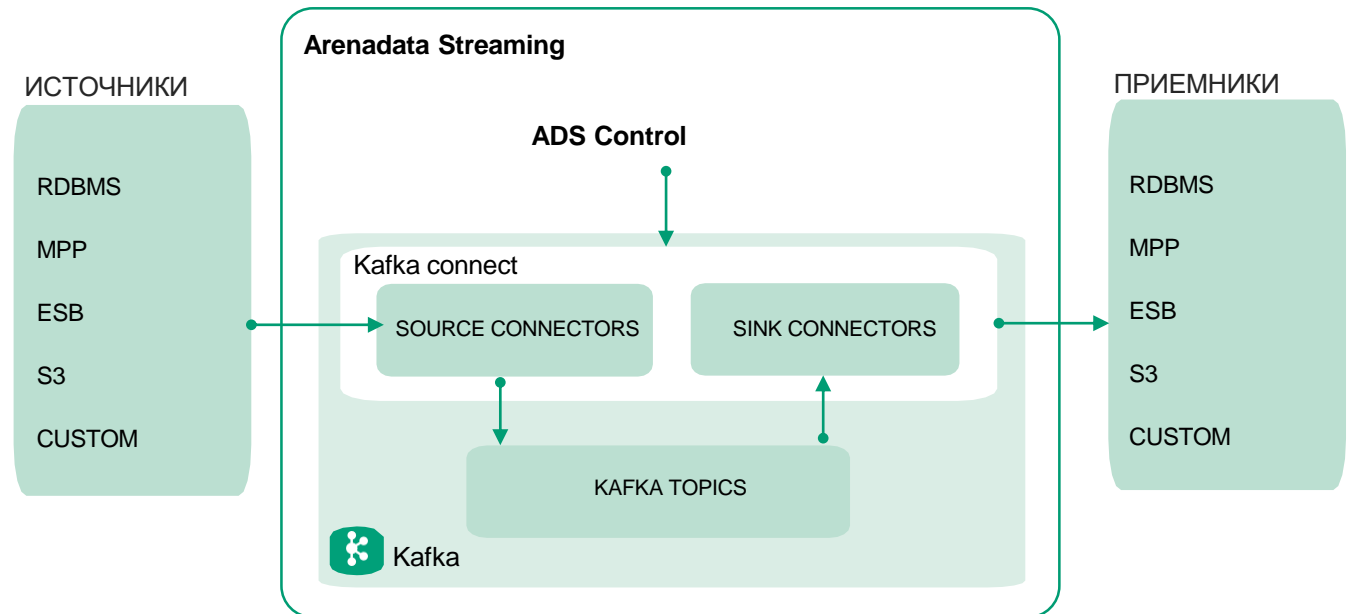
**Основные возможности:**

- Получить быстрый обзор состояния кластера
- Наблюдать и контролировать коннекторы к источникам и приемникам
- Настроить перенос данных в Kafka и из Kafka

# Управление коннекторами

Command Center позволяет настраивать передачу данных в реальном времени между источником и приемником с помощью source и sink коннекторов.

- Простой доступ к Kafka Connect через UI
- Создание и управление коннекторами
- Мониторинг состояния
- Работа с множеством кластеров Kafka
- Широкий набор коннекторов
- Добавление собственных коннекторов



# Mirror Maker 2

- MirrorMaker 2.0—основанный на платформе сервиса Kafka Connect механизм репликации данных из исходного кластера на удаленный
- При архитектуре репликации Active/Active кластеры получают данные непосредственно от источников данных и реплицированные данные, вводимые из удаленного кластера.
- При архитектуре репликации Active/Standby целевой кластер находится в пассивном режиме (не имеет подключенных к нему потребителей и производителей) и получает только реплицированные данные.

MirrorMaker 2.0 в составе Kafka Connect создает специальные коннекторы для включения сложных потоков между кластерами ADS:

- MirrorSourceConnector осуществляет репликацию топиков из исходного кластера в целевой кластер.
- MirrorCheckpointConnector создает контрольные точки смещения потребителя и синхронизирует смещение со служебным топиком `__consumer_offsets` исходного кластера.
- MirrorHeartbeatConnector периодически проверяет подключение между кластерами, создавая сообщения в специальном топике heartbeats в исходном кластере через заданный период времени и считывая их в целевом кластере.

