



Система для поточковой передачи и обработки данных

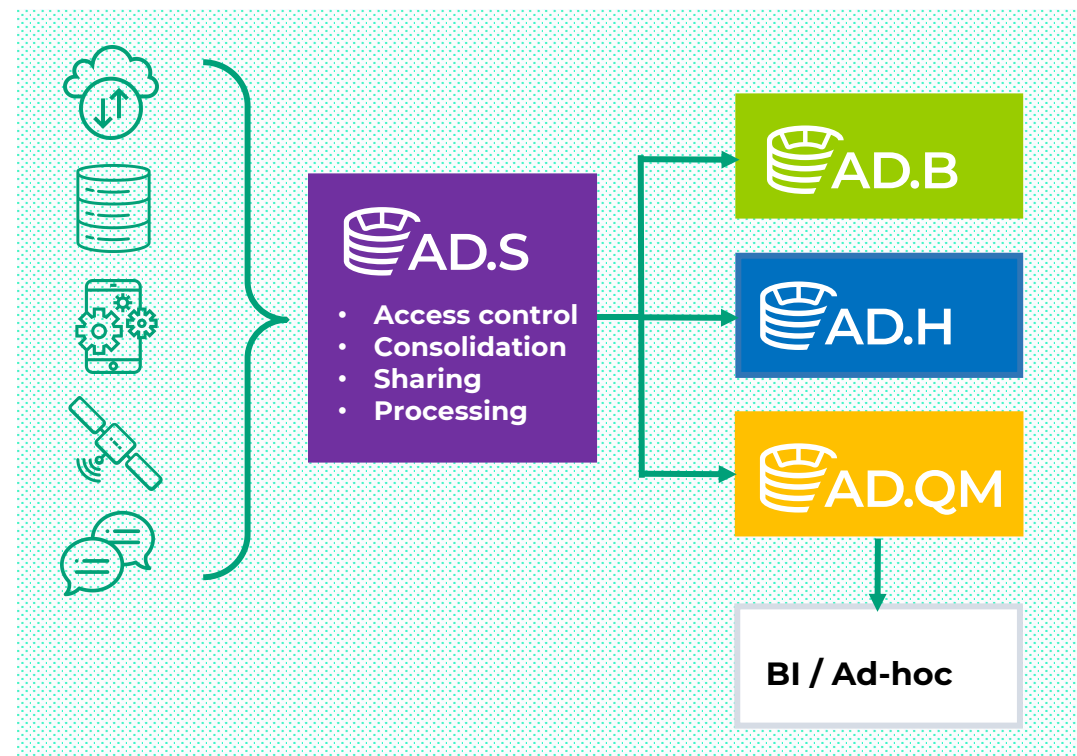
Arenadata Streaming

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.

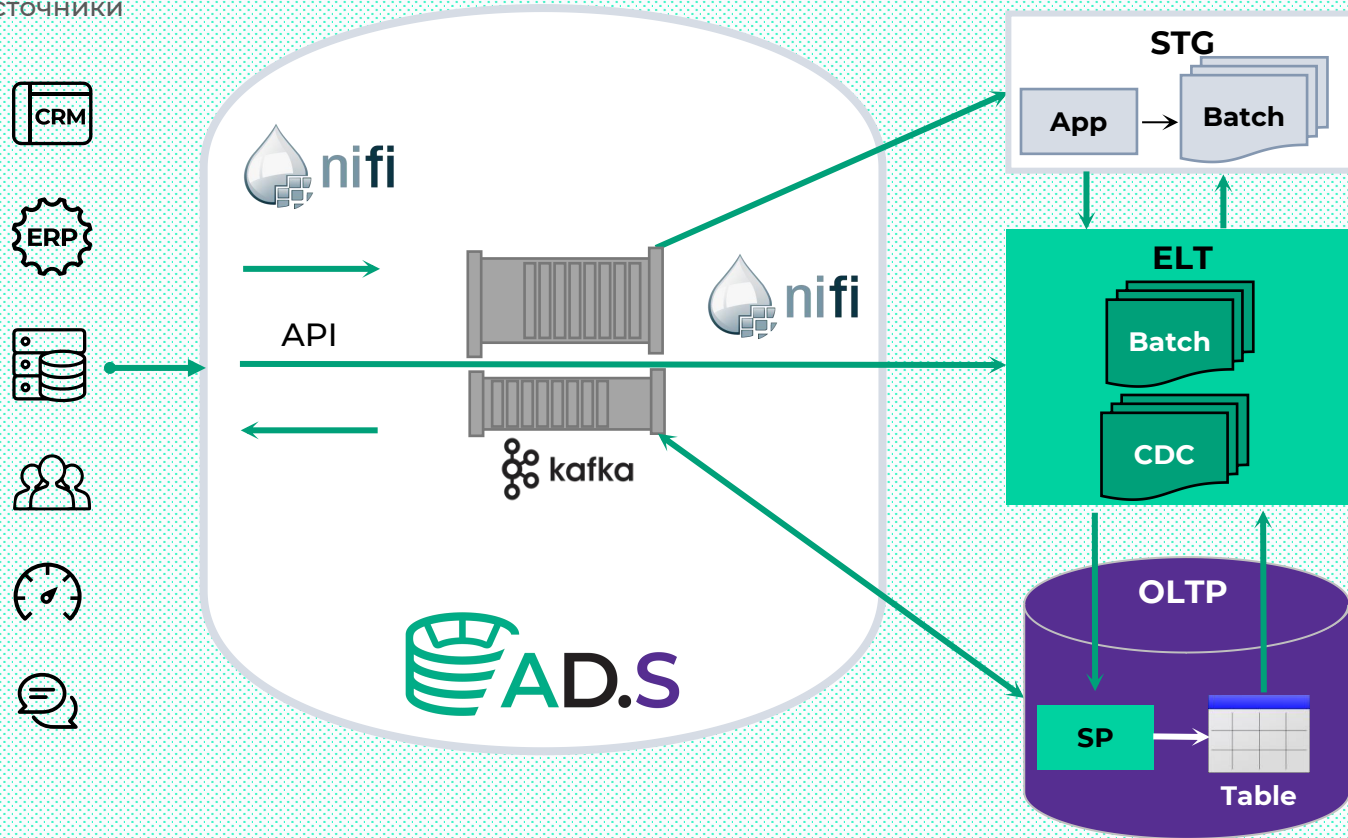
Используется как:

- 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 позволяет

Источники



- 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.

Объединяя лучшее

Message brokers:

- Apache Kafka
- Confluent Platform
- RabbitMQ
- Apache ActiveMQ
- IBM WebSphere MQ



Dataflow platforms:

- Hortonworks DataFlow
- StreamSets
- Informatica BDM

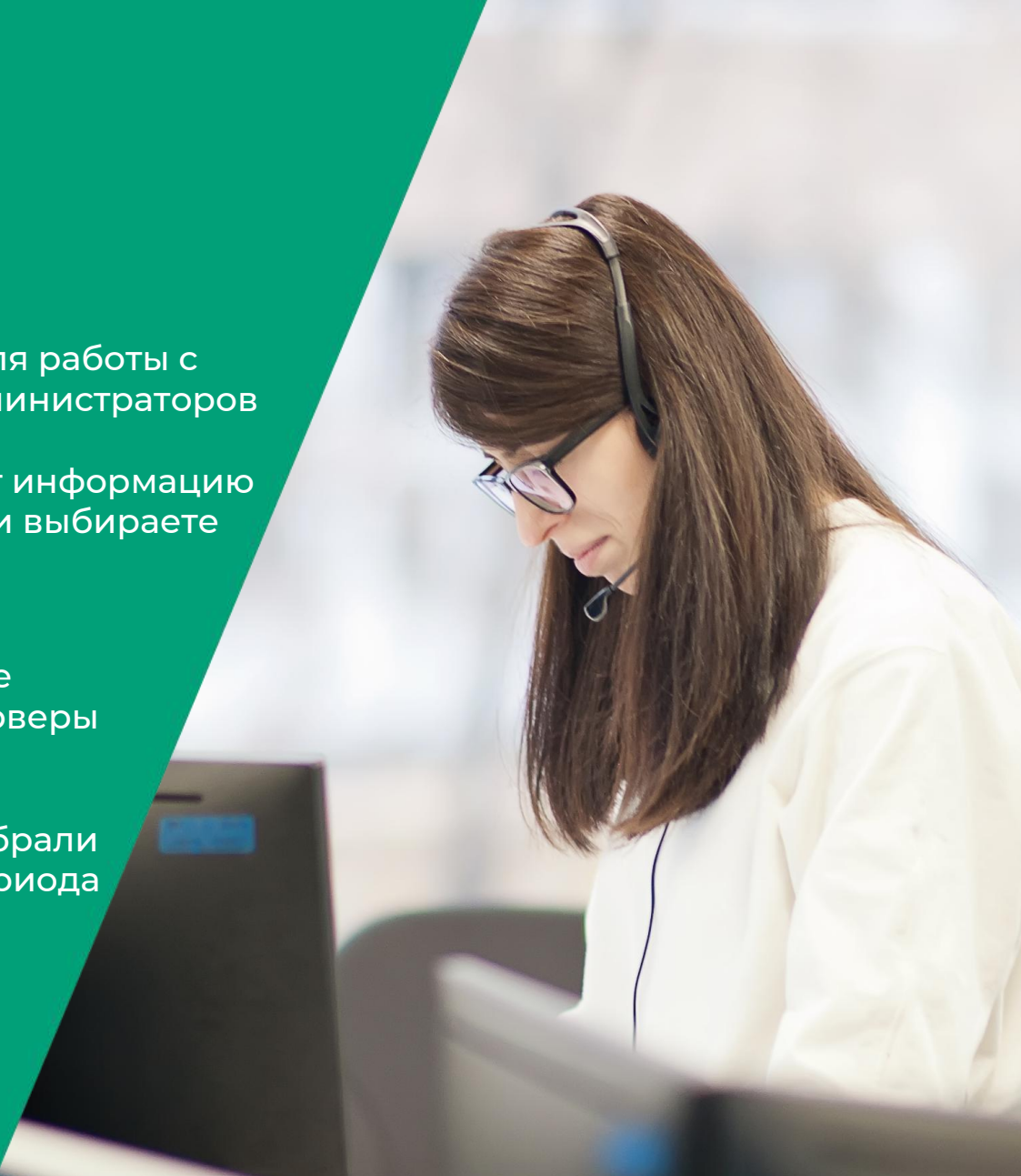


Мы объединили преимущества двух типов систем и разработали систему Arenadata Streaming (ADS) на базе Apache Kafka и Apache Nifi, которая

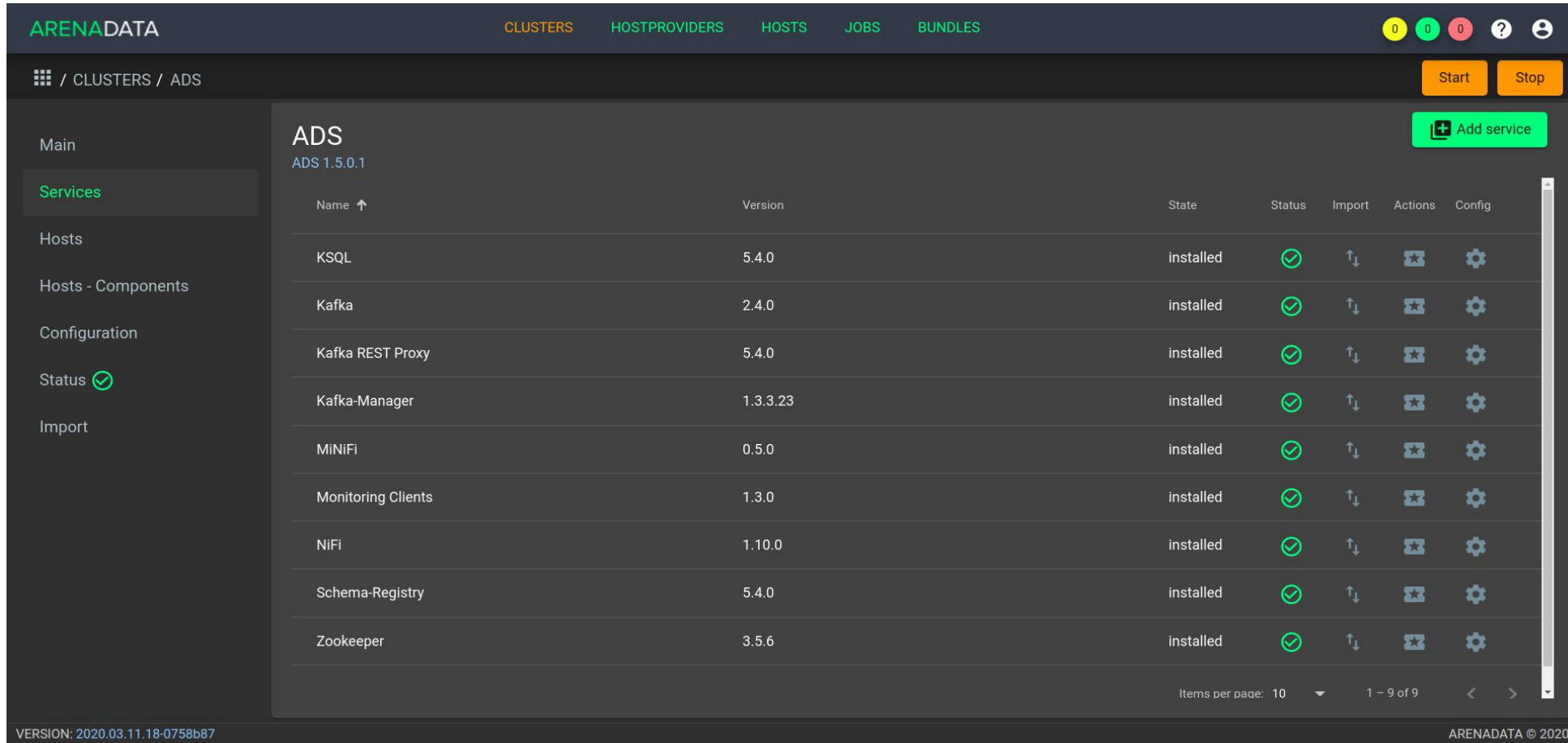
- поддерживает потоковую обработку данных в реальном времени в большом объёме,
- приём данных из разных источников
- и единую точку доступа для всех ваших данных.

Преимущества ADS

- ADS позволяет снизить порог входа в технологию, так как для работы с продуктом не требуется высокий уровень компетенции администраторов
- В отличие от других брокеров сообщений, которые удаляют информацию сразу же после того, как она была доставлена, с ADS вы сами выбираете нужный период хранения данных
- Гибкость и масштабируемость системы позволяет ей расти вместе с вашими потребностями, подстраиваясь под текущие задачи. Так, вы всегда можете добавить в кластер новые серверы без остановки работы системы или же убрать ненужные
- ADS получает и хранит информацию независимо от того, забрали её или нет: она доступна в течение всего обозначенного периода хранения
- Широкий набор API для интеграции с другими внешними системами.



Cluster deployment



The screenshot displays the ARENADATA web interface for managing a cluster. The top navigation bar includes links for CLUSTERS, HOSTPROVIDERS, HOSTS, JOBS, and BUNDLES. The main content area is titled 'ADS 1.5.0.1' and shows a list of services installed on the cluster. The services are listed in a table with columns for Name, Version, State, Status, Import, Actions, and Config. All services are shown as 'installed' with a green checkmark status. The left sidebar contains navigation options: Main, Services (selected), Hosts, Hosts - Components, Configuration, Status (with a green checkmark), and Import. The bottom of the interface shows the version '2020.03.11.18-0758b87' and the copyright 'ARENADATA © 2020'.

Name ↑	Version	State	Status	Import	Actions	Config
KSQL	5.4.0	installed	✓	↑↓	★	⚙️
Kafka	2.4.0	installed	✓	↑↓	★	⚙️
Kafka REST Proxy	5.4.0	installed	✓	↑↓	★	⚙️
Kafka-Manager	1.3.3.23	installed	✓	↑↓	★	⚙️
MINIFI	0.5.0	installed	✓	↑↓	★	⚙️
Monitoring Clients	1.3.0	installed	✓	↑↓	★	⚙️
NIFI	1.10.0	installed	✓	↑↓	★	⚙️
Schema-Registry	5.4.0	installed	✓	↑↓	★	⚙️
Zookeeper	3.5.6	installed	✓	↑↓	★	⚙️

Items per page: 10 1 - 9 of 9

Cluster management

The screenshot displays the ARENADATA web interface for managing a Kafka cluster. The top navigation bar includes tabs for CLUSTERS, HOSTPROVIDERS, HOSTS, JOBS, and BUNDLES. The main content area is titled 'Kafka' and shows the 'server.properties' configuration. The configuration is presented as a table with various properties and their values. A sidebar on the left contains links for Main, Configuration, Status, and Import. The bottom of the interface shows the version number and copyright information.

ARENADATA

CLUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES

/ CLUSTERS / HOLY DON / SERVICES / KAFKA

Check Expand Restart Shrink Start Stop

Main

Configuration

Status ☒

Import

Kafka

Description configuration

ansible update

Search field

Advanced

Save

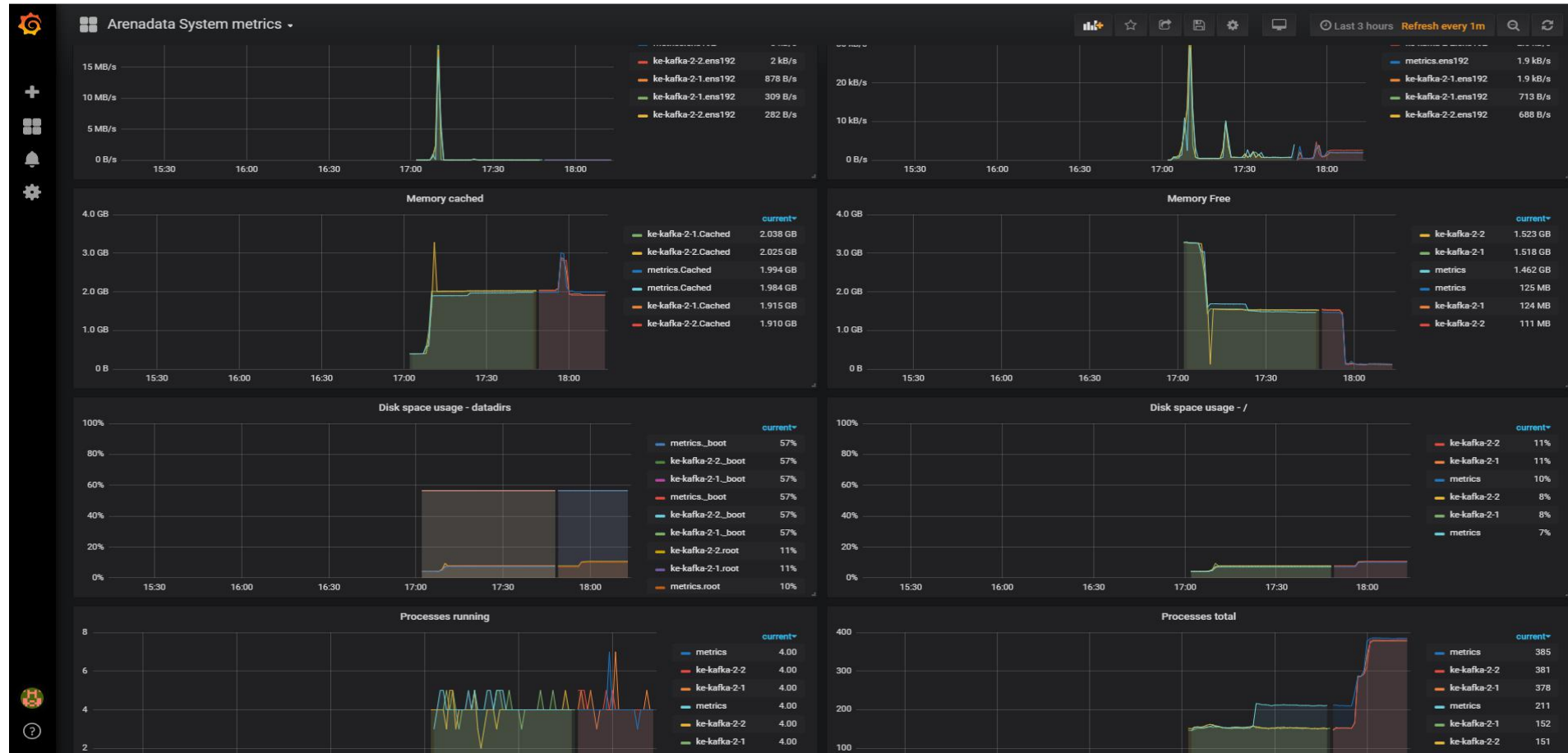
server.properties

auto.leader.rebalance.enable:	☒
queued.max.requests:	500
num.network.threads:	3
num.io.threads:	8
unclean.leader.election.enable:	☐
zookeeper.connection.timeout.ms:	30000
zookeeper.session.timeout.ms:	30000

VERSION: 2019.10.15.18-0305a0e1

ARENADATA © 2019

Cluster monitoring



Broker Metrics

AD Kafka Manager Default Cluster▼ Brokers Topic▼ Preferred Replica Election Reassign Partitions Consumers					
Clusters / Default / Brokers					
← Brokers					
Id	Host	Port	JMX Port	Bytes In	Bytes Out
0	ambari-1.c.arenadata-test.internal	PLAINTEXT:9092	9999	41.00	80.20
1	ambari-2.c.arenadata-test.internal	PLAINTEXT:9092	9999	0.8k	1.9k
2	ambari.c.arenadata-test.internal	PLAINTEXT:9092	9999	41.00	80.20
Combined Metrics					
Rate	Mean	1 min	5 min	15 min	
Messages in /sec	4.71	10.80	10.80	10.77	
Bytes in /sec	0.4k	0.9k	0.9k	0.9k	
Bytes out /sec	0.9k	2.1k	2.1k	2.1k	
Bytes rejected /sec	0.00	0.00	0.00	0.00	
Failed fetch request /sec	0.00	0.00	0.00	0.00	
Failed produce request /sec	0.00	0.00	0.00	0.00	

Topics management

AD Kafka Manager **Default**

Cluster ▾ Brokers Topic ▾ Preferred Replica Election Reassign Partitions Consumers

Clusters / Default / Topics / __consumer_offsets

← __consumer_offsets

Topic Summary

Replication	1
Number of Partitions	50
Sum of partition offsets	78,917
Total number of Brokers	3
Number of Brokers for Topic	3
Preferred Replicas %	100
Brokers Skewed %	0
Brokers Leader Skewed %	0
Brokers Spread %	100
Under-replicated %	0

Config

Value

segment.bytes	104857600
compression.type	producer
cleanup.policy	compact

Operations

Delete Topic

Reassign Partitions

Generate Partition Assignments

Add Partitions

Update Config

Manual Partition Assignments

Partitions by Broker

Broker	# of Partitions	# as Leader	Partitions	Skewed?	Leader Skewed?
0	17	17	(1,4,7,10,13,16,19,22,25,28,31,34,37,40,43,46,49)	false	false
1	16	16	(2,5,8,11,14,17,20,23,26,29,32,35,38,41,44,47)	false	false
2	17	17	(0,3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48)	false	false

Consumers consuming from this topic

console-consumer-58982	KF
------------------------	----

Partitions management

AD Kafka Manager **Default**

Cluster ▾ Brokers Topic ▾ Preferred Replica Election Reassign Partitions Consumers

Clusters / Default / Topics

[←](#) Manual Partition Assignments [Save Partition Assignment](#)

Type to filter topics

__consumer_offsets

Partition 0
Replica 0: Broker 0 ▾

Partition 10
Replica 0: Broker 2 ▾

Partition 24
Replica 0: Broker 0 ▾

Partition 25
Replica 0: Broker 0 ▾

Partition 20
Replica 0: Broker 0 ▾

Partition 29
Replica 0: Broker 2 ▾

Partition 5
Replica 0: Broker 0 ▾

Partition 42
Replica 0: Broker 2 ▾

Partition 37
Replica 0: Broker 0 ▾

Partition 14
Replica 0: Broker 2 ▾

Partition 46
Replica 0: Broker 0 ▾

Partition 1
Replica 0: Broker 2 ▾

Type to filter metrics

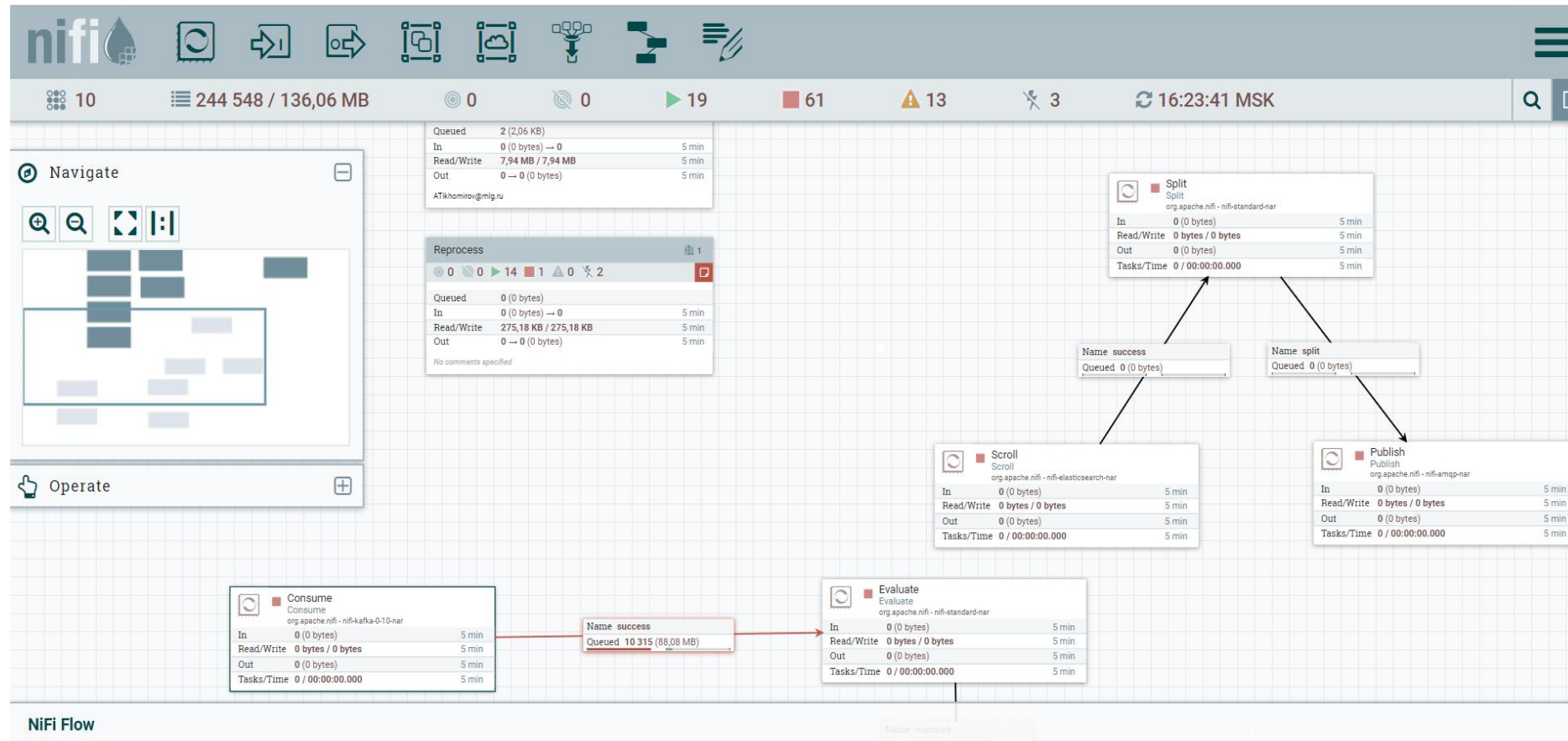
Broker 2

Rate	Mean	15 min
Messages in /sec	0.00	0.00
Bytes in /sec	0.00	0.00
Bytes out /sec	0.00	0.00
Bytes rejected /sec	0.00	0.00
Failed fetch request /sec	0.00	0.00
Failed produce request /sec	0.00	0.00
Process CPU Load	0.00	
System CPU Load	0.00	

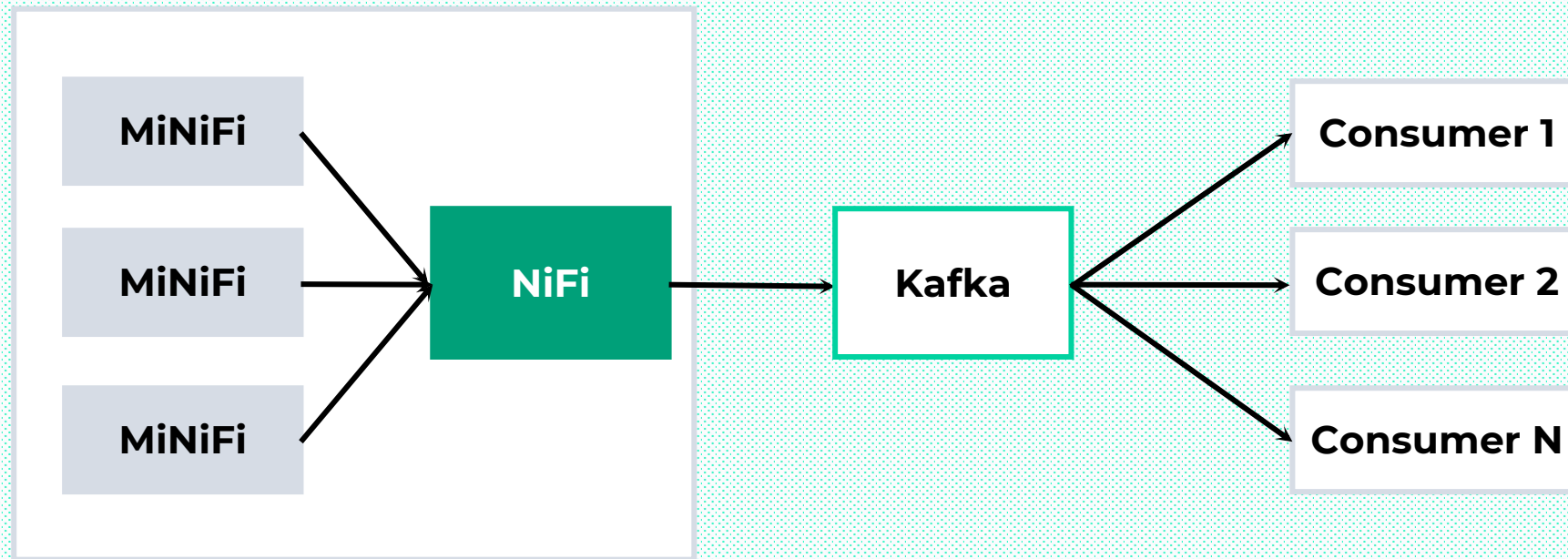
Broker 1

 ARENADATA

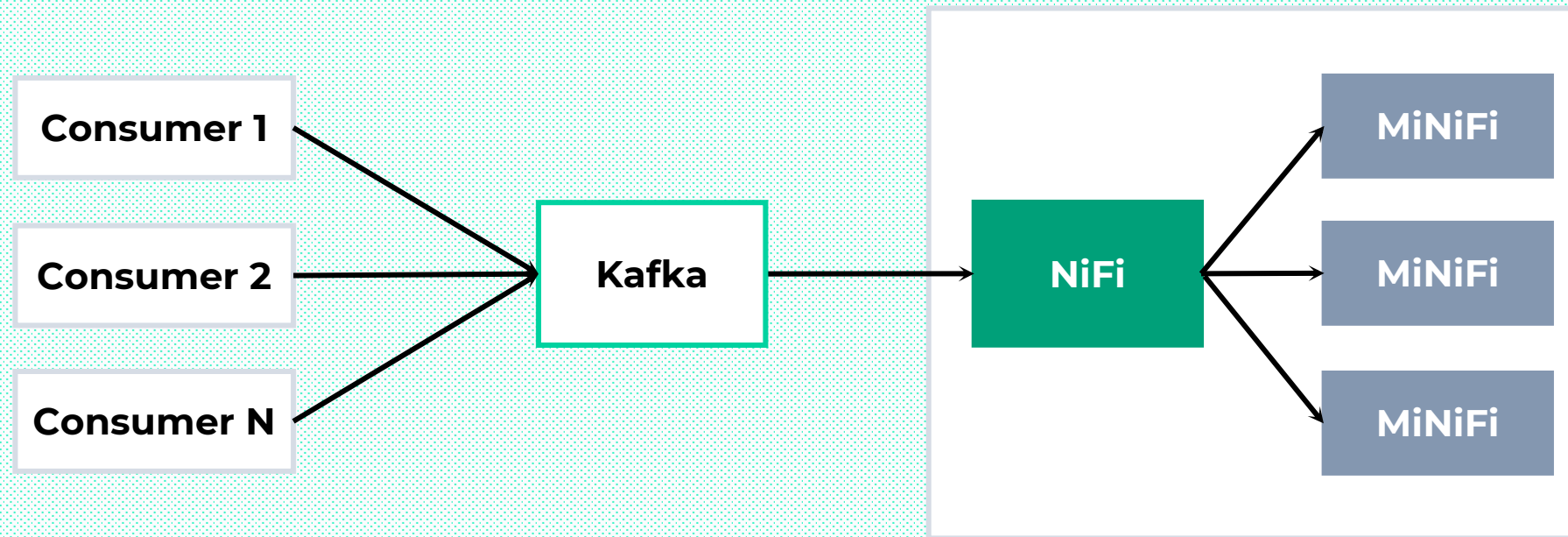
Nifi as ETL



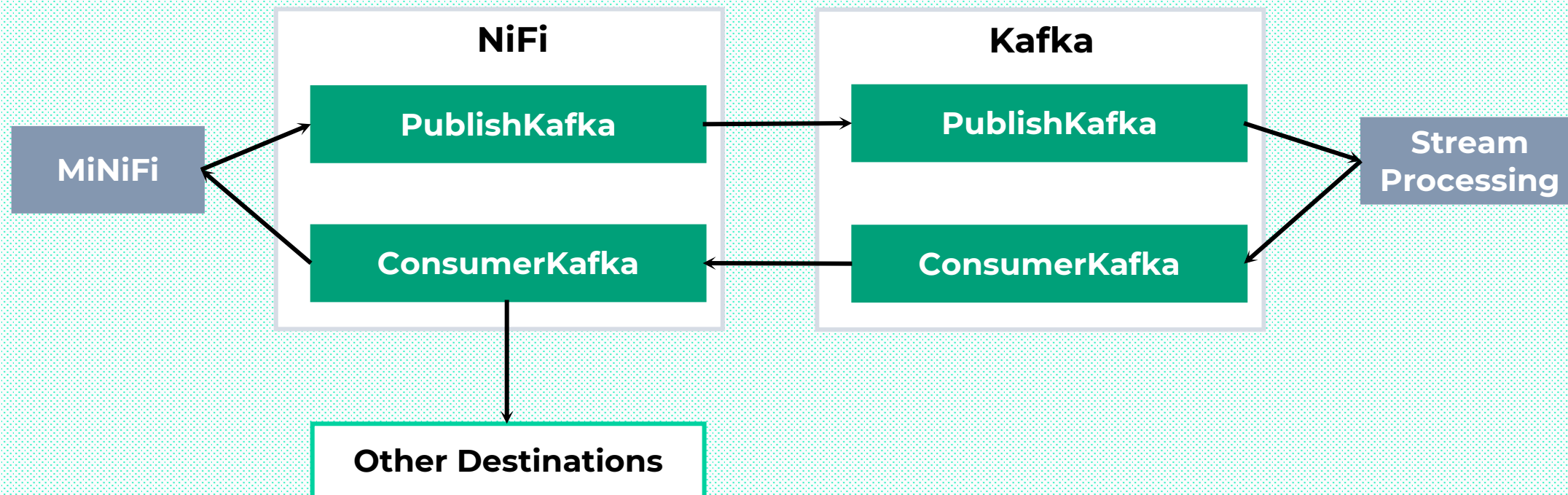
NiFi as producer



NiFi as consumer



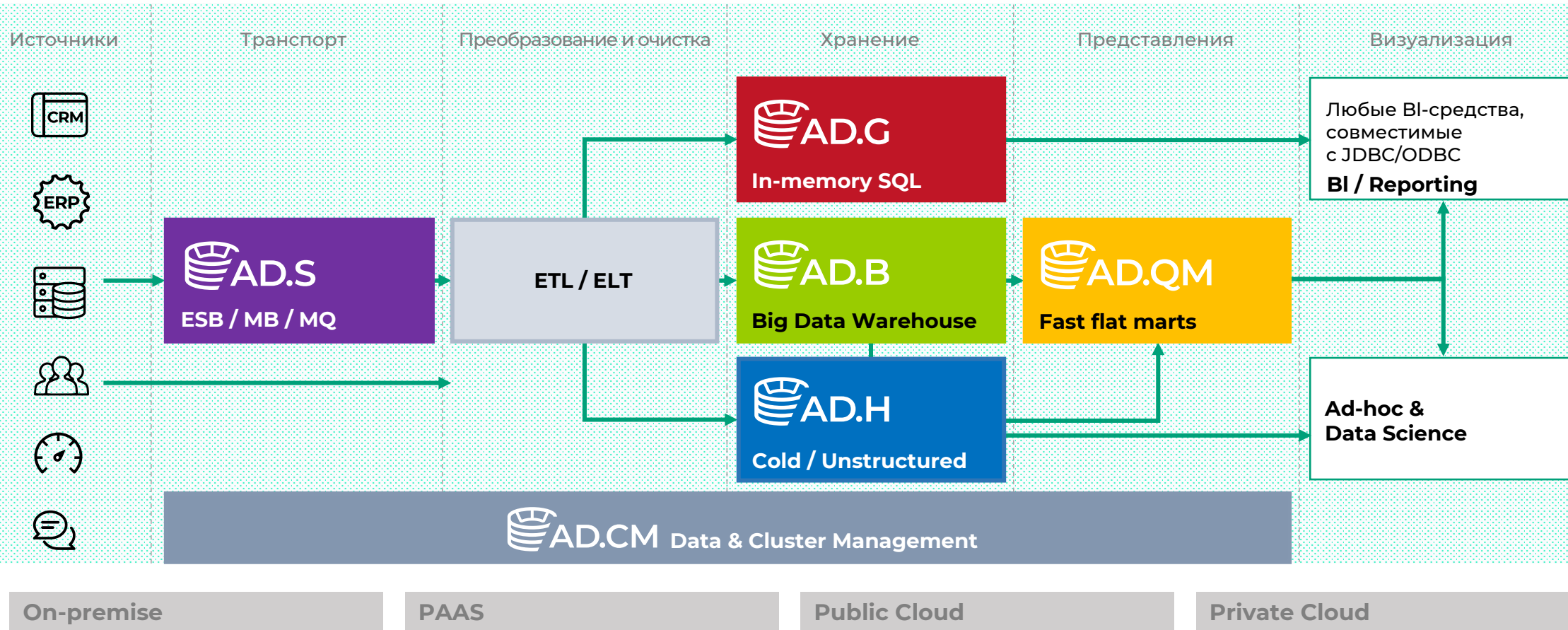
Двунаправленный поток данных



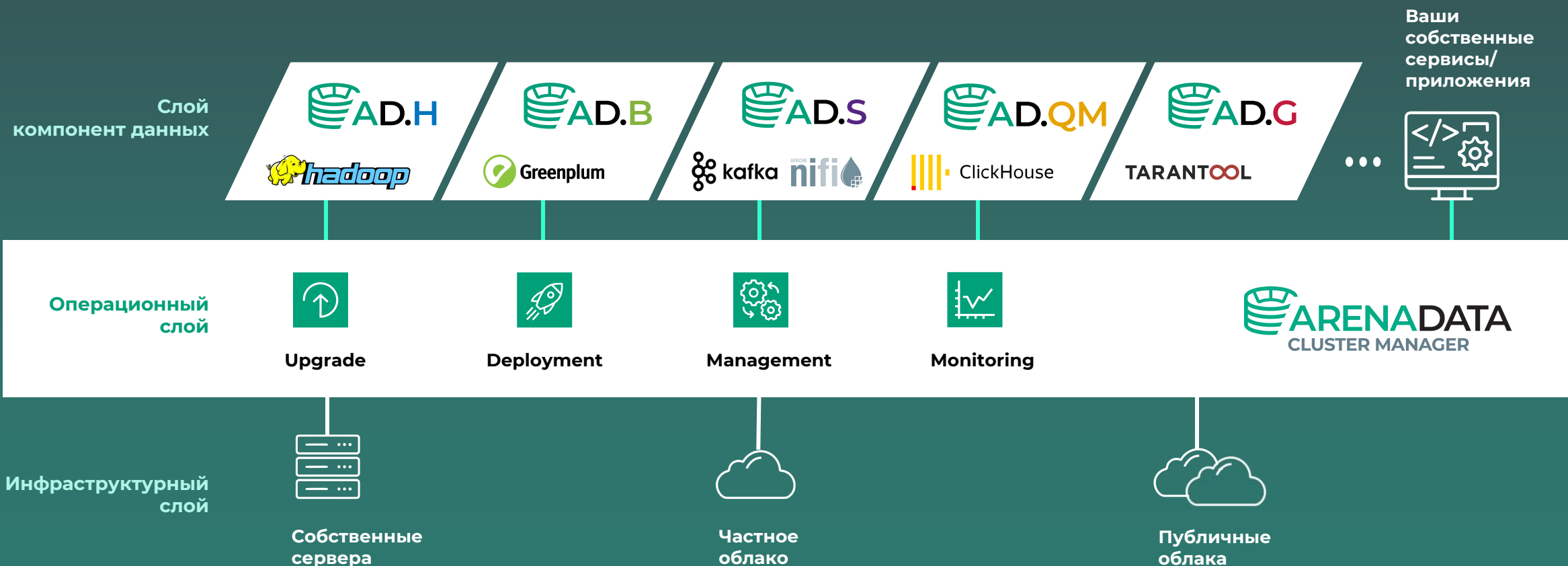
Arenadata Streaming Roadmap



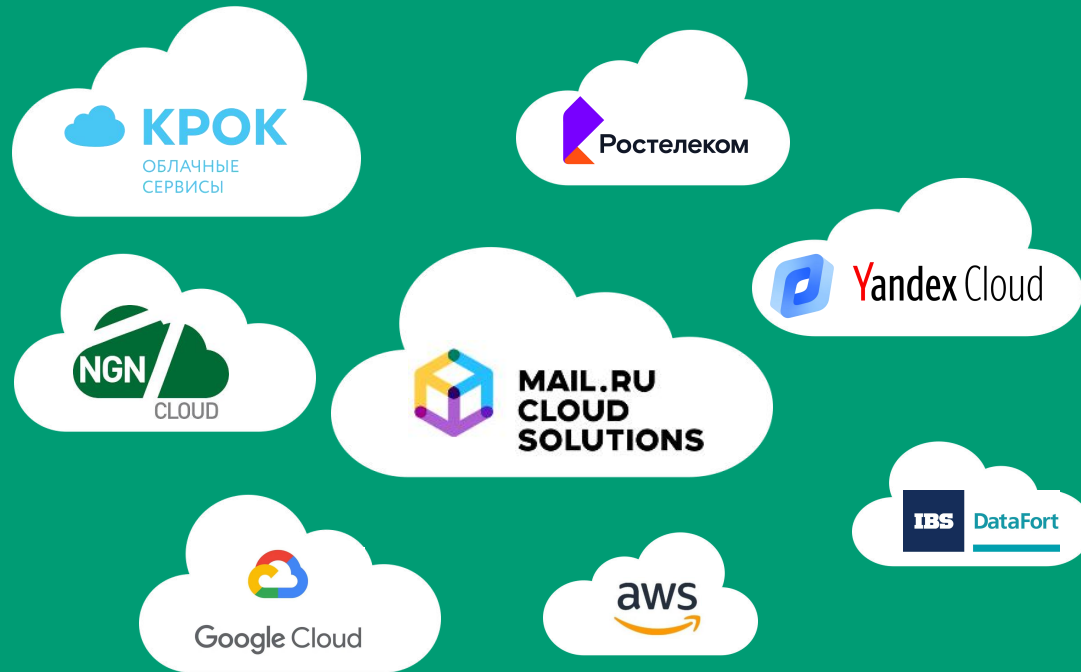
ADS — часть корпоративной платформы данных



Гибридное корпоративное облако



Enterprise Multi-cloud платформа данных



- Вы можете развернуть любой компонент платформы Arendata EDP, в том числе ADS, как на bare-metal оборудовании, так и в облаке, или же воспользоваться востребованным сегодня сценарием — гетерогенной ИТ-инфраструктурой или Multi-clouds.
- Установите ADS распределённо на разные инфраструктуры, обеспечив между ними сетевую доступность.

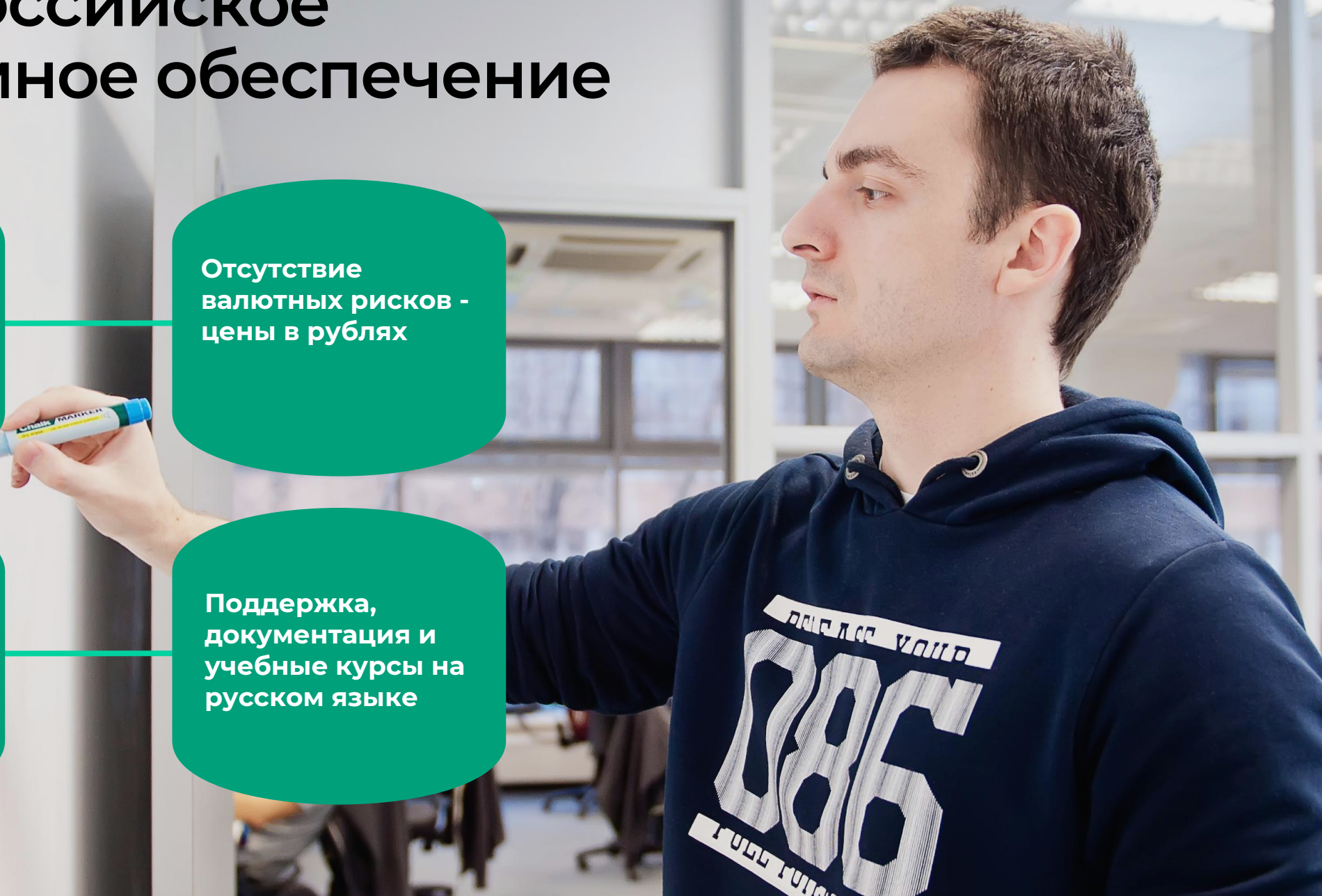
ADS — российское программное обеспечение

Регистрация в
Реестре
Министерства
Связи

Отсутствие
валютных рисков -
цены в рублях

Центры разработки
в РФ - Москва,
Санкт-Петербург,
Хабаровск

Поддержка,
документация и
учебные курсы на
русском языке



Вендорский консалтинг и профессиональные сервисы



Архитектурный контроль

Мы предоставляем проверенную вендорскую точку зрения на проект внедрения, позволяющую принять правильное архитектурное решение в непростых ситуациях.



Аудит цифрового ландшафта

Позволяет ИТ и менеджменту научиться отвечать на вопрос «Как должен измениться ландшафт данных и цифровых сервисов, чтобы обеспечить мои бизнес-потребности сейчас и в будущем?».



Smart Start

Это комплекс услуг по установке и настройке платформы Arenadata и интеграции наших продуктов в инфраструктуру данных заказчика. Мы также проводим инструктаж для администраторов, и выполняем ряд не столь значительных, но полезных вспомогательных действий.



Разгрузка дорогостоящих СУБД

Как вендор, мы поможем вам снизить TCO владения проприетарными СУБД за счёт внедрения и интеграции аналитических MPP-систем под ключ.

Поддержка 24x7

Санкт-Петербург

Москва

Алма-Аты

Хабаровск

Сотрудники
распределены
по часовым
поясам

Услуги по поддержке клиентов

- Vendor support

VIP-пакет

- Выделенный сервис-менеджер
- Регулярный аудит
- Обновление минорных версий



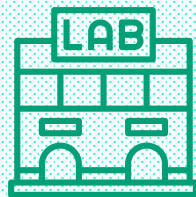
Корпоративная платформа хранения
и обработки больших данных

Курсы по Arenadata Streaming



Знания из первых рук

Наши преподаватели — практикующие специалисты, имеющие за плечами богатую экспертизу и значительное количество реализованных бизнес-кейсов.



Только 40% теории

Все курсы построены на реальных кейсах и практических примерах, а лабораторные работы проходят на стендах, симулирующих рабочую среду.



Сертификация

Пройдя обучение на курсах Arenadata и успешно сдав экзамен, слушатель станет обладателем именного сертификата, подтверждающего полученные знания.

Нам доверяют



Истории успеха



Организация «Умного озера данных» для комплексной системы анализа и управления данными Дирекции региональных продаж на базе Arenadata Hadoop.



Создание единой технологической платформы, способной объединить хранилища данных, создать систему отчётности и аналитики по миллионам клиентов, сокращение TCO на базе массово-параллельной СУБД Arenadata DB и Arenadata Streaming.



Создание платформы данных по всем данным компании с привязкой к системе Data Governance, способной выдерживать высокую нагрузку от пользователей (2000 аналитиков и более 800 конкурентных запросов) с помощью Arenadata DB, Arenadata QuickMarts и Arenadata Streaming.



Построение «Озера данных» для осуществления внешнего государственного аудита на новом качественном уровне. Используются данные всех возможных типов из совершенно разных источников, перечень которых заранее не определён. «Озеро данных» построено с помощью Arenadata Hadoop и Arenadata DB.



Разработка цифровых сервисов и моделей машинного обучения, создание удобного инструмента для вычислений и анализа, а также оптимизация производственных процессов и снижение себестоимости готовой продукции на базе Arenadata Hadoop.



Создание аналитического хранилища, отдельного от системы поддержки управленческой отчётности, объединение трёх «песочниц» и их миграция с SAP HANA на Arenadata DB.

3 шага к платформе Arenadata

1

Расскажите нам о задачах, ИТ-инфраструктуре и получите рекомендации по компонентам платформы и выбору партнёра.

2

Загрузите бесплатно необходимые пакеты для платформы данных с нашего сайта.

3

Разверните нашу платформу в своей инфраструктуре и отправьте нам технические и функциональные требования, чтобы получить индивидуальное предложение.





Корпоративная платформа хранения
и обработки больших данных

Оцените потенциал Big Data и Open Source вместе с Arenadata Enterprise Data Platform

- Узнайте больше на arenadata.tech
- Скачайте бесплатно на store.arenadata.io
- Получите консультацию по почте info@arenadata.io

