

Arenadata™ Streaming

Версия - v1.3-RUS

Установка с помощью ADCM

Оглавление

1	Загрузка дистрибутива платформы	3
2	Предварительные действия	6
3	Загрузка бандла ADS	8
4	Создание кластера ADS	11
4.1	Создание экземпляра кластера	12
4.2	Конфигурация кластера	13
4.3	Добавление хостов	13
4.4	Добавление сервисов	14
4.5	Размещение компонентов сервисов на хостах	17
4.6	Установка кластера	18
5	Инструменты управления	25
5.1	Операции на уровне кластера	25
5.2	Операции на уровне сервиса	25
5.3	Запуск и остановка сервисов	25

Кластер **ADS** устанавливается при помощи **Arenadata Cluster Manager**, описание которого приведено на сайте: [ADCM](#). При этом обеспечивается следующее:

1. Вся логика по установке и конфигурированию находится внутри бандла ADS:
 - Пользователю нет необходимости вручную вводить множество команд в консоли;
 - Сконфигурированы все важные настройки ОС и сервисов.
2. Единый интерфейс доступа и управления конфигурациями:
 - Возможность управления несколькими сервисами через единый интерфейс;
 - Возможность управления несколькими кластерами через единый интерфейс.
3. Возможность использования как облачной, так и физической инфраструктуры;
4. Все rpm-пакеты для ADS берутся из репозитория Arenadata:
 - Все rpm протестированы;
 - Используется фиксированный набор пакетов в помощь пользователям и для составления баг-репортов.

Глава 1

Загрузка дистрибутива платформы

Для загрузки дистрибутива **Arenadata Streaming** необходимо на сайте arenadata.io выбрать вкладку “Скачать”. При этом происходит переход на страницу *Магазин Программного Обеспечения Arenadata* (store.arenadata.io, Рис.1.3), и открывается экранная форма заявки для загрузки продукта (Рис.1.1).

Заполните, пожалуйста, форму для продолжения загрузки

ФИО:

E-Mail:

Компания:

Цель использования:

Отправить

Рис.1.1.: Заявка для загрузки продукта

В экранной форме заявки следует заполнить следующие поля:

- *ФИО* – ФИО пользователя;
- *E-Mail* – адрес электронной почты пользователя;
- *Компания* – наименование компании пользователя;
- *Цель использования* – описание цели использования платформы.

Все поля являются обязательными для заполнения. После ввода данных необходимо нажать кнопку *Отправить*, в результате чего выдается соответствующее сообщение (Рис.1.2).

При закрытии окна происходит переход на страницу *Магазин Приложений* с перечнем продуктов **Arenadata** (Рис.1.3).

При наведении курсора на продукт **ARENADATA STREAMING** предоставляется выбор действий:

Спасибо! Вы можете теперь закрыть окно и продолжить.



Рис.1.2.: Заявка принята

ARENADATA Продукты Документация Поддержка Утилиты lis@arenadata.io

Магазин Приложений

В данном разделе вы можете загрузить программное обеспечение компании Arenadata, а так же получить электронную версию документации по продуктам.

Управление сервисами:

ARENADATA AMBARI

Хранение и обработка данных:

ARENADATA HADOOP **ARENADATA DB** **ARENADATA GRID**

ARENADATA STREAMING

Платформа для обработки потоковых данных

[Скачать Документация](#)

Рис.1.3.: Магазин Приложений

- *Скачать* – переход на страницу загрузки дистрибутива **ADS**;
- *Документация* – переход на страницу он-лайн документации **ADS**.

Для загрузки дистрибутива платформы следует выбрать пункт *Скачать*, при этом происходит переход на соответствующую страницу с кратким описанием продукта, ссылками на компоненты дистрибутива и информацией о релизе (Рис.1.4.).

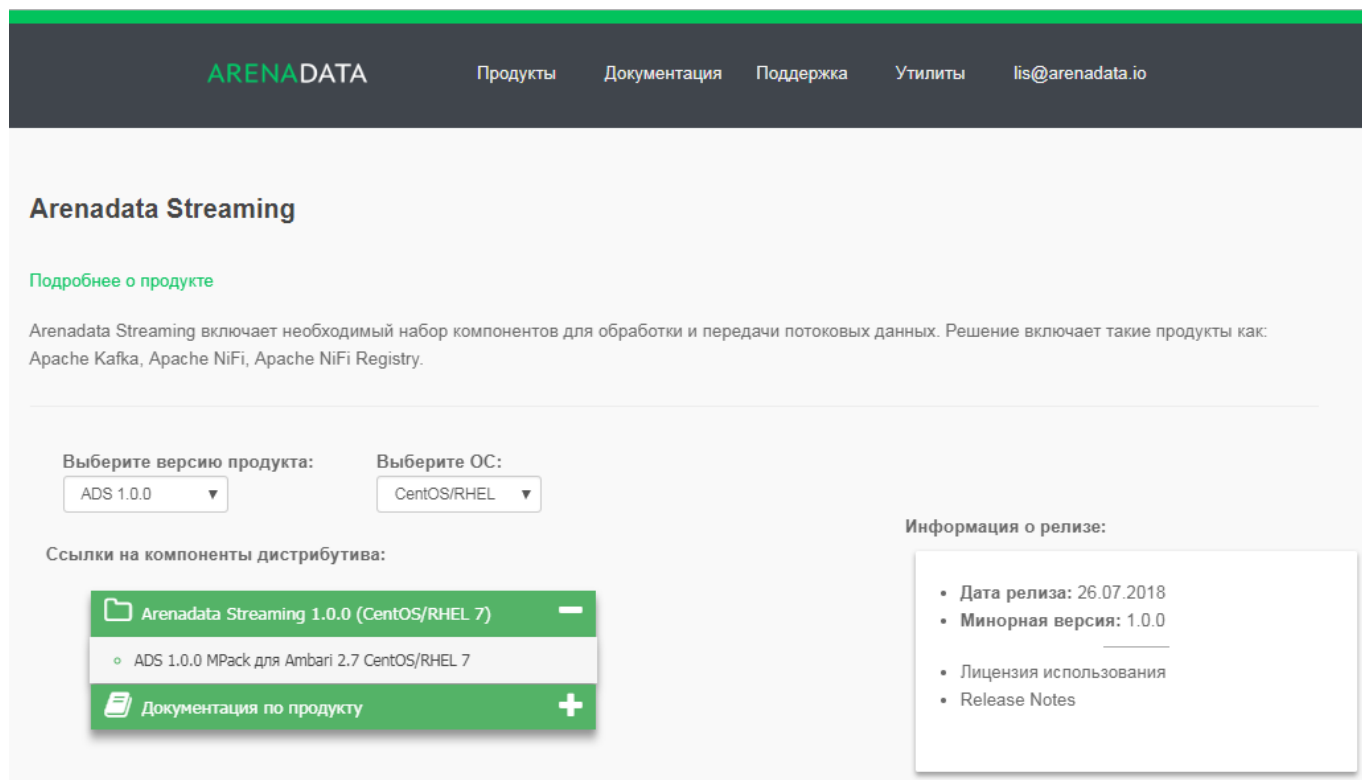


Рис.1.4.: Страница загрузки Arenadata Streaming

В разделе *“Выберите версию продукта”* следует указать интересующую версию **ADS**, а в поле *“Выберите ОС”* – используемую операционную систему. При этом в разделе *“Информация о релизе”* отображается дата выхода и тип релиза выбранной версии, ее номер, ссылки на лицензию использования и Release Notes.

В разделе *Ссылки на компоненты дистрибутива* находятся ссылки для загрузки компонентов выбранной версии платформы. После загрузки компонентов следует установить их, действуя в соответствии с последующими разделами настоящего документа.

Глава 2

Предварительные действия

Для установки ADS посредством ADCM необходимо:

1. Установить ADCM.
2. Проверить настройки ADCM:
 - Перейти во вкладку “Settings”;
 - Проверить на корректность определившийся автоматически URL и при необходимости заменить его на исправный.

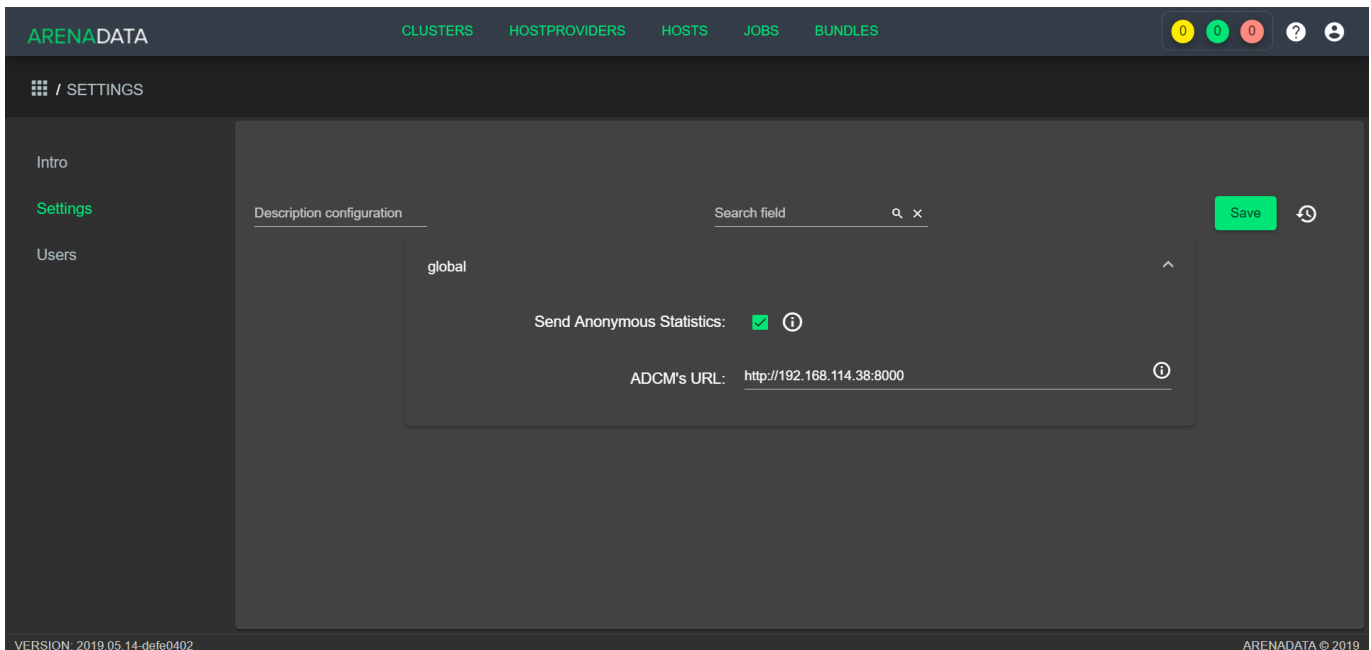


Рис.2.1.: Вкладка “Settings”

Important: Каждый компонент сервиса кластера имеет возможность отсылать статусную информацию о своем состоянии (keep alive) процессу ADCM в докер-контейнере. В ряде случаев ADCM может оказаться за NAT, и тогда исключается очевидный вариант автоматического определения его адреса, видимого со стороны сервисного компонента на хосте кластера. Поэтому данный адрес указывается вручную. Во время установки

ADS адрес ADCM используется для заполнения конфигурации компонентов, отвечающих за передачу статусной информации

3. Создать хосты для кластера ADS:

- Загрузить выбранный бандл хоста. В текущем примере используется бандл *SSH*;
- Инициализировать необходимое количество хостов (в текущем примере данный шаг пропускается, так как используются готовые хосты, а не облачный провайдер);
- Добавить хосты в ADCM. В текущем примере используется 4 хоста: для Zookeeper (*zk*), для Kafka брокеров (*kafka1* и *kafka2*), для Nifi (*nifi*).

4. (Опционально) Создать кластер мониторинга:

- Загрузить бандл мониторинга;
- Создать экземпляр кластера мониторинга и установить его.

Глава 3

Загрузка бандла ADS

Загрузка бандла **ADS** необходима для создания в **ADCM** прототипа кластера, из которого в дальнейшем возможно создание его экземпляров.

Для загрузки бандла следует выполнить следующие действия:

1. Открыть в **ADCM** вкладку “**BUNDLES**” (Рис.3.1.).

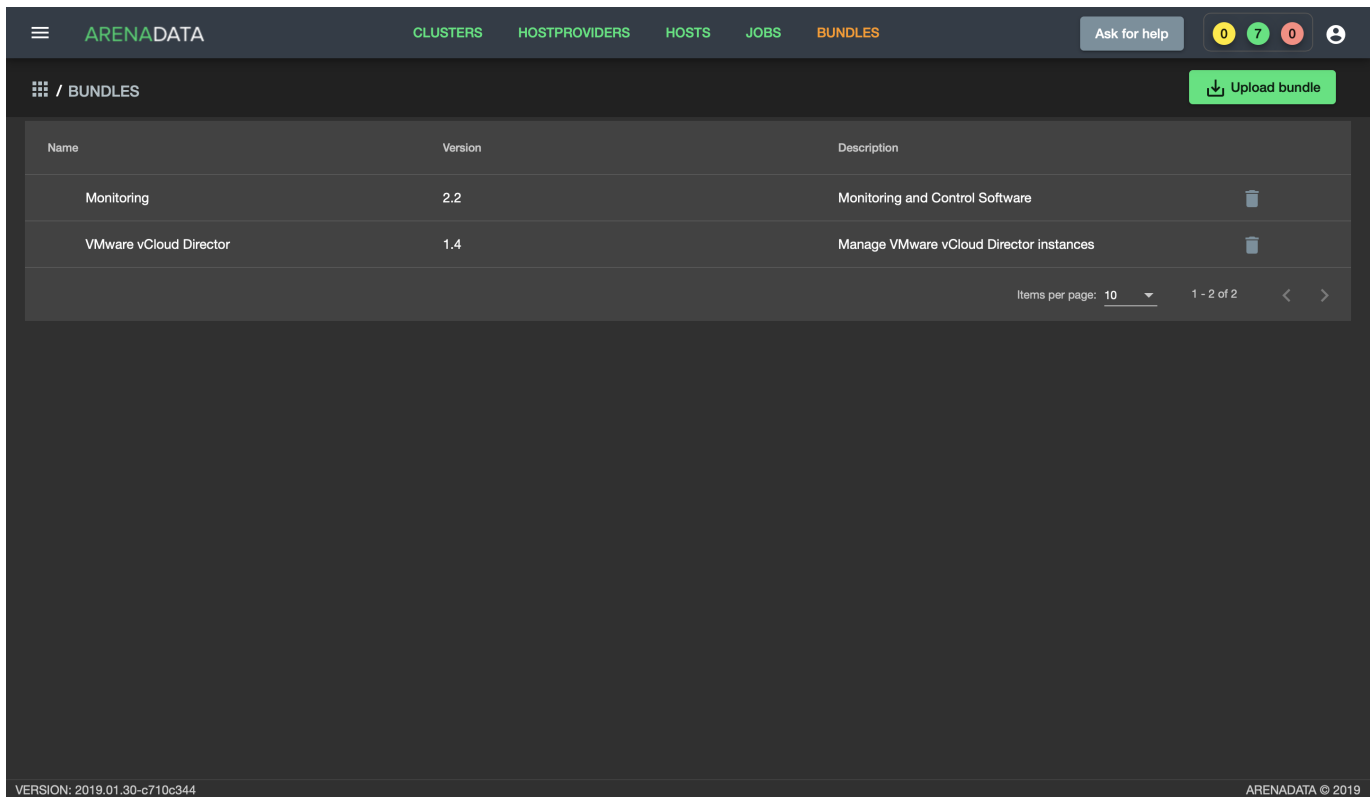


Рис.3.1.: Вкладка “BUNDLES”

2. Нажать “Upload bundle” и в открывшейся форме выбрать файл бандла **ADS** (Рис.3.2.).
3. В результате выполненных действий факт успешной загрузки отображается в общем списке бандлов на вкладке “**BUNDLES**” (Рис.3.3.).

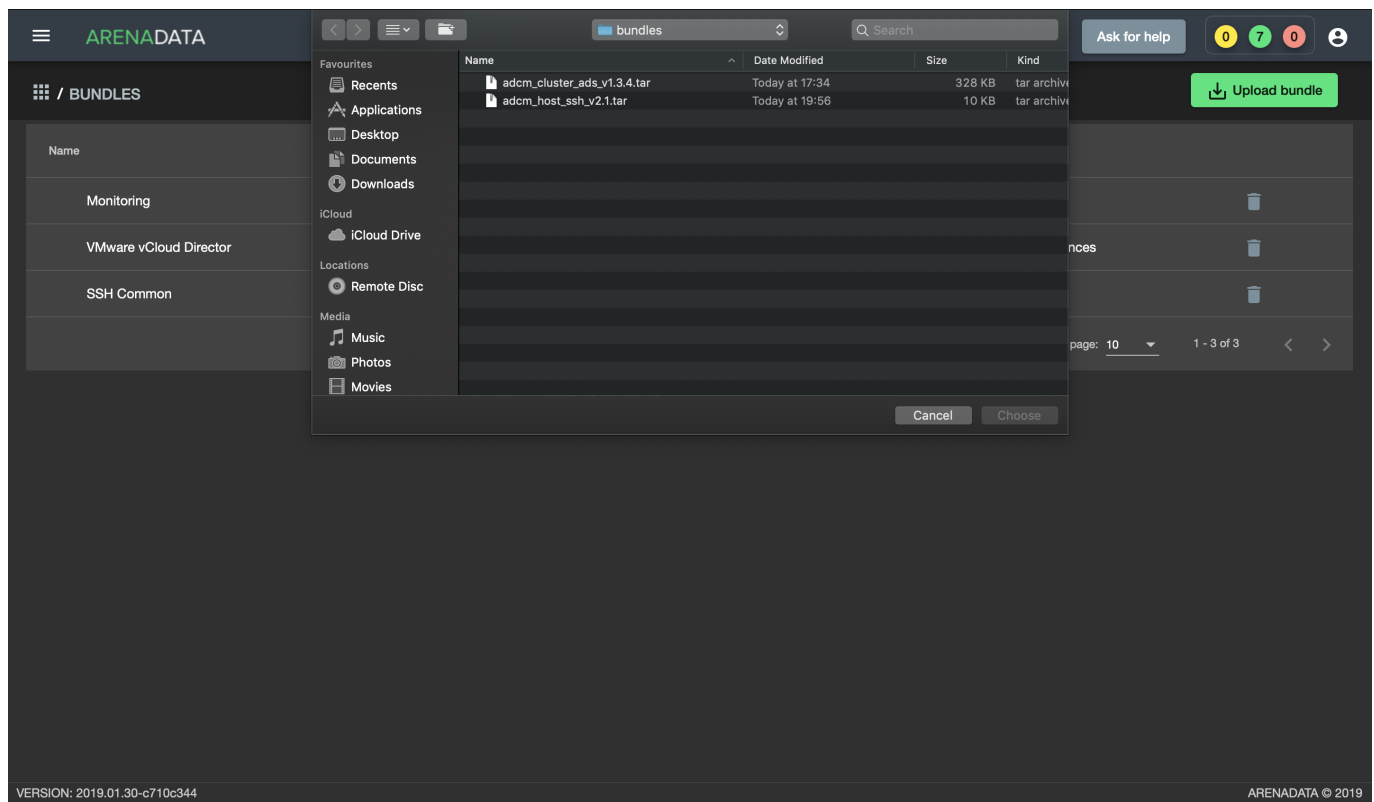


Рис.3.2.: Выбор бандла

The screenshot displays the ARENADATA web interface. At the top, there is a navigation bar with the ARENADATA logo and several menu items: CLUSTERS, HOSTPROVIDERS, HOSTS, JOBS, and BUNDLES. A utility bar on the right contains an 'Ask for help' button and three status indicators (0, 7, 0) along with a user profile icon. Below the navigation bar, the page title is '/ BUNDLES' and there is an 'Upload bundle' button. The main content area features a table with the following data:

Name	Version	Description	
Monitoring	2.2	Monitoring and Control Software	
VMware vCloud Director	1.4	Manage VMware vCloud Director instances	
SSH Common	2.1	Simple ssh hosts	
ADS	1.3.4	Arenadata Streaming Software includes Nifi, Kafka, Zookeeper and Monitoring Clients.	

At the bottom right of the table, there is a pagination control showing 'Items per page: 10' and '1 - 4 of 4' with navigation arrows. The footer of the interface contains the text 'VERSION: 2019.01.30-c710c344' on the left and 'ARENADATA © 2019' on the right.

Рис.3.3.: Результат успешной загрузки бандла

Глава 4

Создание кластера ADS

После выполнения *предварительных действий* и *загрузки бандла* в кластер-менеджере ADCM содержится следующий список объектов (Рис.4.1.):

- Прототип кластера мониторинга и созданный на его основе экземпляр (программа мониторинга развернута на хосте, ADCM содержит записи о его результатах и настройках);
- Прототип кластера *ADS* для создания экземпляров;
- Предварительно подготовленные хосты, в базе данных ADCM хранятся записи о них и их учетных данных – ssh-ключом или паролям.

Fqdn	Provider	Cluster	State	Config	Actions
monitoring	ella	Monitoring	running		
zk	ssh	Assign to cluster	created		
kafka2	ssh	Assign to cluster	created		
nifi1	ssh	Assign to cluster	created		
kafka1	ssh	Assign to cluster	created		

Рис.4.1.: Список объектов в ADCM

Объектам доступен следующий функционал:

- *Создание экземпляра кластера;*
- *Конфигурация кластера;*
- *Добавление хостов;*
- *Добавление сервисов;*
- *Размещение компонентов сервисов на хостах;*
- *Установка кластера.*

4.1 Создание экземпляра кластера

При создании кластера в веб-интерфейсе **ADCM** генерируется новый экземпляр кластера *ADS*, что означает только добавление данных о нем в базу данных **ADCM** – на этом этапе не производится установка *ADS* на хосты.

1. Открыть в **ADCM** вкладку “CLUSTERS” (Рис.4.2.).

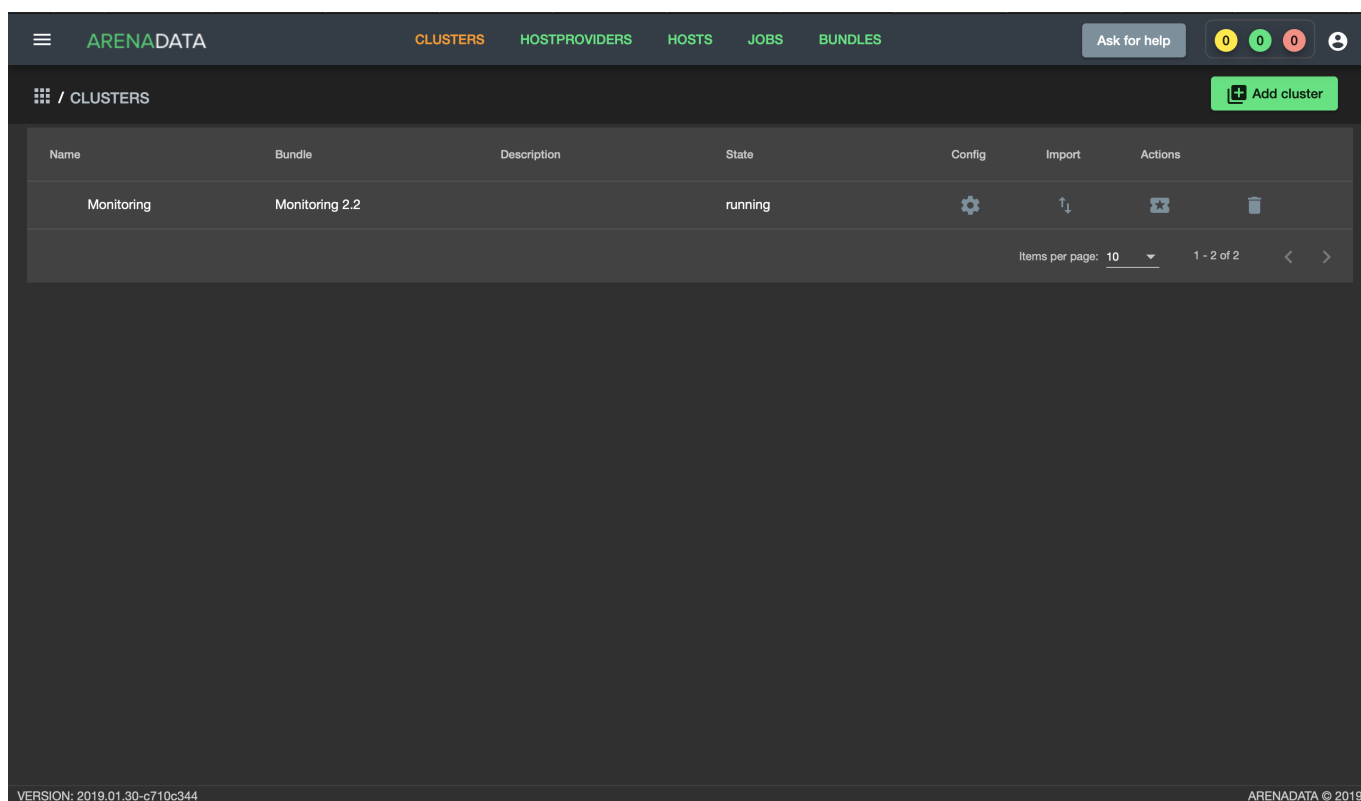


Рис.4.2.: Вкладка “CLUSTERS”

2. Нажать “Add cluster” и в открывшейся форме создать экземпляр кластера из прототипа *ADS*, полученного из бандла (Рис.4.3.).
3. В результате выполненных действий факт создания экземпляра кластера отображается в базе данных **ADCM** на вкладке “CLUSTERS” (Рис.4.4.).

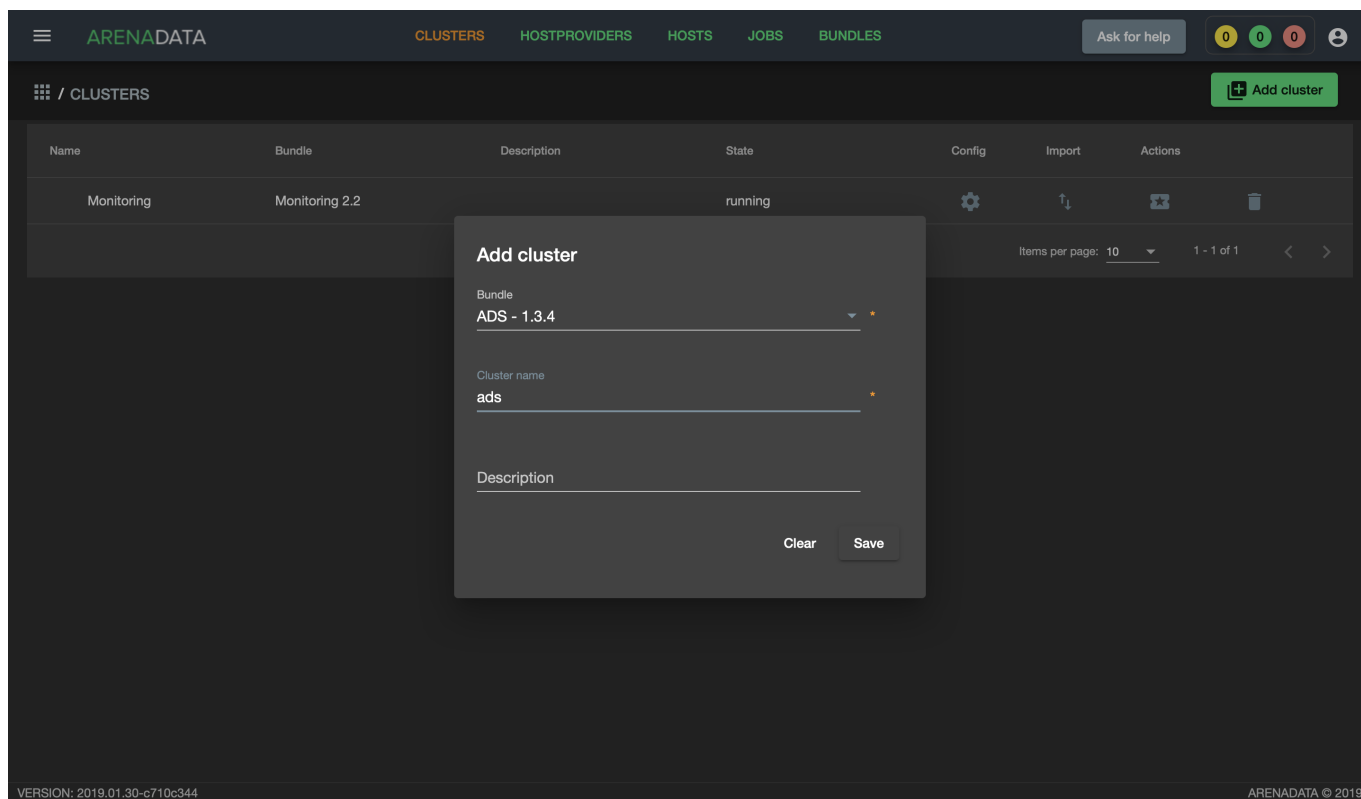


Рис.4.3.: Создание экземпляра кластера

4.2 Конфигурация кластера

Для перехода к настройкам экземпляра кластера *ADS* необходимо нажать кнопку с пиктограммой шестеренки в соответствующей строке вкладки “CLUSTERS” (Рис.4.4.) и перейти в раздел меню “Configuration”. При этом открывается окно конфигурации выбранного экземпляра (Рис.4.5.).

В блоке настроек “Repositories” указываются требуемые для установки *ADS* пакеты из различных ушп-репозиториях, при этом в каждом из параметров можно изменить заданный по умолчанию url:

- *ADS*;
- *monitoring*;
- *epel*.

Если нет необходимости в добавлении того или иного ушп-репозитория, то соответствующие галочки следует снять.

4.3 Добавление хостов

Для добавления хостов в кластер *ADS* необходимо:

1. В меню кластера *ADS* открыть вкладку “Hosts” (Рис.4.6.).
2. Нажать “Add hosts” и в открывшейся форме выбрать необходимые хосты (Рис.4.7.).
3. В результате выполненных действий факт добавления хостов отображается в кластере *ADS* в списке вкладки “Hosts” (Рис.4.8.).

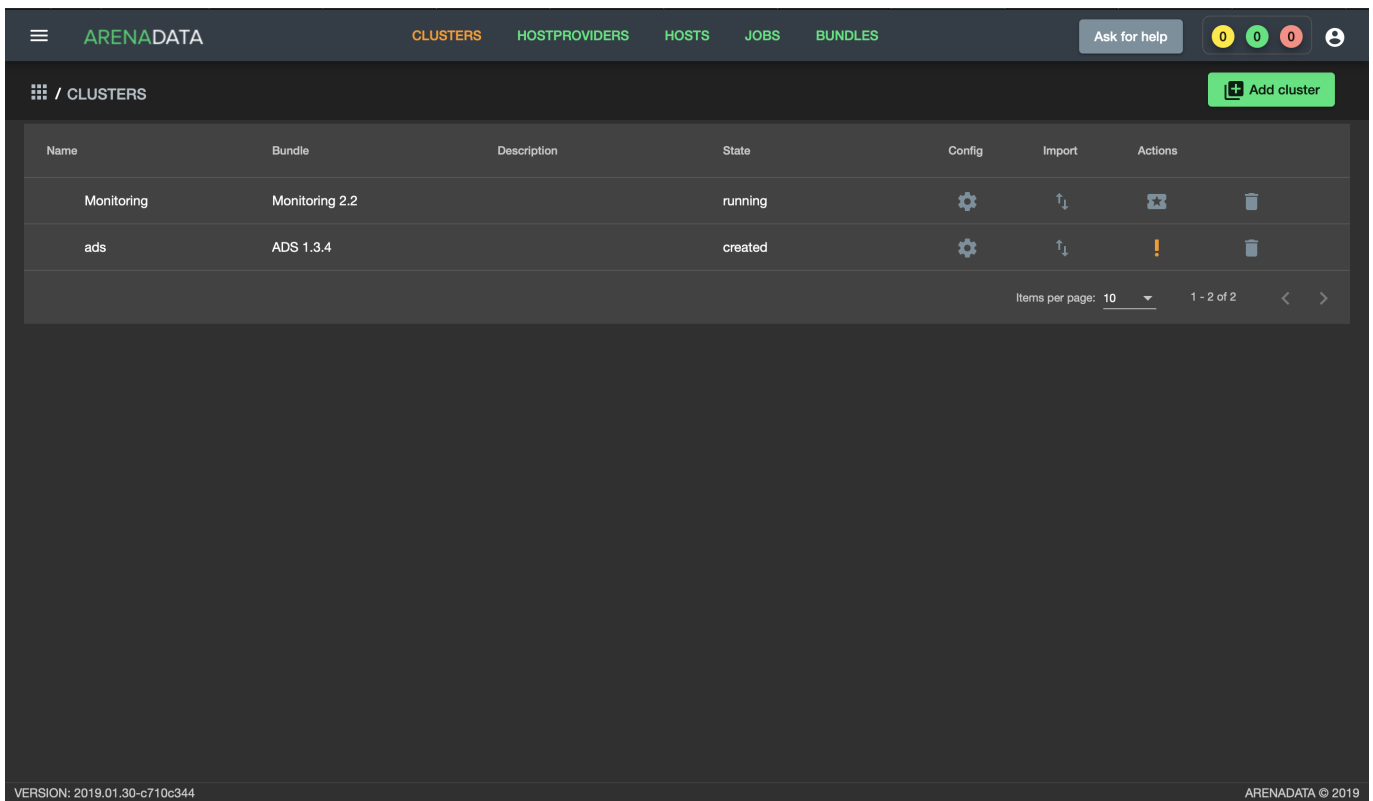


Рис.4.4.: Результат успешного создания экземпляра кластера

4.4 Добавление сервисов

Кластер **ADS** содержит следующие сервисы:

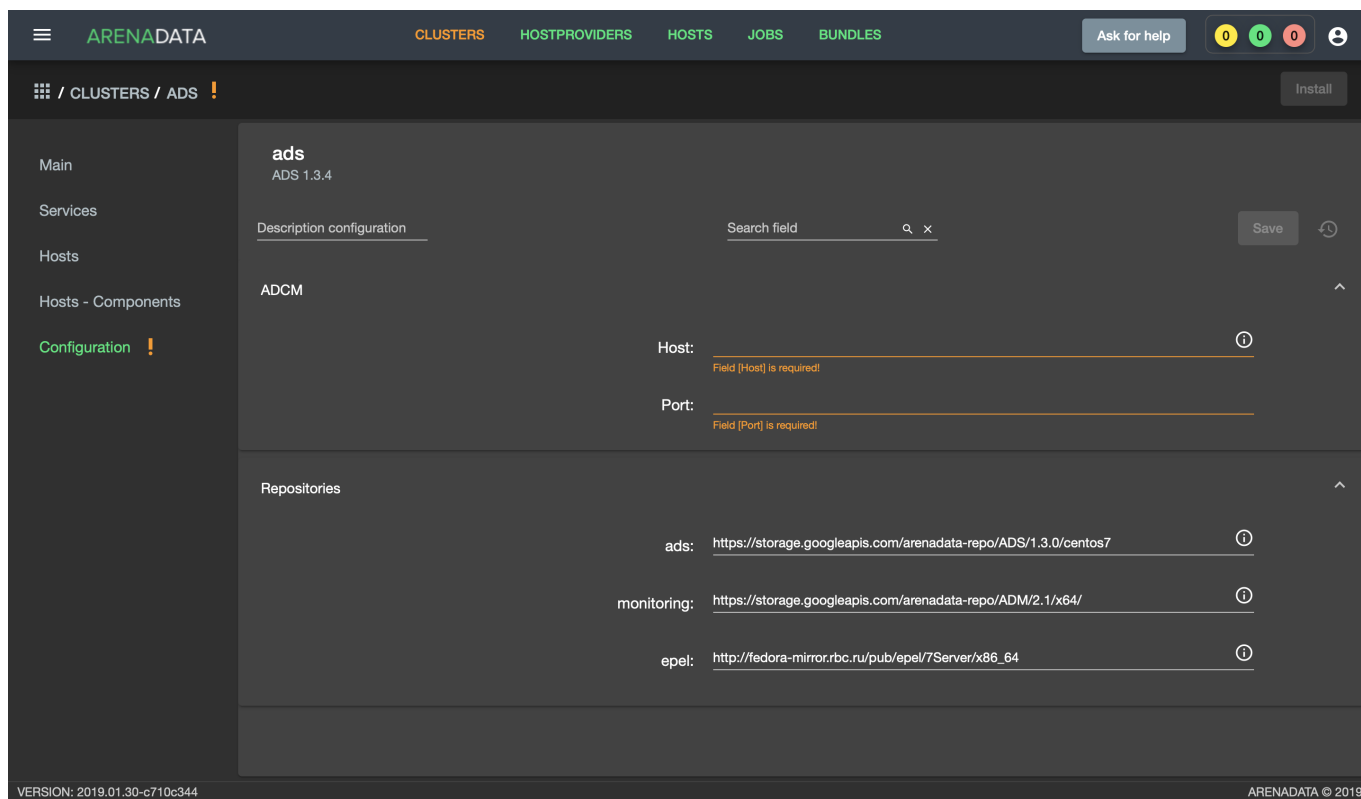


Рис.4.5.: Окно конфигурации кластера

Таблица4.1.: Сервисы ADS

Сервис	Компоненты	Описание
Zookeeper	Zookeeper Server	Сервис, предназначенный для хранения конфигураций, выполнения распределенной синхронизации процессов
Kafka	Kafka Broker	Распределенная платформа для потоковых операций и данных
Nifi	Nifi Server, Nifi Registry	Распределенная платформа, предназначенная для построения и автоматизации потоков данных между различными системами
Schema-registry	Schema-registry	Сервис предоставляет обслуживающий слой для метаданных. Предоставляет интерфейс RESTful для хранения и извлечения схем Avro. Хранит версию историю всех схем, предоставляет несколько параметров совместимости и позволяет изменять схемы в соответствии с настроенными параметрами совместимости и расширенной поддержкой Avro. Предоставляет сериализаторы, подключаемые к клиентам Kafka, которые обрабатывают хранение и извлечение схемы для сообщений Kafka, отправляемых в формате Avro
Kafka REST Proxy	Kafka REST Proxy	Предоставляет RESTful-интерфейс для кластера Kafka, что упрощает создание и потребление сообщений, просмотр состояния кластера и выполнение административных действий без использования собственного протокола Kafka или клиентов
KSQL	KSQL Server, KSQL Client	Сервис представляет собой потоковый SQL-движок с открытым исходным кодом, который обеспечивает обработку данных в реальном времени с использованием Apache Kafka. Состоит из компонентов: Server; Client
Kafka-Manager	Kafka-Manager	Инструмент для управления Apache Kafka
Monitoring Clients	System metrics Agent, JMX metrics Collector	Агенты, отсылающие информацию о хосте и сервисах в мониторинг

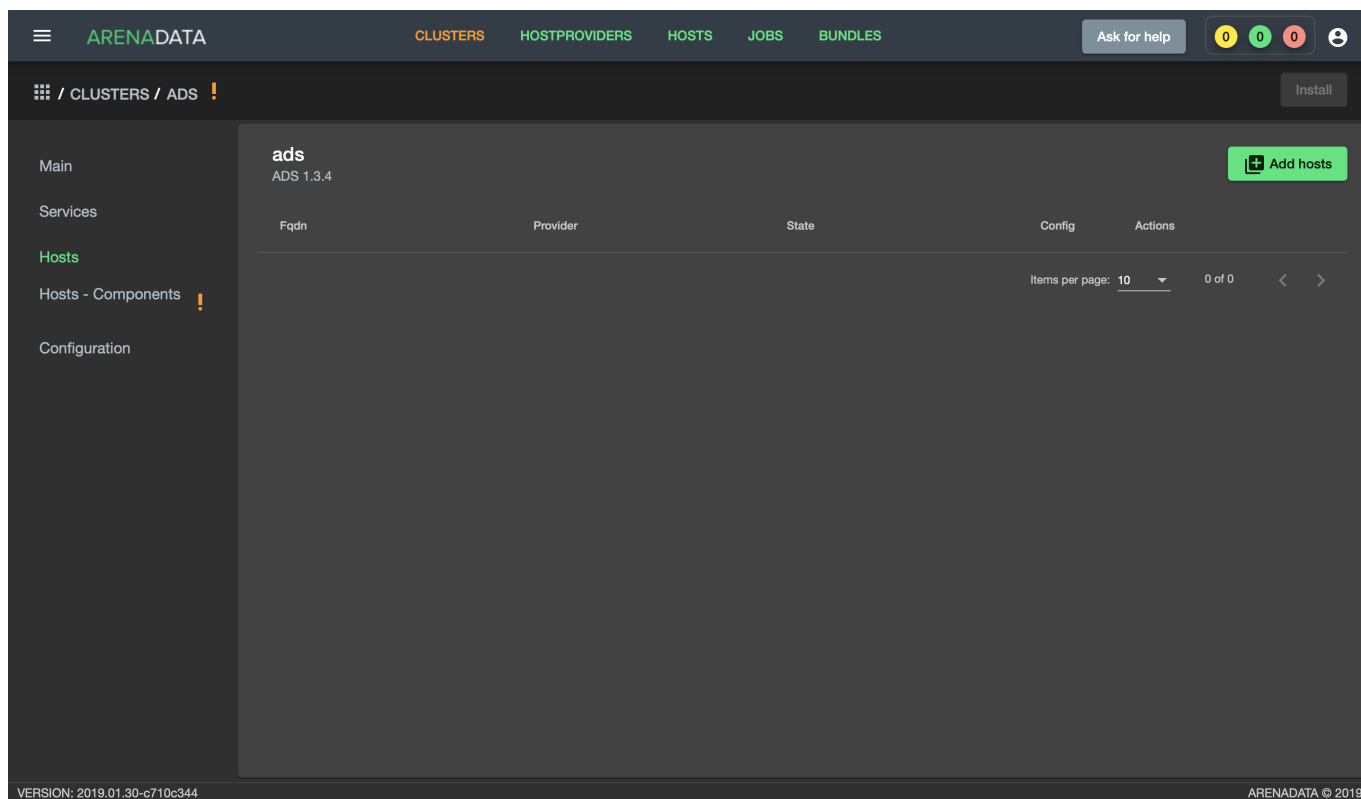


Рис.4.6.: Вкладка “Hosts” кластера ADS

Не все сервисы являются обязательными для установки. Например, если не планируется использование **Nifi**, то нет необходимости добавлять сервис. Или в случае, когда применяется сервис мониторинга (не на базе **Graphite**), незачем ставить агенты из *Monitoring Clients*. Однако, если планируется использование **Kafka**, одноименный сервис и *Zookeeper* обязательны, то же самое можно сказать и про сервис *Nifi*. При этом сервис может состоять из обязательных и необязательных компонентов. Например, сервис *KSQL* состоит из обязательного компонента *Server* и необязательного – *Client*.

Для добавления сервисов в кластер *ADS* необходимо:

1. В меню кластера *ADS* открыть вкладку “Services”.
2. Нажать “Add service” и в открывшейся форме выбрать необходимые сервисы (Рис.4.9).

Important: На текущий момент невозможно удалить из кластера уже добавленный сервис

В настоящем примере в кластер добавлены все сервисы, подробное описание конфигурации которых приведено по следующим ссылкам:

- [Настройка сервиса Zookeeper;](#)
- [Настройка сервиса Kafka;](#)
- [Настройка сервиса Nifi;](#)
- [Настройка сервиса Schema-registry.](#)
- [Настройка сервиса Kafka REST Proxy;](#)
- [Настройка сервиса KSQL;](#)

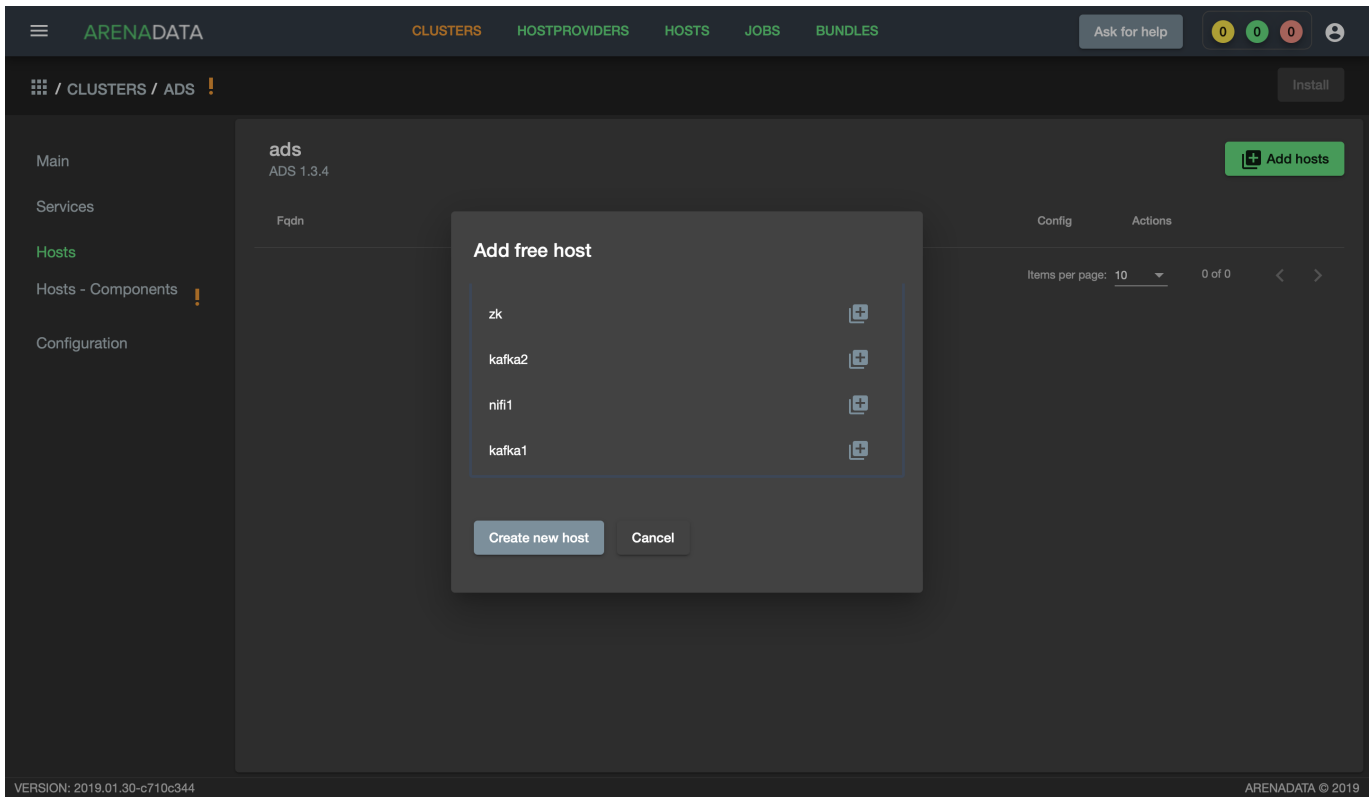


Рис.4.7.: Выбор хостов

- Настройка сервиса Kafka-Manager;
- Настройка сервиса Monitoring Clients.

4.4.1 Особенности сервиса Monitoring Clients

Сервис *Monitoring Clients* требует импорта конфигурационных параметров кластера мониторинга (адреса, логин/пароль) при добавлении в кластер *ADS*.

Для импорта конфигурации мониторинга необходимо открыть в **ADCM** вкладку “CLUSTERS”, выбрать опцию *Import* и отметить импортируемые настройки сервисов с помощью простановки флажков в открывшейся форме (Рис.4.10.).

4.5 Размещение компонентов сервисов на хостах

Каждый сервис состоит из обязательных компонентов, которые должны быть размещены на хостах кластера, и необязательных, которые могут быть не размещены. Для размещения компонентов необходимо на вкладке “Hosts - Components” выбрать компонент в колонке “Components” и определить для него необходимый хост в колонке “Hosts” (Рис.4.11.).

Поскольку все сервисы добавлены в кластер **ADS**, но еще не размещены на хостах, то изначально ни на одном из хостов нет компонентов.

В качестве примера на Рис.4.12. показаны компоненты сервиса *Nifi*, размещенные на хосте *nifi*:

- *Nifi Server* – необходимо добавить на один или более хостов (возможна кластерная конфигурация);
- *Nifi Registry* – необходимо добавить ровно на один хост.

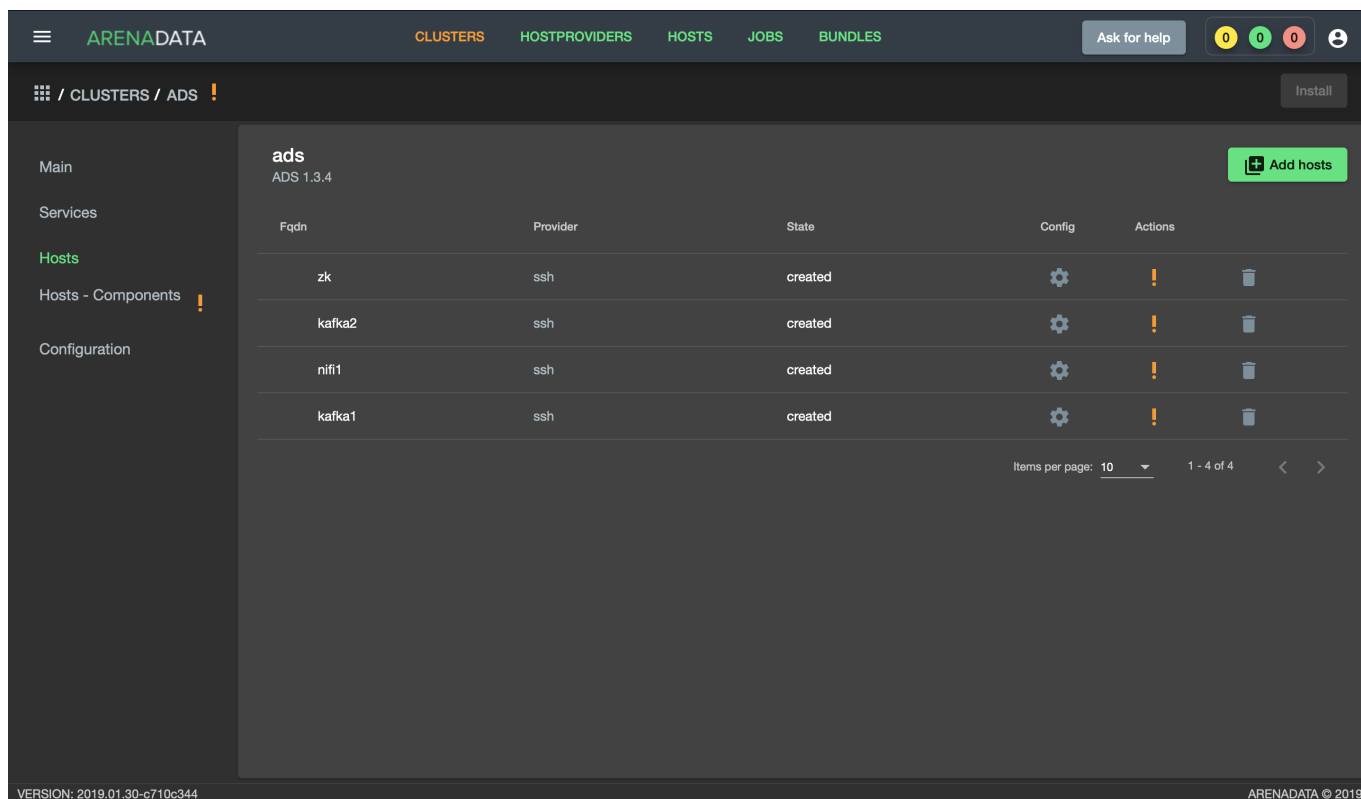


Рис.4.8.: Результат успешного добавления хостов

4.6 Установка кластера

Important: В случае если добавлен сервис *Monitoring Clients*, необходимо импортировать настройки из кластера *Monitoring* (см. *Особенности сервиса Monitoring Clients*)

Прежде чем приступить к установке кластера, необходимо в *ADS* запустить действие *Preinstall* в правом верхнем углу экрана формы (Рис.4.13.).

В результате чего во всплывающем меню следует выбрать необходимые действия из открывшегося списка и нажать *Run* (Рис.4.14.).

На данном этапе все хосты подготовлены и можно приступать непосредственно к установке сервисов кластера:

- *Установка всех сервисов кластера;*
- *Установка сервиса в проинсталлированный кластер.*

4.6.1 Установка всех сервисов кластера

Для установки всех добавленных сервисов в кластере *ADS* необходимо выбрать соответствующий кластер в *ADCM* и выполнить действие *Install* (Рис.4.15.).

По результатам установки все добавленные сервисы меняют состояние с *preinstalled*, на *installed* – установлен (Рис.4.16.).

Для запуска кластера необходимо нажать на кнопку *Start*.

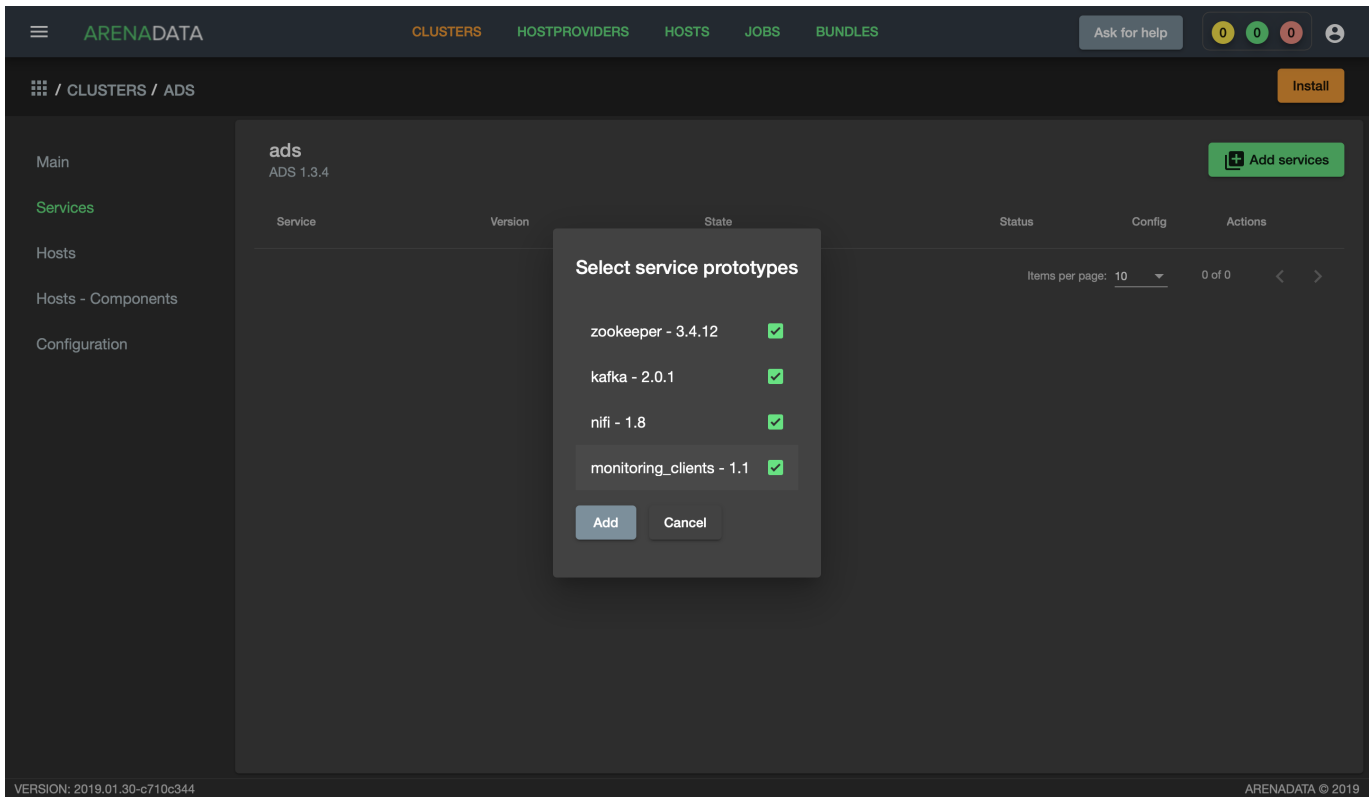


Рис.4.9.: Выбор сервисов

4.6.2 Установка сервиса в проинсталлированный кластер

В ADCM предусмотрена возможность добавления нового сервиса в уже работающий кластер. Для этого сервис необходимо добавить и произвести его установку. Например, для установки сервиса *Kafka* в проинсталлированный кластер необходимо:

- В меню кластера *ADS* открыть вкладку "Services", нажать "Add service" и в открывшейся форме выбрать сервис *Kafka*.
- В строке сервиса *Kafka* в поле "Actions" нажать на пиктограмму и выбрать действие *Install*.
- По результатам инсталляции сервис *Kafka* меняет состояние с *created* – создан, на *installed* – установлен.

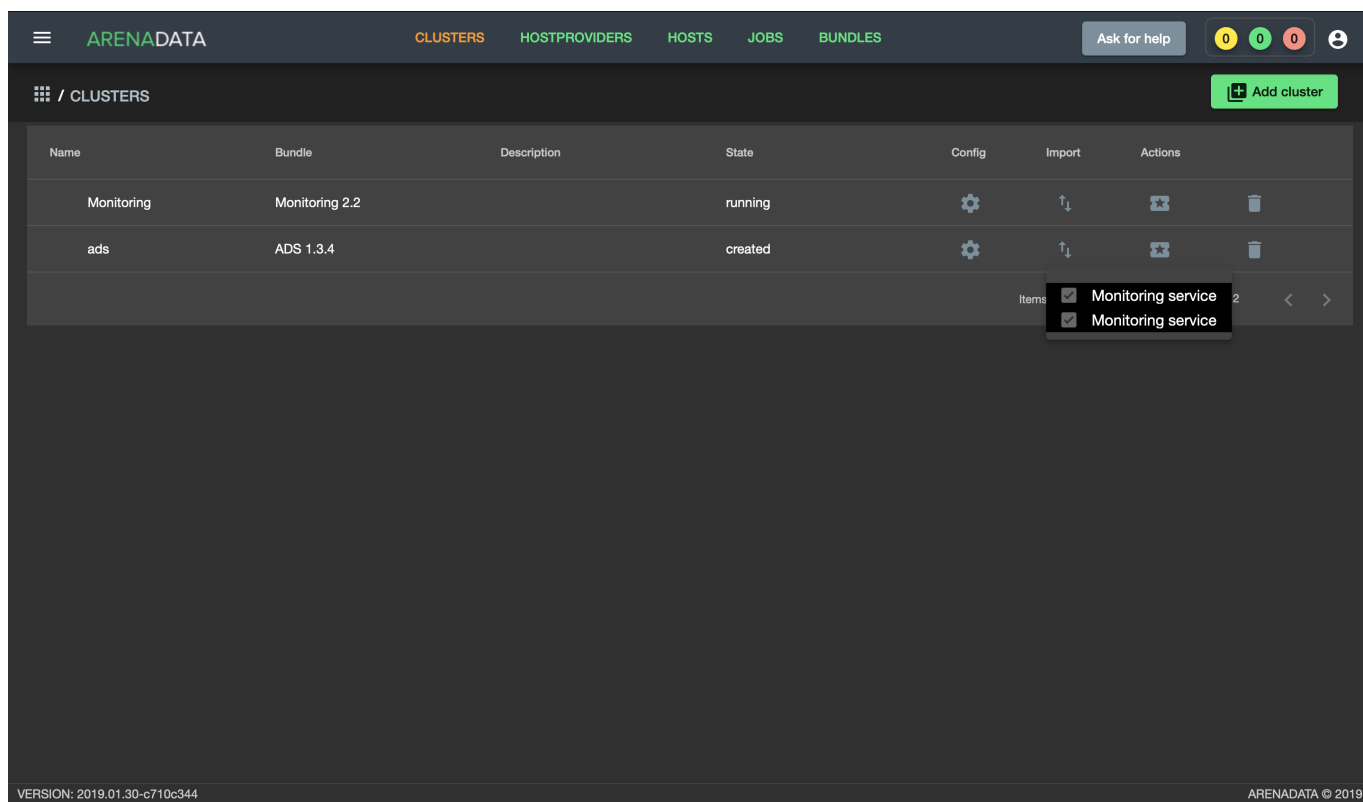


Рис.4.10.: Импорт конфигурации мониторинга

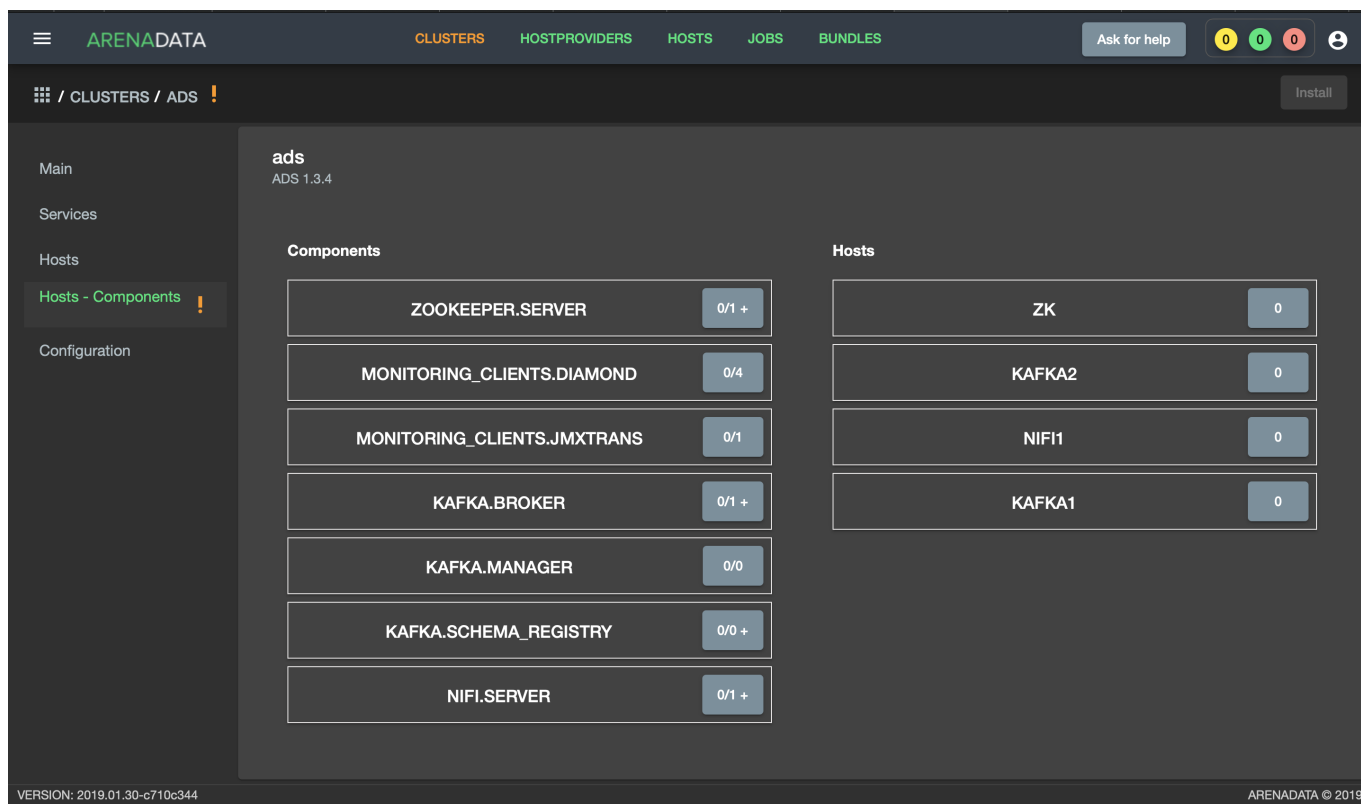


Рис.4.11.: Размещение компонентов сервисов на хостах

The screenshot displays the ARENA DATA web interface for configuring a cluster named 'ads' (version ADS 1.3.4). The interface is divided into a left sidebar with navigation options (Main, Services, Hosts, Hosts - Components, Configuration) and a main content area. The main content area is split into two columns: 'Components' and 'Hosts'. Each component or host entry includes a name and a count with a plus sign for expansion.

Component/Host	Count
ZOOKEEPER.SERVER	1/1 +
MONITORING_CLIENTS.DIAMOND	0/4
MONITORING_CLIENTS.JMXTRANS	0/1
KAFKA.BROKER	2/1 +
KAFKA.MANAGER	1/0
KAFKA.SCHEMA_REGISTRY	1/0 +
NIFI.SERVER	1/1 +
ZK	2
KAFKA2	1
NIFI1	1
KAFKA1	2

VERSION: 2019.01.30-c710c344 ARENADATA © 2019

Рис.4.12.: Компоненты сервиса *Nifi*

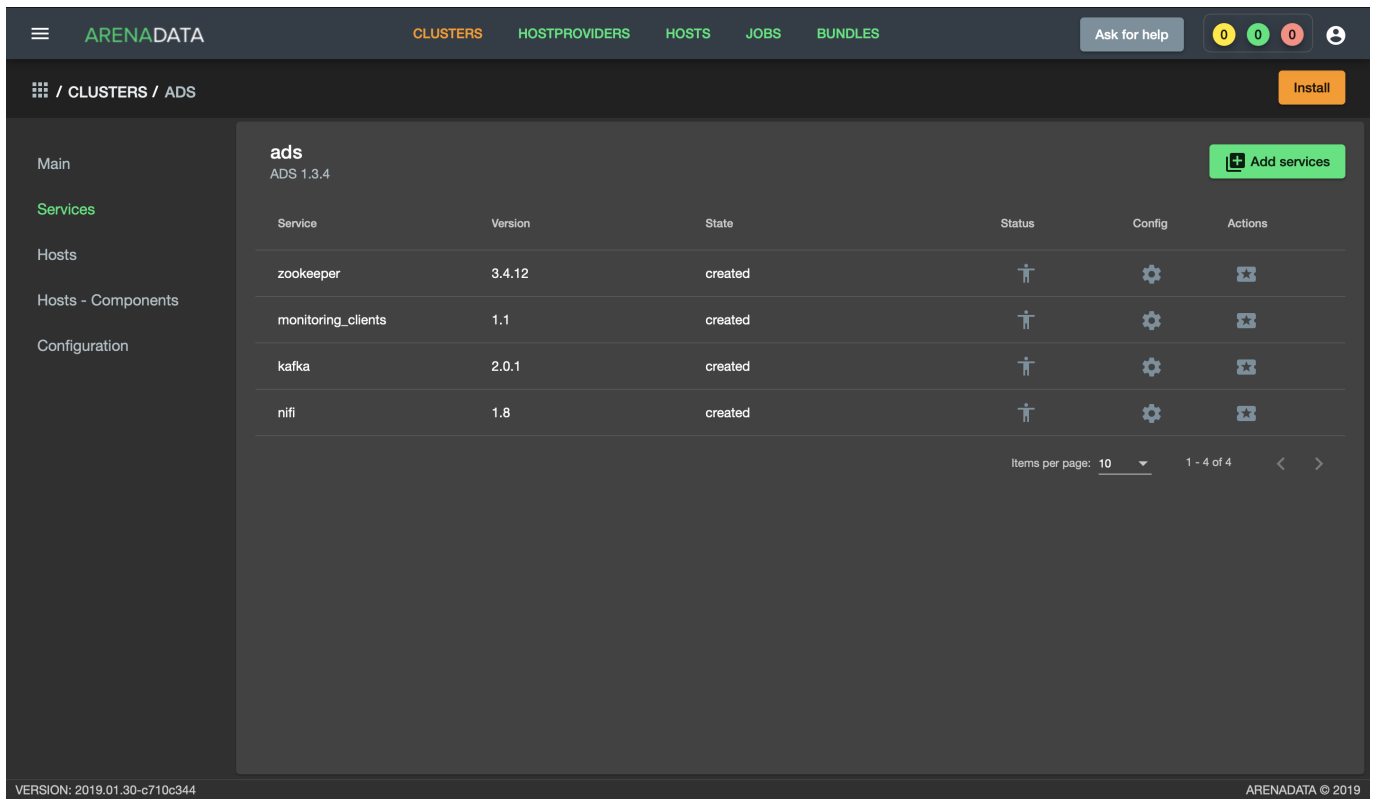


Рис.4.13.: Preinstall

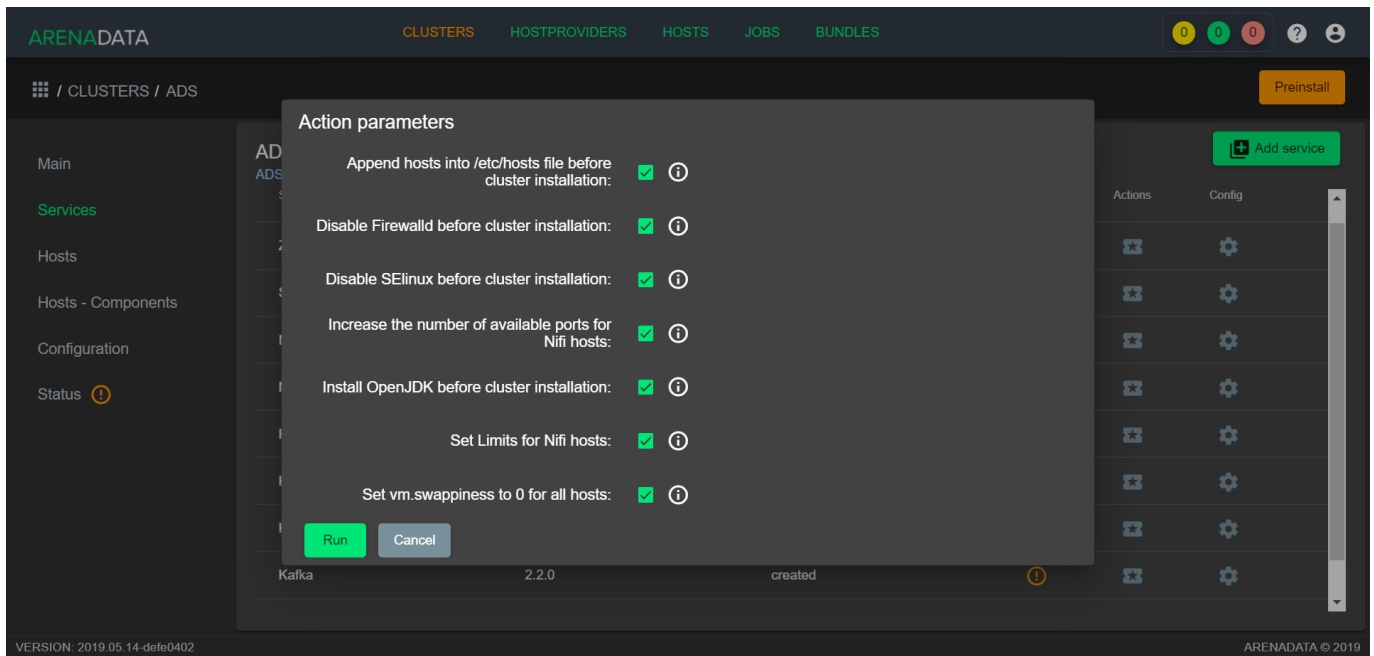


Рис.4.14.: Action parameters

ARENADATA CLUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES

0 0 0 ?

/ CLUSTERS / ADS Install

Main Services

ADS ADS 1.4.7.22 + Add service

Service	Version	State	Status	Actions	Config
Zookeeper	3.4.14	created	!	⊞	⚙
Schema-registry	5.2.1	created	!	⊞	⚙
Nifi	1.9.1	created	!	⊞	⚙
Monitoring Clients	1.1	created	!	⊞	⚙
KSQL	5.2.1	created	!	⊞	⚙
Kafka REST Proxy	5.2.1	created	!	⊞	⚙
Kafka-Manager	1.3.3.21	created	!	⊞	⚙
Kafka	2.2.0	created	!	⊞	⚙

VERSION: 2019.05.14-defe0402 ARENADATA © 2019

Рис.4.15.: Установка всех сервисов кластера

ARENADATA CLUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES

Ask for help 0 0 0

/ CLUSTERS / ADS Restart Start Stop

Main Services

ads ADS 1.3.4 + Add services

Service	Version ↑	State	Status	Config	Actions
kafka	2.0.1	installed	✔	⚙	⊞
zookeeper	3.4.12	installed	✔	⚙	⊞
monitoring_clients	1.1	installed	✔	⚙	⊞
nifi	1.8	installed	✔	⚙	⊞

Items per page: 10 1 - 4 of 4 < >

VERSION: 2019.01.30-c710c344 ARENADATA © 2019

Рис.4.16.: Состояние сервисов кластера

Глава 5

Инструменты управления

В ADS предусмотрено 2 вида операций:

- *Операции на уровне кластера* – предполагают выполнение операций последовательно над всеми сервисами;
- *Операции на уровне сервиса* – предполагают выполнение операций над отдельным сервисом.

5.1 Операции на уровне кластера

Запуск и остановка ADS – существует возможность последовательного запуска и остановки всех сервисов кластера ADS через ADCM. Для этого необходимо открыть в ADCM кластер ADS и нажать кнопку *Start*, *Stop* или *Restart* в зависимости от требуемой работы с кластером (Рис.5.1).

В результате того или иного действия кластер меняет свое состояние на соответствующее – *running* или *stopped*, в том случае если кластер уже проинсталлирован и не находится в состоянии *created*.

5.2 Операции на уровне сервиса

Для каждого из сервисов доступна возможность проверки его работоспособности, а также управления им независимо от остальных. Например, проверка работоспособности сервиса *Kafka* представляет собой создание тестовых топиков и проверку их доступности на каждом из хостов *BROKER*. А проверка работоспособности сервиса *Zookeeper* представляет собой подключение к кворуму *Zookeeper*, создание в нем тестовой *znode* и проверку доступности созданной *znode* каждому из хостов кворума.

Проверка состояний сервисов и вывод результатов действий над ними осуществляется по единому алгоритму, разобранному на примере сервиса *Zookeeper*:

1. В ADCM перейти в кластер ADS. На вкладке “Services” для сервиса *Zookeeper* в поле “Actions” нажать на пиктограмму и выбрать действие *Check* (Рис.5.2).
2. Открыть вкладку “JOBS” (Рис.5.3).
3. Выбрать последнее действие над кластером ADS и в открывшемся окне проверить результаты (Рис.5.4).

5.3 Запуск и остановка сервисов

Для каждого из сервисов есть возможность управления им независимо от остальных, выполняя такие операции как *Stop*, *Start*, *Restart*.

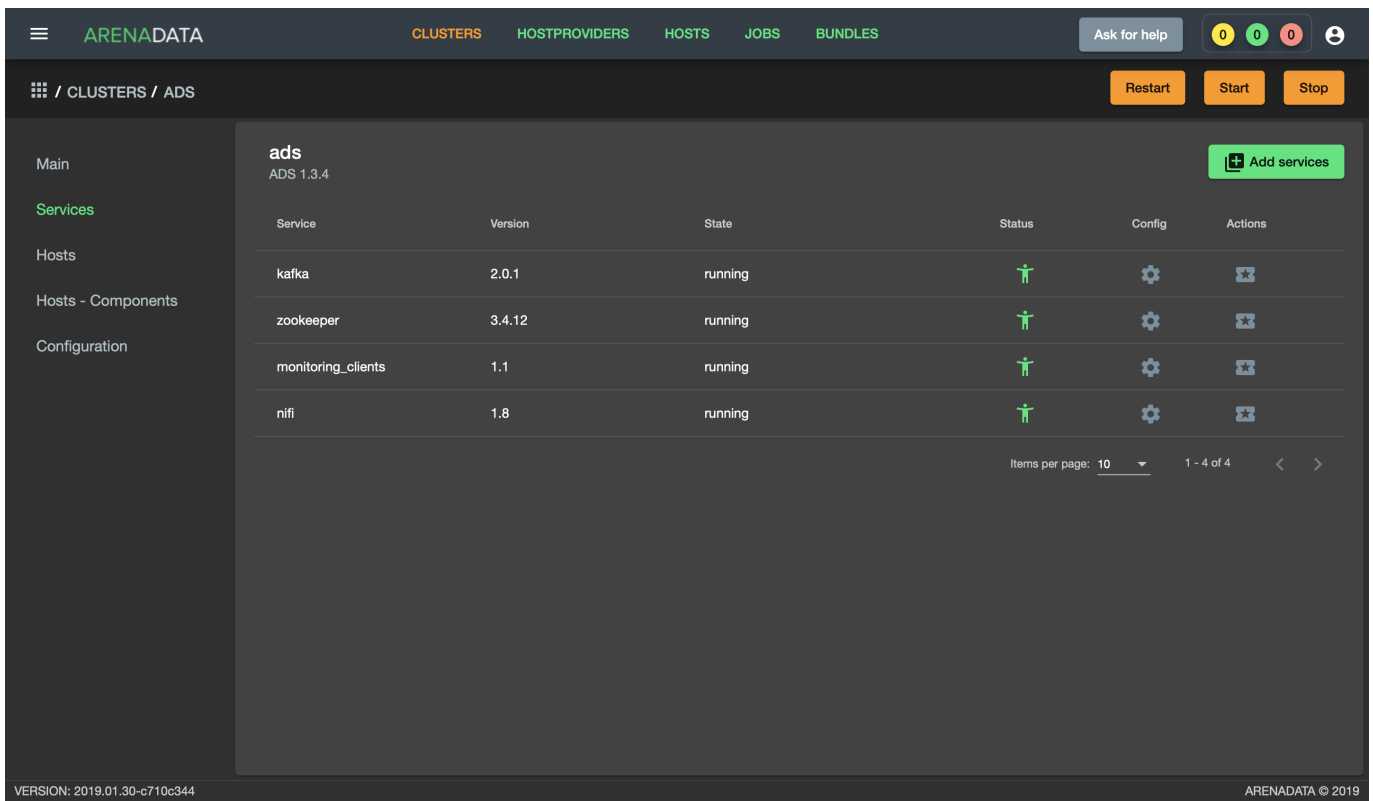
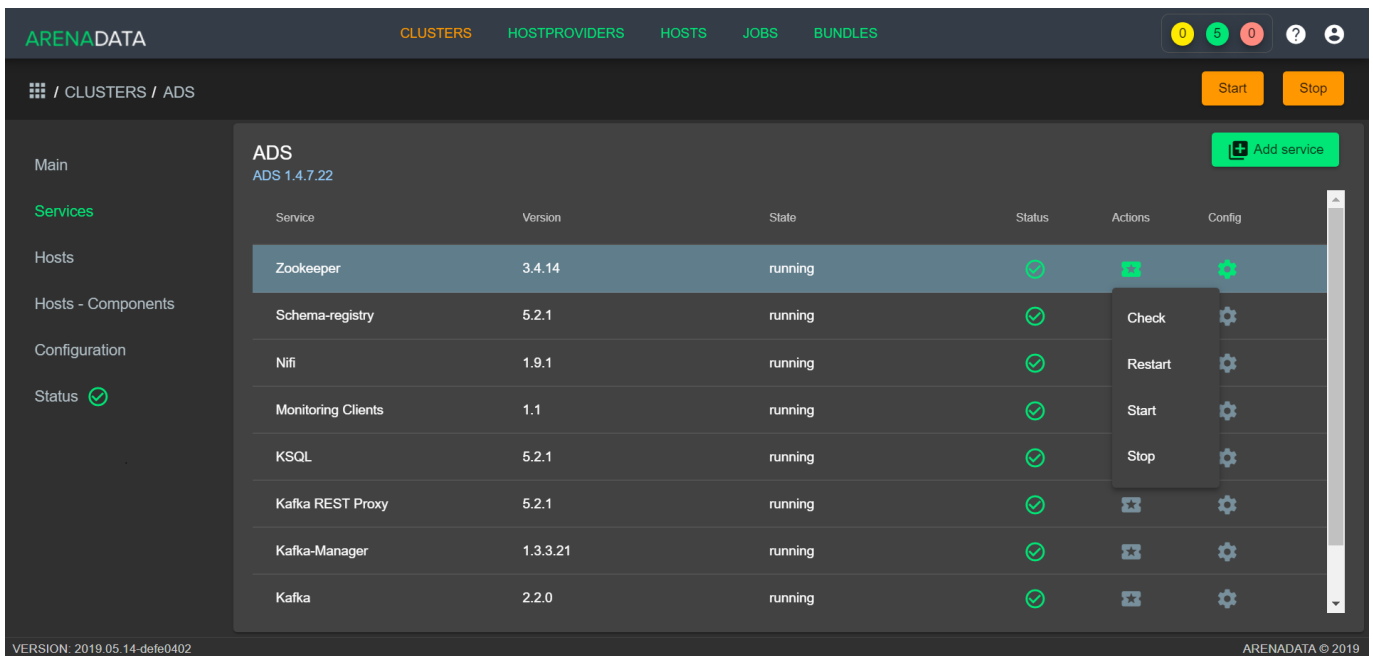


Рис.5.1.: Операции на уровне кластера

Рис.5.2.: Запуск проверки состояния сервиса *Zookeeper*

The screenshot shows the ARENADATA interface with the 'JOBS' tab selected. The table displays the following data:

#	Action	Objects	Start date	Finish Date	Status
27	CHECK	ADS	21 мая 2019 г., 16:44:16	21 мая 2019 г., 16:44:36	success
26	CHECK	ADS	21 мая 2019 г., 16:35:27	21 мая 2019 г., 16:36:02	success
25	RESTART	ADS	21 мая 2019 г., 15:38:37	21 мая 2019 г., 15:40:15	success
24	RESTART	ADS	21 мая 2019 г., 15:38:33	21 мая 2019 г., 15:39:33	success
23	RESTART	ADS	21 мая 2019 г., 15:38:28	21 мая 2019 г., 15:39:35	success
22	RESTART	ADS	21 мая 2019 г., 15:38:24	21 мая 2019 г., 15:39:38	success
21	RESTART	ADS	21 мая 2019 г., 15:38:19	21 мая 2019 г., 15:40:25	success
19	RESTART	ADS	21 мая 2019 г., 15:38:11	21 мая 2019 г., 15:39:15	success
18	RESTART	ADS	21 мая 2019 г., 15:38:01	21 мая 2019 г., 15:38:44	success

At the bottom of the table, it indicates 'Items per page: 10' and '1 - 10 of 27'. The footer shows 'VERSION: 2019.05.14-defe0402' and 'ARENADATA © 2019'.

Рис.5.3.: Вкладка “JOBS”

The screenshot shows the ARENADATA interface with the 'CHECK' log for a job. The log content is as follows:

```

Check for "ads" [ cluster ]
zookeeper 3.4.12

[ Success ] -- Znode zk_service_check availability check
Zookeeper service check is Passed
    
```

There is a 'Refresh log' button in the top right corner of the log area. The footer shows 'VERSION: 2019.01.30-c710c344' and 'ARENADATA © 2019'.

Рис.5.4.: Проверка состояния сервиса *Zookeeper*

Например, чтобы перезапустить все компоненты сервиса *Kafka*, необходимо в **ADCM** перейти в кластер *ADS*, на вкладке “Services” для сервиса *Kafka* в поле “Actions” нажать на пиктограмму и выбрать действие *Restart*.

В результате того или иного действия статус сервиса меняется на *running* или *stopped*, в том случае если сервис уже проинсталлирован и не находится в состоянии *created*.

Important: Контактная информация службы поддержки – e-mail: info@arenadata.io
