# ARENADATA

# $\mathbf{Arenadata}^{\mathsf{TM}} \mathbf{Streaming}$

Версия - v1.6-RUS

Инструкция по установке

# Оглавление

| 1 | Загрузка дистрибутива платформы               |  |  |  |  |  |  |  |  |  |  |  |  |
|---|---|--|--|--|--|--|--|--|--|--|--|--|--|
| 2 | Р. Предварительные действия                   |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Загрузка бандла ADS                           |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Создание кластера АDS                         |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.1 Создание экземпляра кластера              |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.2 Конфигурация кластера                     |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.3 Добавление хостов                         |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.4 Добавление сервисов                       |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.5 Размещение компонентов сервисов на хостах |  |  |  |  |  |  |  |  |  |  |  |  |
|   | 4.6 Установка кластера                        |  |  |  |  |  |  |  |  |  |  |  |  |

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 несколькоми сервисами через единый интерфейс;
  - Возможность управления несколькими кластерами через единый интерфейс.
- 3. Возможность использования как облачной, так и физической инфраструктуры;
- 4. Все грт-пакеты для ADS берутся из репозитория Arenadata:
  - Все грт протестированы;
  - Используется фиксированный набор пакетов в помощь пользователям и для составления баг-репортов.

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

# 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: |  |  |   |
|                     |  |  |   |
|                     |  |  |   |
|                     |  |  | 1 |
|                     |  |  |   |
| 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 (Puc.1.2.).

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

Спасибо! Вы можете теперь закрыть окно и продолжить.

Рис.1.2.: Заявка принята

| ARENADATA  | ſ  | Тродукты    | Документация                       | Поддержка       | Утилиты       | Пользователь  | #  |
|--|--|-------------|------------------------------------|-----------------|---------------|---|----|
| Магазин Приложени<br>В данном разделе вы можете загрузи<br>продуктам.                                | <b>Й</b><br>ть программное обеспечение ком               | ипании Aren | adata, а так же по                 | лучить электроі | ную версин    | о документации п  | 10 |
| Управление сервисами:  |  |             |                                    |                 |               |   |    |
| Ф Arenadata Cluster<br>Manager<br>Платформа управления,<br>развертывания и мониторинга<br>кластеров. | Cepвис для автоматического<br>развертывания стека Hadoop |             |                                    |                 |               |   |    |
| Документация   | Документация   |             |                                    |                 |               |   |    |
| Хранение и обработка д<br>Ф Arenadata Hadoop   | анных:<br>Ф Arenadata DB                                 |             | S Arenadata<br>Streaming           |                 | •             | Arenadata<br>QuickMarts                                     |    |
| Набор сервисов Наdoop стека  | MPP-система на базе<br>Greenplum DB                      |             | Платформа для обр<br>потоковых дан | работки<br>ных  | Кла<br>систея | стерная колоночная<br>иа управления базам<br>данных (СУБД). | NN |
| Документация   | Документация   |             | Документац                         | ия              |               | Документация  |    |



×

Для загрузки дистрибутива платформы следует перейти на страницу продукта, кликнув на продукт Arenadata Streaming. При этом происходит переход на соответствующую страницу с кратким описанием продукта, ссылками на компоненты дистрибутива и информацией о релизе (Puc.1.4.).

| ARENADATA  | Продукты      | Документация                       | Поддержка     | Утилиты     | Пользователь     | #       |
|--|---------------|------------------------------------|---------------|-------------|------------------|---------|
|  |               |                                    |               |             |                  |         |
| S Arenadata Streaming  |               |                                    |               |             |                  |         |
| Arenadata Streaming включает необходимый набор компоненто<br>как: Apache Kafka, Apache NiFi, Apache NiFi Registry. | ов для обрабо | отки и передачи по                 | токовых даннь | іх. Решение | включает такие п | родукты |
| Выберите версию:   | I             | Выберите ОС:                       |               |             |                  |         |
| ADS 1.5.0  | T             | CentOS/RHEL 7 x8                   | 36_64         |             |                  | T       |
| Ссылки на компоненты дистрибутива:   |               | Информация о рели                  | 136:          |             |                  |         |
| Arenadata Streaming Platform   | +             | <ul> <li>Дата релиза: 2</li> </ul> | 27.04.2020    |             |                  |         |
| • Arenada Streaming Bundle for ADCM  |               | • Тип релиза: М                    | ajor Release  |             |                  |         |
| <i>🗐</i> Документация по продукту  | +             | • минорная вер                     | СИЯ: 1.5.0.1  |             |                  |         |
|  |               |                                    |               |             |                  |         |

Рис.1.4.: Страница загрузки Arenadata Streaming

В разделе "Выберите версию продукта" следует указать интересующую версию **ADS**, а в поле "Выберите *OC*" – используемую операционную систему. При этом в разделе "Информация о peлuse" 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.

# Предварительные действия

Для установки ADS посредством ADCM необходимо:

- 1. Установить АDCM (документация продукта приведена по ссылке).
- 2. Проверить настройки ADCM:
  - Перейти во вкладку "Settings";
  - Проверить на корректность определившийся автоматически URL и при необходимости заменить его на исправный (Рис.2.1.).

| ARENADATA                      | CLUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES           | • • • • • •      |
|--------------------------------|---|------------------|
| <b>III</b> /                   |   |                  |
| Intro                          |   |                  |
| Settings                       | Description configuration Search field Q X Advanced | Save 🕠           |
| Users                          | Global Options                                      | ^                |
|                                | Send Anonymous Statistics: 🔽                        | (j) C            |
|                                | ADCM's URL: http://10.92.6.55:8000                  | () C             |
|                                | Google Auth   | ^                |
|                                | client_id:  | C                |
|                                | secret:   | C -              |
| VERSION: 2020.03.11.18-0758b87 |   | ARENADATA © 2020 |

Рис.2.1.: Проверка настроек ADCM

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.
- 5. (Опционально) Для установки без доступа к сети Интернет (offline):
  - Запросить дистрибутив Arenadata Enterprise Tools;
  - Развернуть Arenadata Enterprise Tools на выделенном хосте.

# 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.).

| ARENADATA                       |         | HOSTPROVIDERS |  | BUNDLES                         | o o o ? 🔒        |
|---------------------------------|---------|---------------|--|---------------------------------|------------------|
| III / BUNDLES                   |         |               |  |                                 | Upload bundles   |
| Name                            | Version |               |  | Description                     |                  |
| Yandex Compute VM Provider      | 2.5-1   |               |  | Manage Yandex Compute Cloud VMs | Î                |
| Monitoring                      | 2.10    |               |  | Monitoring and Control Software | Î                |
|                                 |         |               |  | tterns per page: 10             |                  |
|                                 |         |               |  |                                 |                  |
|                                 |         |               |  |                                 |                  |
|                                 |         |               |  |                                 |                  |
|                                 |         |               |  |                                 |                  |
|                                 |         |               |  |                                 |                  |
|                                 |         |               |  |                                 |                  |
| VERSION: 2019 10 08 16-6450aa80 |         |               |  |                                 | ΔRΕΝΔΠΔΤΔ @ 2019 |

#### Рис.3.1.: Вкладка "BUNDLES"

- 2. Нажать "Upload bundle" и в открывшейся форме выбрать файл бандла ADS (Рис.3.2.).
- 3. После успешной загрузки бандла необходимо принять лицензионное соглашение (Рис.3.3.).
- 4. В результате выполненных действий факт успешной загрузки отображается в общем списке бандлов на вкладке "BUNDLES" (Рис.3.4.).

| ARENADATA    |      |             |                                |  |                               |                             |                                 | 0 0 0 | 0       | θ  |
|--------------|------|-------------|--------------------------------|--|-------------------------------|-----------------------------|---------------------------------|-------|---------|----|
| <b>III</b> / |      |             |                                |  |                               |                             |                                 | Uploa | l bundl | es |
|              |      | Отмен       |                                | Открытие файлов<br>Ф denis Документы bundles I | <u>م</u>                      | Открыть                     |                                 |       |         |    |
|              | 2.10 | ф д         | Домашняя папка<br>Рабочий стол | Имя<br>adcm_cluster_ads_v1.5.0.1_ce.tgz        | Размер<br>901.1 кБ<br>83.3 кБ | Изменён 🔻<br>12:47<br>2 апр | Monitoring and Control Software |       |         |    |
|              |      | •∎ В<br>С Д | Видео<br>Документы             |  | 05.5 10                       | z unp.                      |                                 |       |         |    |
|              |      | ⇒ 3<br>⊡ /  |                                |  |                               |                             |                                 |       |         |    |
|              |      |             | Музыка<br>Другие места         |  |                               |                             |                                 |       |         |    |
|              |      | Ì           | 7                              |  | Пользователь                  | ские файлы 🔻                |                                 |       |         |    |
|              |      |             |                                |  |                               |                             |                                 |       |         |    |

Рис.3.2.: Выбор бандла

| ARENADATA                | CLUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES  | 0 0 0    | ?        | θ      |
|--------------------------|--|----------|----------|--------|
| III /                    | Accept license agreement   | Uploa را | d bundle | es     |
|                          | LICENSE AGREEMENT  | 4        |          |        |
|                          | A REFERENCE THE TERMS AND REFERENCESSOR IN THE ARRESPENT ARE RECEIPTED TH CERTAIN A  |          |          |        |
| Yandex Compute VN        | <ol> <li>DEFINITIONS. THE TERMS AND DEFINITIONS USED IN THE ROKEEMENT ARE DESCRIBED IN SECTION 1.</li> <li>1.1. "Confidential Information" shall mean all disclosed information with regard to this Agreement or the Product<br/>whether furnished in oral, written or other tangible or intangible form. The Confidential Information shall include, but</li> </ol>   |          |          |        |
|                          | not limited to, components of business plans, products, inventions, design plans, financial plans, customer related<br>information, strategies and other information of similar nature.<br>1.2. "Documentation" shall mean all user manuals and administrator quides as well as other technical documents  |          |          |        |
|                          | 1.3. "Product" shall mean any version of software for computers and data bases, including, but not limited to, the computer software Arenadata Streaming Platform, that are built on an open SW basis managed by the Apache Software Foundation licensed under the terms and conditions of the Apache 2.0 License (http://www.apache.org/licenses/LICENSE-   | 3 of 3   |          |        |
|                          | 2.0).<br>1.4. "Intellectual Property" shall mean all systems, methods, algorithms, structures, libraries, applications<br>(supplementary software), components/parts of the Product (including texts and fonts), all modules, other elements of th<br>Product copied and/or incorporated in all working software, any copies, documentation, authorship, logos and other<br>information included in the Product.<br>1.5. Right Owner - Arenadata Software LLC.   | e        |          |        |
|                          | 2. INTELLECTUAL PROPERTY EXCLUSIVE RIGHTS<br>2.1. The Product, systems, methods, algorithms, structures, libraries, applications (supplementary software),<br>components/parts of the Product (including texts and fonts), all modules, other elements of the Product copied and/or<br>incorporated in all working software, any copies, documentation, authorship, logos and other information included in the<br>Product shall be objects of intellectual property and commercial secret of the Right Owner, i.e. Arenadata Software LLC<br>and/or its affiliates and shall be protected according to the Russian Federation effective legislation on intellectual<br>property, commercial secret as well as the provisions of this Agreement. | Ť        |          |        |
|                          | Do you accept the license agreement? Yes No  |          |          |        |
|                          |  |          |          |        |
| VERSION: 2020.03.11.18-0 | 7758687  | ARE      | NADATA ( | © 2020 |

Рис.3.3.: Лицензионное соглашение

| ARENADATA                      |       | CLUSTERS | HOSTPROVIDERS | HOSTS     | JOBS | BUNDLES |                                 | 0 0 0 | ?      | θ      |
|--------------------------------|-------|----------|---------------|-----------|------|---------|---------------------------------|-------|--------|--------|
| <b>III</b> /                   |       |          |               |           |      |         |                                 | Uploa | d bunc | lles   |
|                                |       |          |               |           |      |         |                                 |       |        |        |
| Monitoring                     | 2.10  |          |               | communit  |      |         | Monitoring and Control Software |       | Î      |        |
| Yandex Compute VM Provider     | 2.8-1 |          |               | community |      |         | Manage Yandex Compute Cloud VMs |       | Î      |        |
|                                |       |          |               | communit  |      |         | Arenadata Streaming Software.   |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
|                                |       |          |               |           |      |         |                                 |       |        |        |
| VERSION: 2020.03.11.18-0758b87 |       |          |               |           |      |         |                                 | AREN  | ADATA  | © 2020 |

Рис.3.4.: Результат успешной загрузки бандла

# 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-ключах или паролях.

| ARENADATA                       |          | CLUSTERS | HOSTPROVIDERS     | HO | STS     | JOBS | BUNDLES |           |                    | 0      | 0 0        | ) ?      | 8        |
|---------------------------------|----------|----------|-------------------|----|---------|------|---------|-----------|--------------------|--------|------------|----------|----------|
| III / HOSTS                     |          |          |                   |    |         |      |         |           |                    |        |            | Create I | nost     |
| Fqdn                            | Provider |          | Cluster           |    | State   |      |         | Status    | Actions            | Config |            |          |          |
| monitoring                      | ssh      |          | Assign to cluster |    | running |      |         | $\oslash$ | 22                 | \$     | Î          |          |          |
| nifi                            | ssh      |          | Assign to cluster |    | running |      |         | $\otimes$ | 52                 | \$     | Î          |          |          |
| kafka1                          | ssh      |          | Assign to cluster |    | running |      |         | $\oslash$ | 52                 | \$     | Î          |          |          |
| kafka2                          | ssh      |          | Assign to cluster |    | running |      |         | $\oslash$ | <b>5</b>           | \$     | Î          |          |          |
| zk                              | ssh      |          | Assign to cluster |    | running |      |         | $\oslash$ | 52                 | \$     | Î          |          |          |
|                                 |          |          |                   |    |         |      |         |           | Items per page: 10 |        | 1 - 5 of 5 |          |          |
|                                 |          |          |                   |    |         |      |         |           |                    |        |            |          |          |
|                                 |          |          |                   |    |         |      |         |           |                    |        |            |          |          |
| 12/ERS1/2002/019.05.14-defe0402 |          |          |                   |    |         |      |         |           |                    |        | ARE        | NADAT    | A © 2019 |

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.).

| ARENADATA                |                | CLUSTERS    | HOSTPROVIDERS | HOSTS JOB | S BUNDLI |        |                    | 0      | 0 0        | ?         | 8      |
|--------------------------|----------------|-------------|---------------|-----------|----------|--------|--------------------|--------|------------|-----------|--------|
| III / CLUSTERS           |                |             |               |           |          |        |                    |        | Crea       | te cluste | er     |
| Name                     | Bundle         | Description | State         | Status    | Actions  | Import | Upgrade            | Config |            |           |        |
| Monitoring               | Monitoring 3.0 |             | running       | $\otimes$ | 53       |        |                    | \$     | Î          |           |        |
|                          |                |             |               |           |          |        | Items per page: 10 |        | 1 - 1 of 1 |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
|                          |                |             |               |           |          |        |                    |        |            |           |        |
| VERSION: 2019.05.14-defe | 0402           |             |               |           |          |        |                    |        | AREN       | ADATA @   | © 2019 |

Рис.4.2.: Вкладка "CLUSTERS"

- 2. Нажать "Create 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.).

### 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 *RedHat repos* указываются требуемые для установки *ADS* пакеты из различных уит-репозиториев. В блоке *AltLinux repos* указываются требуемые для установки *ADS* пакеты из различных apt-репозиториев. При этом в каждом из параметров можно изменить заданный по умолчанию url:

- ADS;
- monitoring;

| ARENADATA                      |                          |                       |    |           |           |    | 0 | • •    | ?       | θ      |
|--------------------------------|--------------------------|-----------------------|----|-----------|-----------|----|---|--------|---------|--------|
| <b>III</b> /                   |                          |                       |    |           |           |    |   | 🛨 Crea | te clus | ter    |
|                                | Bundle                   | cluster               |    |           |           |    |   |        |         |        |
| Monitoring                     | Monitoring 2.10 c<br>Bui |                       |    |           | $\otimes$ | 53 |   | \$     |         |        |
|                                |                          | S - 1.5.0.1 community |    | <u> </u>  |           |    |   |        |         |        |
|                                | Clu                      | ster name *<br>IS     |    |           |           |    |   |        |         |        |
|                                |                          |                       |    |           |           |    |   |        |         |        |
|                                | De<br>—                  | scription             |    |           |           |    |   |        |         |        |
|                                |                          |                       | Sa | ve Cancel |           |    |   |        |         |        |
|                                |                          |                       |    |           |           |    |   |        |         |        |
|                                |                          |                       |    |           |           |    |   |        |         |        |
| VERSION: 2020.03.11.18-0758b87 |                          |                       |    |           |           |    |   | AREN   | ADATA   | © 2020 |

Рис.4.3.: Создание экземпляра кластера

| ARENADATA                      | CLUSTERS                | HOSTPROVIDERS | HOSTS | JOBS | BUNDLES |              |                         | 0       | 0 0    | ?         | 8    |
|--------------------------------|-------------------------|---------------|-------|------|---------|--------------|-------------------------|---------|--------|-----------|------|
| <b>III</b> /                   |                         |               |       |      |         |              |                         |         | 🛨 Crea | te cluste | r    |
| Name 🕇                         | Bundle                  | Descriptio    |       |      |         |              | tions Import            | Upgrade | Config |           |      |
| ADS                            | ADS 1.5.0.1 community   |               |       |      | created |              | 🗙 t <sub>t</sub>        |         | \$     |           |      |
| Monitoring                     | Monitoring 2.10 communi | ty            |       |      | running | $\bigotimes$ | <b>₹</b> † <sub>↓</sub> |         | \$     |           |      |
|                                |                         |               |       |      |         | ltems pe     | er page: <u>10</u>      |         | 2 of 2 |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
|                                |                         |               |       |      |         |              |                         |         |        |           |      |
| VERSION: 2020.03.11.18-0758b87 |                         |               |       |      |         |              |                         |         | AREN   | ADATA ©   | 2020 |

Рис.4.4.: Результат успешного создания экземпляра кластера

| ARENADATA                      |                           | CLUSTERS | HOSTPROVIDERS       | HOSTS          | JOBS          | BUNDLES                   |          | 0 0 | 0 O     |
|--------------------------------|---------------------------|----------|---------------------|----------------|---------------|---------------------------|----------|-----|---------|
| III / CLUSTERS / ADS           |                           |          |                     |                |               |                           |          |     | Install |
| Main                           | ADS<br>ADS 1.5.0.1        |          |                     |                |               |                           |          |     |         |
| Services                       | Description configuration |          | Search fiel         |                |               |                           | Advanced | Sav | • •     |
| Hosts - Components             | RedHat repos              |          |                     |                |               |                           |          |     |         |
|                                | ADS repo url:             |          | http://downloads.ar | enadata.io/AD  | S/1.5.0/cen   | itos/7/community/x86_64   |          | (j  |         |
| Status 🕕                       | Monitoring repo url:      |          | http://downloads.ar | enadata.io/ADI | M/2.2/cent    | os/7/community/x86_64     |          |     | C       |
| Import                         | Zookeeper repo url:       |          | http://downloads.ar | enadata.io/zoc | okeeper/3.5   | .6/centos/7/community/x8  | 6_64     |     |         |
|                                | Altlinux repos            |          |                     |                |               |                           |          |     | ^       |
|                                | ADS repo url:             |          | http://downloads.ar | enadata.io/AD  | S/1.5.0/altli | inux/8.0/community        |          | 0   |         |
|                                | Monitoring repo url:      |          | http://downloads.ar | enadata.io/ADI | M/2.2/altlin  | ux/8.0/community          |          |     | G       |
|                                | Zookeeper repo url:       |          | http://downloads.ar | enadata.io/zoc | okeeper/3.5   | .6/altlinux/8.0/community |          | Ū   |         |
|                                |                           |          |                     |                |               |                           |          |     |         |
| VERSION: 2020 03 11 18-0758687 |                           |          |                     |                |               |                           |          | ADE |         |

#### Рис.4.5.: Окно конфигурации кластера

- Zookeeper;
- Ranger plugins. Только для Enterprise версии.

Important: При установке в окружении без доступа к сети Интернет (offline) необходимо предварительно развернуть кластер Arenadata Enterprise Tools и импортировать настройки HTTP Mirorr

В таком случае настройки offline-репозиториев применяются автоматически. Для этого на странице кластера *ADS* необходимо перейти на вкладку *Import* и выбрать сервис http кластера Enterprise Tools (Puc.4.6.). Альтернативно, в эти поля можно указать url локальных репозиториев с необходимыми пакетами.

Important: Arenadata Enterprise Tools предоставляет локальную копию только репозиториев Arenadata Streaming и Arenadata Monitoring. Доступ к репозиториям CentOS должен быть организован отдельно

### 4.3 Добавление хостов

Для добавления хостов в кластер ADS необходимо:

- 1. В меню кластера ADS открыть вкладку "Hosts" (Рис.4.7.).
- 2. Нажать "Add hosts" и в открывшейся форме выбрать необходимые хосты (Рис.4.8.).

**Important:** Необходимо использовать иерархические (FQDN) имена хостов. В противном случае возможны ошибки при установкеэксплуатации кластера.

| ARENADATA                       | CLUSTERS                | HOSTPROVIDERS | HOSTS JOB       | S BUNDLES | 0                          | • • ?    | 8        |
|---------------------------------|-------------------------|---------------|-----------------|-----------|----------------------------|----------|----------|
| III / CLUSTERS / ADS            |                         |               |                 |           |                            |          | Ç!5      |
| Main                            | ADS<br>ADS 1.4.10.11    |               |                 |           |                            | Save     |          |
| Services                        | graphite                | grafana       |                 | zookeeper | http                       |          |          |
| Hosts                           | Etools cluster/Graphite | Etools        | cluster/Grafana |           | Etools cluster/HTTP Mirror |          |          |
| Hosts - Components              | Graphite 1.1.5          | Grafana 5     | .4.3            |           | HTTP Mirror 2019071709     |          |          |
| Configuration                   |                         |               |                 |           |                            |          |          |
| Status 🚺                        |                         |               |                 |           |                            |          |          |
|                                 |                         |               |                 |           |                            |          |          |
|                                 |                         |               |                 |           |                            |          |          |
|                                 |                         |               |                 |           |                            |          |          |
|                                 |                         |               |                 |           |                            |          |          |
|                                 |                         |               |                 |           |                            |          |          |
| VERSION: 2019.10.08.16-6450aa80 |                         |               |                 |           |                            | ARENADAT | A © 2019 |





Рис.4.7.: Вкладка "Hosts" кластера ADS

| ARENADATA            | CLUSTERS      | HOSTPROVIDERS H | OSTS JOBS | BUNDLES | o o o o o |
|----------------------|---------------|-----------------|-----------|---------|-----------|
| III / CLUSTERS / ADS |               |                 |           |         | Install   |
| Main ADS<br>ADS 1.5  | <b>5</b> .0.1 |                 |           |         | Add hosts |
|                      |               |                 |           |         |           |
|                      | Add 1         | ree host        |           |         |           |
|                      |               |                 |           | •       |           |
| Configuration        | nifi          |                 |           |         |           |
| Status (!)           | kaf           | ka1             |           |         |           |
|                      | kaf           | ka2             |           |         |           |
|                      | zk            |                 |           |         |           |
|                      |               |                 |           |         |           |
|                      |               |                 |           |         |           |
|                      |               |                 |           |         |           |
|                      |               |                 |           |         |           |
|                      |               |                 |           |         |           |



3. В результате выполненных действий факт добавления хостов отображается в кластере *ADS* в списке вкладки "Hosts" (Рис. 4.9.).

### 4.4 Добавление сервисов

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

Кластер **ADS** содержит следующие сервисы:

| Import   |         |
|--|---------|
| Main       ADS<br>ADS 1.5.0.1         Services       FODN       Provider       State       State       Actions       Config         Hosts       Infi       Ssh       running       O       I       I       O       I       O       I       I       O       I       I       O       I       I <td>Install</td>  | Install |
| HostsPowlerStateStateStateActionsConfigHosts - ComponentsnifisshrunningImage: StateImage: StateIma | osts    |
| nifi     ssh     running     or     ss       Hosts - Components     kafka1     ssh     running     or     ss       Configuration     kafka2     ssh     running     or     ss       Status ()     zk     ssh     running     or     ss     or  |         |
| kafka1     ssh     running     Import  |         |
| kafka2     ssh     running     Import       Import     Items per page: 10     1 - 4 of 4   |         |
| Import   |         |
| Import   |         |
|  |         |



| 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  | Распрделенная платформа, предназначенная для построения и     |  |  |  |  |  |
|                    | Registry           | 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 версионную 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  | Сервис представляет собой потоковый SQL-движок с открытым     |  |  |  |  |  |
|                    | Client             | 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                        |  |  |  |  |  |
| MiNifi             | MiNiFi C2 Server,  | Агент для сбора и отсылки данных, который пополняет основные  |  |  |  |  |  |
|                    | MiNiFi Agent       | принципы NiFi в управлении потоками данных, уделяя особое     |  |  |  |  |  |
|                    |                    | внимание сбору данных у их источника                          |  |  |  |  |  |
| Monitoring Clients | System metrics     | Агенты, отсылающие информацию о хосте и сервисах в            |  |  |  |  |  |
|                    | Agent, JMX metrics | мониторинг  |  |  |  |  |  |
|                    | Collector          | 17  |  |  |  |  |  |

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 – *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.4.10.).

| ARENADATA   | CL                           | LUSTERS HOSTPROVIDERS HOSTS JOBS BUNDLES  | 3 0 0 9 🖯   |
|---|------------------------------|---|---|
| III / CLUSTERS / ADS  |                              |   | Install   |
| Main<br>Services<br>Hosts<br>Hosts - Components<br>Configuration<br>Status ()<br>Import | ADS<br>ADS 1.5.0.1<br>Name ↑ | Add service         Kafka - 2.4.0         MiNiFi - 0.5.0         Monitoring Clients - 1.3.0         NiFi - 1.10.0 | E Add service<br>State Status Import Actions Config<br>tems per page: 10 		 0 of 0 	< > |
|   |                              | Schema-Registry - 5.4.0   |   |
| VERSION: 2020.03.11.18-0758b87  |                              |   | ARENADATA © 2020  |

Рис.4.10.: Выбор сервисов для кластера ADS

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

#### 4.4.1 Особенности сервиса Monitoring Clients

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

Для импорта конфигурации мониторинга необходимо в кластере *ADS* выбрать сервис *Monitoring Clients*, затем перейти во вкладку *Import* и отметить импортируемые настройки сервисов (Puc. 4.11.).

| III / CLUSTERS / LONGEST DESNA ! / SERVICES / MONITORING CLIENTS ! |                          |                               |  |  |  |  |  |
|--|--------------------------|-------------------------------|--|--|--|--|--|
| Main   | Monitoring Clients       |                               |  |  |  |  |  |
| Configuration  | graphite                 | grafana                       |  |  |  |  |  |
| Status 🧭   | This import is required! | This import is required!      |  |  |  |  |  |
| Import !   | Graphite 1.1.5           | Mad Oka/Grafana Grafana 5.4.3 |  |  |  |  |  |
| Actions  |                          |                               |  |  |  |  |  |
|  |                          |                               |  |  |  |  |  |
|  |                          |                               |  |  |  |  |  |

Рис.4.11.: Импорт конфигурации мониторинга

### 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" (Puc.4.12.).

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.4.13. 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.

С версии **ADS 1.6.0.0** есть поддержка Service Dependency Tree. Данный функционал не позволяет пользователю забыть о зависимых сервисах. Так, для работы сервиса **Schema-Registry** необходим обязательный сервис **Kafka** (Puc.4.14.).

### 4.6 Установка кластера

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

| ARENADATA  |  | s BUNDLES 0 0 ? S |     |
|--|--|-------------------|-----|
| III / CLUSTERS / ADS   |  |                   |     |
| Main Asservices<br>Hosts<br>Hosts - Components !<br>Configuration<br>Status ()<br>Import | ADS<br>ADS 1.5.0.1<br>*KSQL Server 0/1<br>KSQL Client 0<br>*System metrics Agent 0/1<br>JMX metrics Collector 0/1<br>*NiFi Server 0/1<br>NiFi Registry 0/1<br>*Schema-Registry 0/1 | Save Restore      |     |
|  | * Zookeeper Server 0/1   |                   |     |
| VERSION: 2020.03.11.18-0758b87   |  | ARENADATA © 20    | 020 |

Рис.4.12.: Размещение компонентов сервисов на хостах

| ARENADATA                      | CLUSTERS HOSTPROVIDERS HOS        | ts Jobs But | NDLES  | • • • • • • •    |
|--------------------------------|-----------------------------------|-------------|--------|------------------|
| III / CLUSTERS / ADS           |                                   |             |        |                  |
| Main<br>Services               | ADS<br>ADS 1.5.0.1<br>KSQL Server | 0 / 1       |        | Save Restore     |
| Hosts                          | KSQL Client                       | 0           | nifi   | 1                |
| Hosts - Components             | * System metrics Agent            | 0 / 1       | kafka1 |                  |
| Status 🚺                       | JMX metrics Collector             | 0 / 1       | kafka2 | •                |
| Import                         | NiFi Server                       | 1/1         | zk     |                  |
|                                | NiFi Registry                     | 0 / 1       |        |                  |
|                                | * Schema-Registry                 | 0 / 1       |        |                  |
|                                | * Zookeeper Server                | 0 / 1       |        |                  |
| VERSION: 2020.03.11.18-0758b87 |                                   |             |        | ARENADATA © 2020 |

Рис.4.13.: Компоненты сервиса Nifi

| III / CLUSTERS / LONGEST D | esna ! 🕼      |   |                |                    |           | Run action () |
|----------------------------|---------------|---|----------------|--------------------|-----------|---------------|
|                            | Longest Desna |   |                |                    |           | Restore Save  |
|                            | Componente    |   |                | lleste i           |           |               |
|                            | components    |   |                |                    |           |               |
|                            | * Schema-     | Registry  | 0 / 1          |                    | demo-host | 0             |
|                            |               |   |                |                    |           |               |
|                            |               | This component cannot be installed with                             | out the follov | ving dependencies. |           |               |
|                            |               |   |                |                    |           |               |
|                            |               | <ul> <li>Kafka         <ul> <li>Kafka Broker</li> </ul> </li> </ul> |                |                    |           |               |
|                            |               |   | _              |                    |           |               |
|                            |               |   | It is c        | lear Install All   |           |               |
|                            |               |   |                |                    |           |               |
|                            |               |   |                |                    |           |               |

Рис.4.14.: Управление плагином

На данном этапе все хосты подготовлены и можно приступать 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.).

| ARENADATA                               |                    | CLUSTERS | HOSTPROVIDERS | HOSTS    | JOBS | BUNDLES |  |         |              |              | 0 0 | • •     | 9       |
|---|--------------------|----------|---------------|----------|------|---------|--|---------|--------------|--------------|-----|---------|---------|
| III / CLUSTERS / ADS                    |                    |          |               |          |      |         |  |         |              |              |     |         | Install |
| Main                                    | ADS<br>ADS 1.5.0.1 |          |               |          |      |         |  |         |              |              | Ľ   | Add sei | vice    |
|   |                    |          |               |          |      |         |  |         |              |              |     |         | Â       |
| Hosts                                   | KSQL               |          |               | 5.4.0    |      |         |  | created | $\otimes$    |              | 53  | \$      |         |
| Hosts - Components                      | Kafka              |          |               | 2.4.0    |      |         |  | created | $\bigotimes$ |              | 23  | \$      |         |
| Configuration Status ()                 | Kafka REST Proxy   |          |               | 5.4.0    |      |         |  | created | $\otimes$    |              | 53  | \$      |         |
|   | Kafka-Manager      |          |               | 1.3.3.23 |      |         |  | created | $\otimes$    |              | 53  | \$      |         |
| Import                                  | MiNiFi             |          |               | 0.5.0    |      |         |  | created | $\otimes$    |              | 52  | \$      |         |
|   | Monitoring Clients |          |               | 1.3.0    |      |         |  | created | $\otimes$    |              | 22  | \$      |         |
|   | NiFi               |          |               | 1.10.0   |      |         |  | created | $\otimes$    |              | 23  | \$      |         |
|   | Schema-Registry    |          |               | 5.4.0    |      |         |  | created | $\oslash$    |              | 53  | \$      |         |
|   | Zookeeper          |          |               | 3.5.6    |      |         |  | created | $\otimes$    |              | 22  | ۵       |         |
|   |                    |          |               |          |      |         |  |         |              | <b>-</b> 1 · |     |         | > •     |
| VERSION: 2020.03.11.18-0758b97 ARENADAT |                    |          |               |          |      |         |  |         | TA © 2020    |              |     |         |         |

Рис.4.15.: Установка всех сервисов кластера

В появившемся диалоговом окне предоставляется выбор опций (Рис.4.16.):

- Disable SELinux before cluster installation отключение SELinux на добавляемых хостах. Для того, чтобы данная настройка применилась, после завершения операции Install необходимо перезагрузить хосты вручную;
- Disable Firewalld before cluster installation выключение firewalld на хостах;
- Install OpenJDK before cluster installation установка пакета java-1.8.0-openjdk на хостах;
- Set vm.swappiness to 0 for all hosts отключение swapping на хостах;
- Append hosts into /etc/hosts file before cluster installation запись добавляемых нод в /etc/hosts на хостах кластера. Данную опцию рекомендуется отключить, если настроен DNS.

| ARENA <b>DATA</b>    |                            |          |  |  |   |           | ( | 0 1 ( | 999        |
|----------------------|----------------------------|----------|--|--|---|-----------|---|-------|------------|
| III / CLUSTERS / ADS |                            |          |  |  |   |           |   |       | Install    |
| Main                 | ADS                        |          |  |  |   |           |   |       | dd service |
| Services             |                            |          |  |  |   |           |   |       |            |
| Hosts                |                            |          |  |  |   |           |   |       |            |
| Hosts - Components   | Kafka                      |          |  |  |   | created   |   |       |            |
| Configuration        | Run an action [ Install ]? |          |  |  |   |           |   |       |            |
| Status ①             | Disable SElinux:           | <b>_</b> |  |  |   | ന്ദ       |   |       |            |
| Import               |                            |          |  |  |   | 00        |   |       |            |
|                      | Disable Firewalld:         |          |  |  |   | () C      |   |       |            |
|                      | Install OpenJDK:           | <b></b>  |  |  |   | (j) C     |   |       |            |
|                      | Set vm.swappiness to 0:    |          |  |  |   | (j) C     |   |       |            |
|                      | Configure /etc/hosts:      |          |  |  |   | (i) C     |   |       |            |
|                      |                            |          |  |  | c | ancel Run |   |       |            |
|                      |                            |          |  |  |   |           |   |       |            |

Рис.4.16.: Доступные при кластерной установке настройки

По результатам установки все добавленные сервисы меняют состояние с created нa installed – установлен (Рис. 4.17.).

Для запуска кластера необходимо нажать кнопку Start.

#### 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 проинсталлиованный 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.

| ARENADATA   |                    | CLUSTERS | HOSTPROVIDERS | HOSTS    | JOBS | BUNDLES |           |              | 0 0       | • ?     | 9 8       |
|---|--------------------|----------|---------------|----------|------|---------|-----------|--------------|-----------|---------|-----------|
| III / CLUSTERS / ADS                                      |                    |          |               |          |      |         |           |              | s         | tart    | Stop      |
| Main  | ADS<br>ADS 1.5.0.1 |          |               |          |      |         |           |              | Ē         | Add ser | vice      |
|   |                    |          |               |          |      |         |           |              |           |         | Â         |
| Hosts   | KSQL               |          |               | 5.4.0    |      |         | installed | $\bigotimes$ | 22        | \$      |           |
| Hosts - Components<br>Configuration<br>Status 🥑<br>Import | Kafka              |          |               | 2.4.0    |      |         | installed | $\otimes$    | 83        | \$      |           |
|   | Kafka REST Proxy   |          |               | 5.4.0    |      |         | installed | $\otimes$    | 53        | \$      |           |
|   | Kafka-Manager      |          |               | 1.3.3.23 |      |         | installed | $\otimes$    | 52        | \$      |           |
|   | MiNiFi             |          |               | 0.5.0    |      |         | installed | $\oslash$    | <b>53</b> | \$      |           |
|   | Monitoring Clients |          |               | 1.3.0    |      |         | installed | $\otimes$    | 53        | \$      |           |
|   | NiFi               |          |               | 1.10.0   |      |         | installed | $\otimes$    | 53        | \$      |           |
|   | Schema-Registry    |          |               | 5.4.0    |      |         | installed | $\oslash$    | 22        | \$      |           |
|   | Zookeeper          |          |               | 3.5.6    |      |         | installed | $\oslash$    | 53        | \$      |           |
|   |                    |          |               |          |      |         |           |              |           |         | > •       |
| VERSION: 2020.03.11.18-0758b87                            |                    |          |               |          |      |         |           |              |           | ARENADA | TA © 2020 |

Рис.4.17.: Состояние сервисов кластера