

Инструкция по установке Arenadata Hadoop Platform T4 (ADH T4)

Содержание:

1	Online-установка.....	4
1.1	Установка ADCM.....	4
1.1.1	Шаг 1. Предварительные требования	4
1.1.2	Шаг 2. Установка ADCM	7
1.1.3	Шаг 3. Запуск ADCM	8
1.1.4	Шаг 4. Проверка установки.....	8
1.1.5	Шаг 5. Обновление ADCM	12
1.2	Подготовка хостов.....	12
1.2.1	Добавление хостпровайдера SSH	12
1.2.2	Добавление хостов	14
1.3	Установка кластера ADH T4.....	15
1.3.1	Создание кластера	15
1.3.2	Добавление сервисов.....	17
1.3.3	Добавление хостов в кластер	18
1.3.4	Добавление компонентов	18
1.3.5	Настройка сервисов	19
1.3.6	Настройка кластера	20
1.3.7	Установка кластера	21
1.4	Установка мониторинга	24
1.4.1	Создание кластера	24
1.4.2	Добавление сервисов.....	25
1.4.3	Добавление хостов в кластер	26
1.4.4	Добавление компонентов	26
1.4.5	Настройка сервисов	27
1.4.6	Установка кластера	28
1.4.7	Интеграция с кластером ADH.....	29
2	Offline-установка	32
2.1	Установка ADCM.....	32
2.1.1	Шаг 1. Предварительные требования	32
2.1.2	Шаг 2. Установка ADCM	35
2.1.3	Шаг 3. Запуск ADCM	35
2.1.4	Шаг 4. Проверка установки.....	36
2.1.5	Шаг 5. Обновление ADCM	38
2.2	Подготовка хостов.....	39
2.2.1	Добавление хостпровайдера SSH	39
2.2.2	Добавление хостов	41
2.3	Установка кластера Enterprise Tools	43
2.3.1	Создание кластера	43
2.3.2	Добавление сервисов.....	44
2.3.3	Добавление хостов в кластер	44
2.3.4	Добавление компонентов	45
2.3.5	Установка кластера	46
2.4	Установка кластера ADH T4.....	47

2.4.1	Создание кластера	47
2.4.2	Добавление сервисов	48
2.4.3	Добавление хостов в кластер	49
2.4.4	Добавление компонентов	50
2.4.5	Настройка сервисов	52
2.4.6	Настройка кластера	52
2.4.7	Импорт настроек ET	53
2.4.8	Установка кластера	54
2.5	Установка мониторинга	57
2.5.1	Настройка кластера Enterprise Tools	57
2.5.2	Интеграция с кластером ADH.....	61

1 Online-установка

Раздел посвящен online-установке Arenadata Hadoop Platform T4 (ADH T4). Данный способ установки подходит для хостов с доступом к Интернету. Процесс online-установки через Arenadata Cluster Manager (ADCM) включает следующие шаги:

- Установка ADCM (пп. 1.1);
- Подготовка хостов (пп. **Ошибка! Источник ссылки не найден.**);
- Установка кластера ADH T4 (пп. 1.3);
- Установка мониторинга (пп. 1.4).

1.1 Установка ADCM

Для развертывания ADCM требуется отдельный хост.

1.1.1 Шаг 1. Предварительные требования

1.1.1.1 Требования к software и hardware

ADCM поставляется в виде Docker-образа. Для его установки необходимы:

- Операционная система версии CentOS 7/RHEL 7/Alt Linux SP 8/Astra Linux SE 1.7 Орел. Обратите внимание, что архивные версии не поддерживаются и не тестируются.
- Пользователь с правами sudo.
- Доступ к официальным репозиториям CentOS Extras и CentOS Base. По умолчанию, эти репозитории доступны в ОС. Однако, если они были отключены, потребуются их предварительное включение (например, путем редактирования конфигурационного файла в директории /etc/yum.repos.d/).
- Менеджер программных пакетов YUM/APT.
- Хост, отвечающий следующим требованиям:
 - CPU: 4 ядра CPU;
 - RAM: 16 ГБ;
 - HDD: 50 ГБ.

Приведенные системные требования являются минимальными. Целевой сайзинг необходимо рассчитывать исходя из требований конкретной организации.

Также перед началом установки необходимо отредактировать файл `/etc/sudoers`, чтобы впоследствии пользователь, выполняющий установку ADCM, обладал необходимыми правами. Для этого выполните команду:

```
$ sudo vi /etc/sudoers
```

В этом файле необходимо удалить символ `#` в начале строки `WHEEL_USERS ALL=(ALL) ALL`. Чтобы сохранить это изменение и вернуться в терминал, нажать `Esc`, после чего набрать `:wq` либо `:wq!`.

1.1.1.2 Установка Docker

Установка Docker с официального сайта не рекомендуется. Вместо этого советуется использовать официальные репозитории дистрибутива Linux. Репозитории, предоставляемые разработчиками операционных систем, как правило, более стабильны и безопасны.

В зависимости от операционной системы и менеджера пакетов, которые используются, последовательность установки включает следующие шаги:

- YUM:

1. Обновите все программные пакеты, установленные в системе:

```
$ sudo yum update -y
```

2. Установите пакеты, необходимые для Docker:

```
$ sudo yum install -y yum-utils docker device-mapper-persistent-data lvm2
```

3. Запустите Docker:

```
$ sudo systemctl start docker
```

4. Включите Docker в качестве системного сервиса:

```
$ sudo systemctl enable docker
```

- APT:

1. Обновите все программные пакеты, установленные в системе:

```
$ sudo apt-get update  
$ sudo apt-get dist-upgrade
```

2. Установите пакеты, необходимые для Docker:

```
$ sudo apt-get install docker-engine
```

3. Запустите Docker:

```
$ sudo systemctl start docker
```

4. Включите Docker в качестве системного сервиса:

```
$ sudo systemctl enable docker
```

1.1.1.3 Отключение SELinux (опционально)

Для долговременного отключения SELinux следует прописать `SELINUX=disabled` в конфигурационном файле `/etc/selinux/config` и перезагрузить систему.

Файл может быть изменен с помощью редактора `vi` (или `vim`):

```
$ sudo vi /etc/selinux/config
```

Содержимое обновленного файла:

```
# This file controls the state of SELinux on the system.
# SELINUX= can take one of these three values:
#     enforcing - SELinux security policy is enforced.
#     permissive - SELinux prints warnings instead of enforcing.
#     disabled - SELinux is fully disabled.
SELINUX=disabled
# SELINUXTYPE= type of policy in use. Possible values are:
#     targeted - Only targeted network daemons are protected.
#     strict - Full SELinux protection.
SELINUXTYPE=targeted
```

Кроме этого, доступно временное отключение SELinux с помощью следующей команды:

```
$ sudo setenforce 0
```

Тем не менее, мы рекомендуем отключить SELinux на постоянной основе, чтобы он не перезапускался при каждой перезагрузке системы.

Не забудьте перезагрузить хост после отключения SELinux.

1.1.1.4 Остановка firewall

Если на хосте используется сервис `firewalld`, требуется его остановка. Можно использовать команду:

```
$ sudo systemctl stop firewalld
```

Альтернативный вариант — отключить применение правил к сетевым пакетам при помощи команды `disable`:

```
$ sudo systemctl disable firewalld
```

1.1.1.5 Настройка доступа к Docker CLI без прав root

Для настройки доступа к Docker CLI без использования прав `root` необходимо создать группу пользователей `docker` (если она еще не существует) и добавить текущего пользователя в эту группу с определенными правами:

```
$ sudo groupadd docker
$ sudo usermod -a -G docker $USER
```

После запуска команд требуется перезайти в систему под текущим пользователем.

1.1.2 Шаг 2. Установка ADCM

ПРИМЕЧАНИЕ:

- В продуктовой среде рекомендуется использовать PostgreSQL в качестве внешней базы данных вместо встроенной SQLite;
- Для просмотра полного списка доступных образов можно обратиться к репозиторию ADCM на Docker Hub;
- Все данные ADCM сохраняются в папке `/adcm/data` внутри Docker-контейнера. Во избежание потерь данных рекомендуем примонтировать к контейнеру том (Docker Volume) с последующей настройкой его бэкапов.

Для установки ADCM необходимо выполнить следующие команды:

1. Скачайте Docker-образ из Arenadata Docker Registry.

```
$ sudo docker pull hub.arenadata.io/adcm/adcm:<version>
```

где `<version>` — это версия Docker-образа в одном из следующих форматов:

- `<major>.<minor>.<patch>` — если необходим конкретный патч ADCM. Пример: 2.0.0.
- `<major>.<minor>` — если необходим последний патч в рамках выбранной версии ADCM. Пример: 2.0.

Обратите внимание, что до выпуска ADCM 2.0.0 для версионирования использовался следующий формат: `YYYY.MM.DD.HH`.

2. Создайте Docker-контейнер на основе загруженного образа. Приведенная ниже команда создает контейнер с ADCM на порте 8000. Все данные контейнера будут храниться в директории `/opt/adcm/`.

```
$ sudo docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data  
hub.arenadata.io/adcm/adcm:<version>
```

Если требуется запустить ADCM с включенным SELinux, при создании Docker-контейнера необходимо указать опцию `:Z` в параметре, содержащем имя каталога данных:

```
$ sudo docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data:Z  
hub.arenadata.io/adcm/adcm:<version>
```

Опционально при создании контейнера можно указать переменные окружения, перечисленные ниже. Каждая переменная указывается в формате `-e <name>=<value>`. Пример:

```
$ sudo docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data  
hub.arenadata.io/adcm/adcm:<version> -e LOG_LEVEL="INFO" -e  
DEFAULT_ADCM_URL="http://10.92.40.40:8000/"
```

1.1.3 Шаг 3. Запуск ADCM

Поскольку ADCM представляет собой контейнер под управлением Docker, за его запуск и установку отвечают стандартные Docker-команды:

- Запуск ADCM:

```
$ sudo docker start adcm
```

- Остановка ADCM:

```
$ sudo docker stop adcm
```

- Автозапуск ADCM. Чтобы настроить автоматический запуск Docker-контейнера в случае непредвиденных ошибок, используйте следующую команду:

```
$ sudo docker update --restart=on-failure adcm
```

ПРИМЕЧАНИЕ:

- Дополнительная информация по политикам перезапуска Docker-контейнеров (restart policies) может быть получена из документации Docker;
- Если для текущего пользователя настроен доступ к Docker CLI без использования прав `root`, слово `sudo` в предыдущих командах необязательно.

1.1.4 Шаг 4. Проверка установки

1.1.4.1 Проверка статуса Docker-контейнера

Чтобы убедиться в успешной установке Docker-контейнера `adcm`, выполните следующую команду.

```
$ sudo docker container ls
```

Результат команды приведен ниже. Убедитесь, что в столбце `STATUS` выводится значение `Up`:

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
74b5628146e5	hub.arenadata.io/adcm/adcm:2.0	"/etc/startup.sh"	4 days ago	Up 21 minutes	0.0.0.0:8000->8000/tcp	adcm

1.1.4.2 Проверка доступности порта 8000

После установки и запуска ADCM его web-интерфейс должен стать доступным на 8000 порте развернутого контейнера. Для проверки этого выполните команды:

- `netstat` — для вывода информации о состоянии порта.

```
$ sudo netstat -ntpl | grep 8000
```

Результат команды:

```
tcp6      0      0 :::8000          :::*              LISTEN
1514/docker-proxy-c
```

- `curl` — для проверки URL-соединения.

```
$ curl http://localhost:8000
```

Результат команды:

```
<!doctype html>
<html lang="en">
<head>
  <meta charset="utf-8">
  <title>Arenadata Cluster Manager</title>
  <base href="/">
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <link rel="icon" type="image/x-icon" href="assets/favicon.ico">
<link rel="stylesheet" href="styles.10db6328264e0907c52f.css"></head>
<body>
  <app-root></app-root>
<script src="runtime-es2015.7eb406ed18bf0258cd35.js" type="module"></script><script
src="runtime-es5.7eb406ed18bf0258cd35.js" nomodule defer></script><script
src="polyfills-es5.2e224d70daec4412d3c2.js" nomodule defer></script><script
src="polyfills-es2015.37e0553ac06970d6a5b5.js" type="module"></script><script
src="main-es2015.39851da0ebf9ed6fec45.js" type="module"></script><script src="main-
es5.39851da0ebf9ed6fec45.js" nomodule defer></script></body>
</html>
```

1.1.4.3 Проверка web-интерфейса ADCM

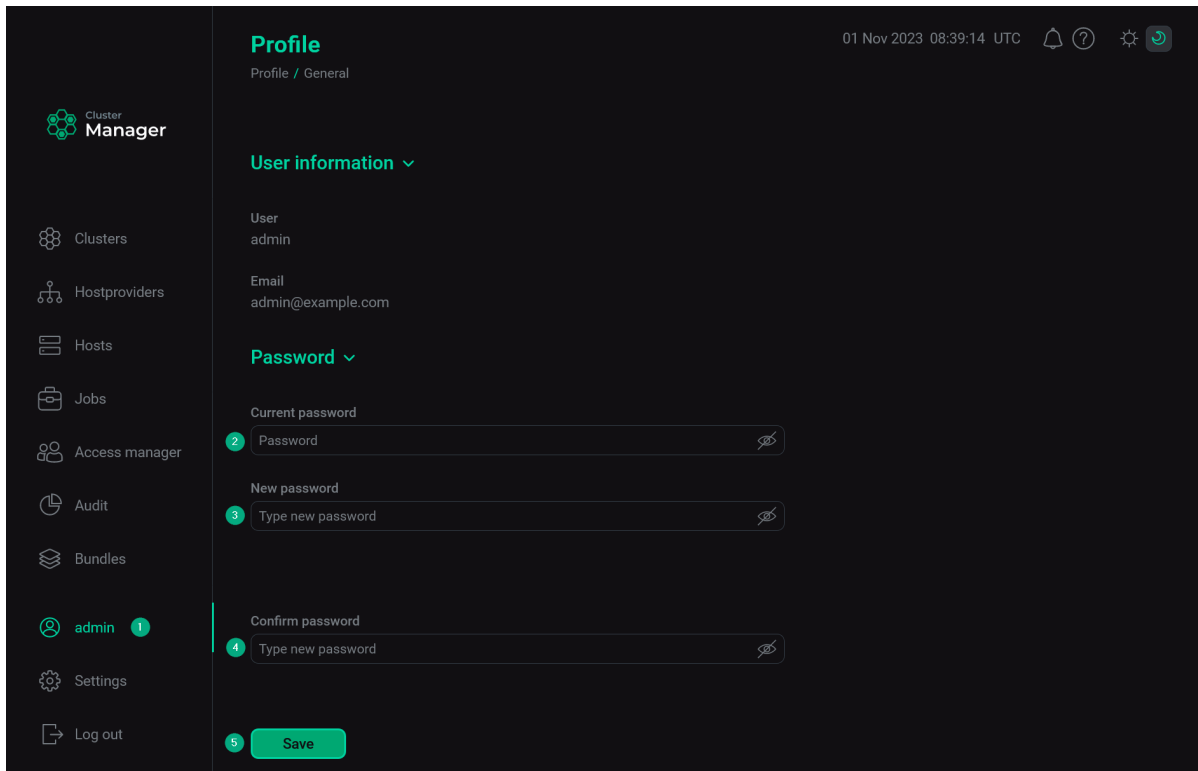
При установке ADCM автоматически создается пользователь с административными правами. Чтобы зайти в ADCM, используйте следующие учетные данные:

- *User:* admin
- *Password:* admin

Для изменения пароля требуется выполнить следующие действия:

1. Откройте профиль текущего пользователя в левом навигационном меню.
2. Введите текущий пароль в поле *Current password*.

3. Укажите новый пароль в поле *New password*.
4. Подтвердите новый пароль в поле *Confirm password*.
5. Нажмите *Save*.

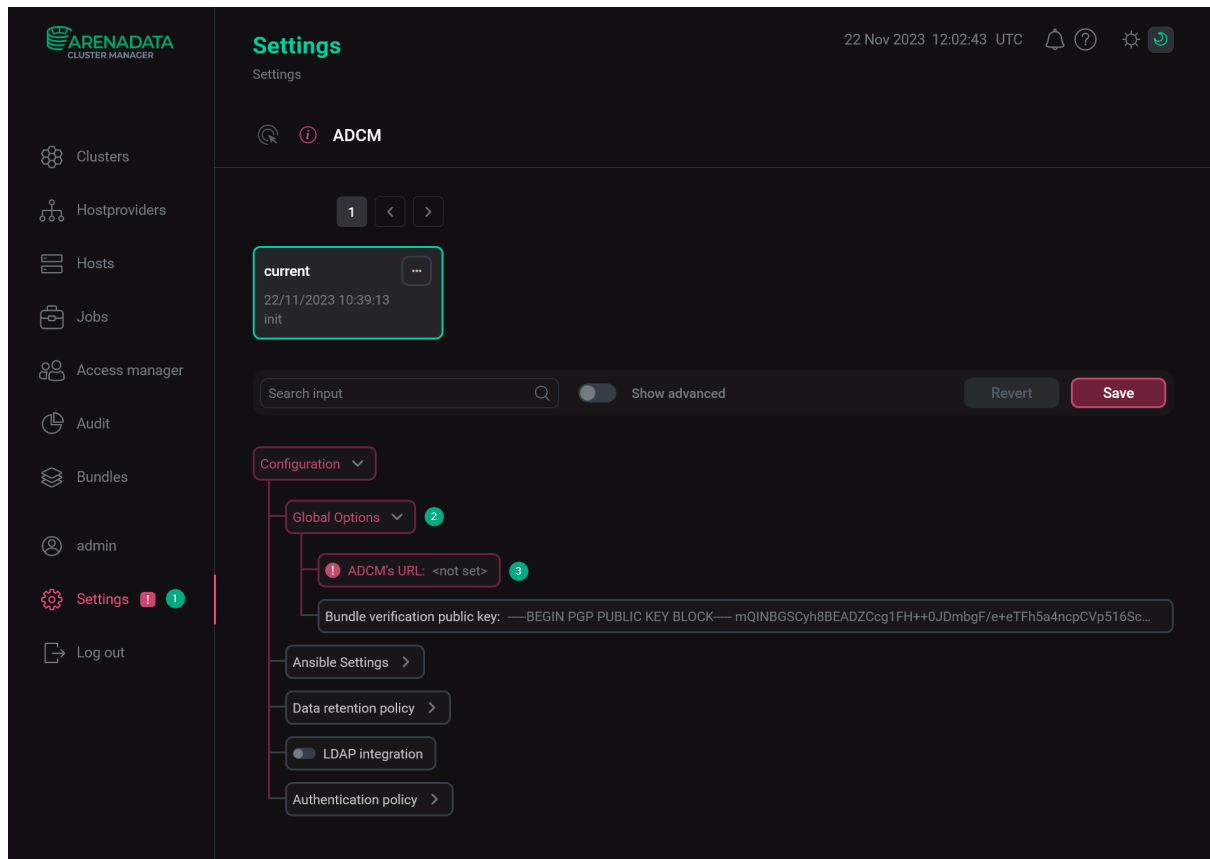


The screenshot shows the 'Profile' page in the Cluster Manager interface. The page title is 'Profile' with a subtitle 'Profile / General'. The user information is displayed as 'User: admin' and 'Email: admin@example.com'. The 'Password' section contains three input fields: 'Current password' (with a red '2' next to it), 'New password' (with a red '3' next to it), and 'Confirm password' (with a red '4' next to it). A 'Save' button is located at the bottom of the form, with a red '5' next to it. The left sidebar contains navigation items: Clusters, Hostproviders, Hosts, Jobs, Access manager, Audit, Bundles, admin (with a red '1' next to it), Settings, and Log out.

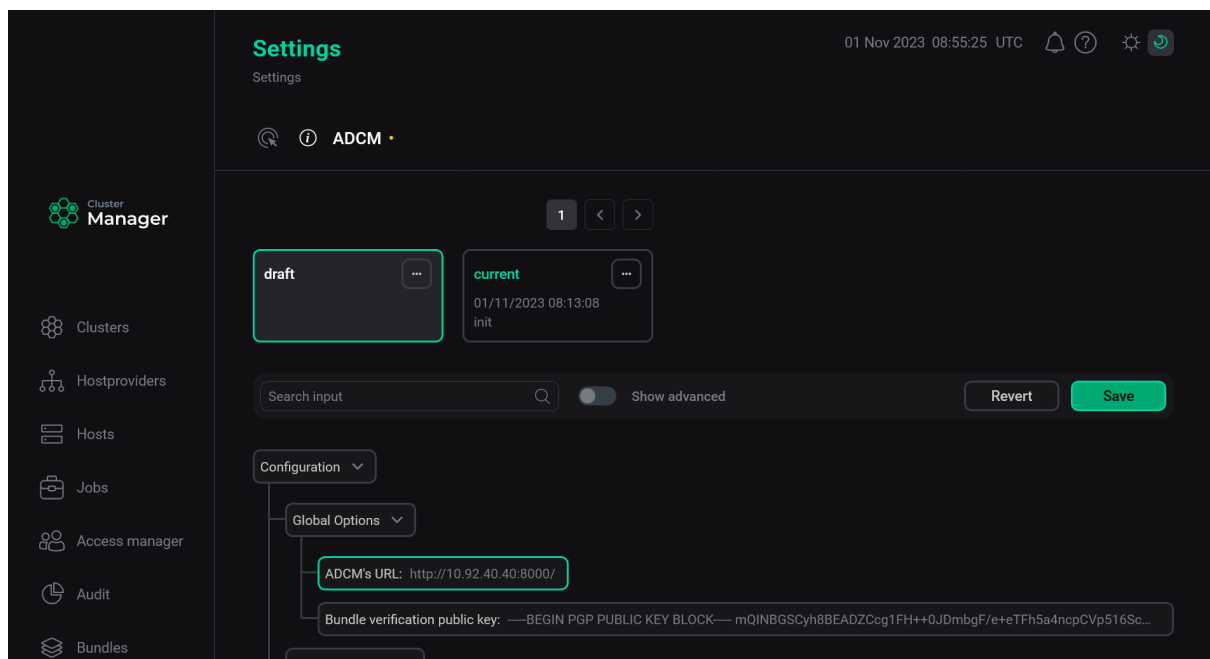
1.1.4.4 Установка URL ADCM

Для успешного обмена данными между ADCM и развернутыми с его помощью кластерами важно корректное определение внешнего адреса ADCM. В дальнейшем этот адрес будет использоваться для отправки в ADCM статусов кластерных компонентов. Для указания URL ADCM используется переменная окружения `DEFAULT_ADCM_URL` при создании контейнера ADCM. Если переменная не была указана, необходимо установить URL вручную в web-интерфейсе ADCM. Для этого выполните шаги:

1. Выберите пункт *Settings* в левом навигационном меню.
2. Раскройте узел *Global Options* в дереве конфигурационных настроек.
3. Нажмите на значение поля *ADCM's URL*.



4. В открывшемся окне измените значение *ADCM's URL* и нажмите *Apply*.
5. Нажмите *Save* для сохранения внесенных изменений.



1.1.5 Шаг 5. Обновление ADCM

Перед выполнением обновления рекомендуется сделать бэкап данных ADCM.

Для обновления ADCM до нужной версии необходимо выполнить шаги:

1. Остановите Docker-контейнер `adcm`.

```
$ sudo docker stop adcm
```

2. Удалите Docker-контейнер `adcm`.

```
$ sudo docker container rm adcm
```

3. Загрузите из Arenadata Docker Registry необходимый Docker-образ.

```
$ sudo docker pull hub.arenadata.io/adcm/adcm:<version>
```

где `<version>` — это версия Docker-образа в одном из следующих форматов:

- `<major>.<minor>.<patch>` — если необходим конкретный патч ADCM. Пример: 2.0.0.
- `<major>.<minor>` — если необходим последний патч в рамках выбранной версии ADCM. Пример: 2.0.

Обратите внимание, что до выпуска ADCM 2.0.0 для версионирования использовался следующий формат: `YYYY.MM.DD.HH`.

4. Создайте новый Docker-контейнер на базе загруженного образа.

```
$ sudo docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data  
hub.arenadata.io/adcm/adcm:<version>
```

5. Запустите новый Docker-контейнер `adcm`.

```
$ sudo docker start adcm
```

1.2 Подготовка хостов

1.2.1 Добавление хостпровайдера SSH

Перед установкой нового кластера необходимо добавить информацию о его будущих хостах в ADCM. Это задача хостпровайдеров — специальных плагинов, отвечающих за взаимодействие между хостами и ADCM.

ADCM поддерживает несколько типов хостпровайдеров. Эта глава описывает добавление хостпровайдера SSH, который позволяет подключаться к уже созданным (в облаке или on-premises) хостам с использованием SSH-протокола. Шаги для добавления хостпровайдера SSH в ADCM приведены ниже.

1.2.1.1 Шаг 1. Загрузка бандла

Дистрибутивы хостпровайдеров поставляются в виде [бандлов](#). В случае хостпровайдера SSH бандл представляет собой обычный архив, описывающий логику установки SSH-соединения к заданному хосту.

Для добавления бандла выполните шаги:

- На сайте продукта Arenadata Cluster Manager нажмите Скачать. Другой способ — перейдите на сайт <https://network.arenadata.io/> и выберите продукт Arenadata Cluster Manager.
- В открывшемся списке релизов нажмите Download в строке с необходимой версией ADCM.
- На следующей странице загрузите необходимый бандл. Бандлы для хостпровайдера SSH называются SSH Common Bundle.

1.2.1.2 Шаг 2. Загрузка бандла хостпровайдера в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.

1.2.1.3 Шаг 3. Создание хостпровайдера на базе загруженного бандла

Для создания в ADCM нового хостпровайдера на базе загруженного бандла необходимо выполнить шаги:

1. Выберите пункт *Hostproviders* в левом навигационном меню и нажмите *Create provider*.
2. В открывшемся окне:
 - Выберите загруженный бандл в поле *Type*.
 - Выберите версию бандла в поле *Version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа хостпровайдеров.
 - Введите имя хостпровайдера в поле *Name*.
 - При необходимости добавьте описание хостпровайдера в поле *Description*.
 - Нажмите *Create*.

3. После успешного добавления хостпровайдер отображается на странице *Hostproviders*.

1.2.2 Добавление хостов

В ADCM под хостом понимается любая физическая либо виртуальная машина. При этом не имеет значения, какая инфраструктура для нее используется — облако или bare metal.

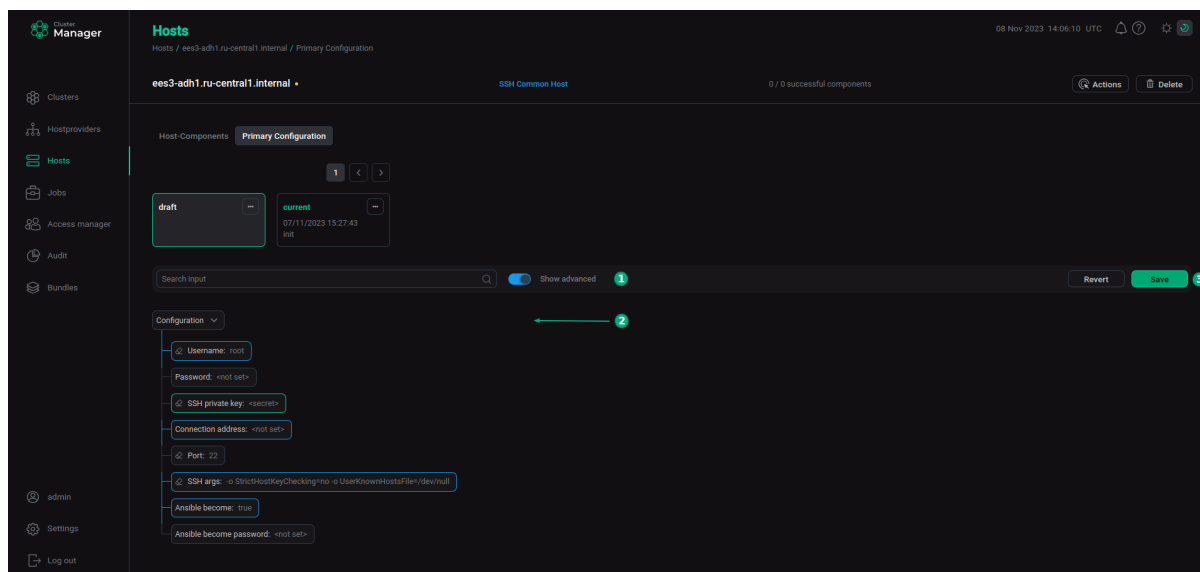
Процесс настройки хостов в ADCM отличается в зависимости от выбранного типа хостпровайдера. Эта статья описывает, как добавлять и настраивать хосты в ADCM на базе хостпровайдера SSH, который является простейшим коннектором к существующим хостам.

Перед добавлением хостов в ADCM на основе хостпровайдера SSH убедитесь в следующем:

- Хостпровайдер SSH установлен в ADCM;
- Хосты с возможностью подключения по протоколу SSH предварительно созданы в облаке или on-premises. Обратите внимание, что эта операция выполняется вручную и не поддерживается текущим хостпровайдером.

Шаги для добавления хостов в ADCM на базе хостпровайдера SSH:

1. Выберите пункт левого навигационного меню *Hosts* и нажмите *Create host*.
2. В открывшемся окне:
 - Выберите хостпровайдер в поле *Hostprovider*. Обратите внимание, что хостпровайдеры отображаются в списке с именами, присвоенными им на этапе добавления в ADCM.
 - Введите FQDN хоста в поле *Name*. FQDN должен содержать имя хоста и домен.
 - Нажмите *Create*.
3. Вернитесь на страницу *Hosts* и перейдите к настройке добавленного хоста. Для этого нажмите на имя хоста в столбце *Name*.
4. В открывшемся окне выберите вкладку *Primary Configuration*, переведите в активное состояние переключатель *Show advanced*, заполните необходимые параметры и нажмите *Save*.



5. Вернитесь на страницу *Hosts*. В строке, содержащей добавленный хост, нажмите на иконку в столбце *Actions* и выберите пункт выпадающего меню *Check connection*. Затем подтвердите действие в открывшемся окне.
6. Дождитесь завершения проверки соединения. Если проверка выполнена успешно, это означает, что хост настроен корректно и может быть добавлен в новый кластер.
7. После завершения проверки соединения установите *statuschecker*. Это специальный демон, который будет периодически проверять состояние служб и компонентов, установленных на хосты кластера через ADCM. Чтобы установить этот демон, необходимо повторно нажать на иконку в столбце *Actions* и выбрать пункт выпадающего меню *Install statuschecker*. Затем подтвердите действие в открывшемся окне.
8. Дождитесь завершения установки *statuschecker*.

1.3 Установка кластера ADH T4

1.3.1 Создание кластера

В ADCM кластер означает набор сервисов, запущенных на группе хостов. Дистрибутивы кластеров для ADCM поставляются в виде бандлов. В общем случае бандл представляет обычный архив, описывающий логику развертывания кластера.

Шаги для добавления нового кластера приведены далее.

1.3.1.1 Шаг 1. Загрузка бандла

На сайте продукта Arenadata Hadoop нажмите *Скачать*.

Другой способ — перейдите на сайт <https://network.arenadata.io/> и выберите продукт Arenadata Hadoop.

На открывшейся странице скачайте бандл ADH нужной версии.

1.3.1.2 Шаг 2. Загрузка бандла в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.

1.3.1.3 Шаг 3. Создание кластера на базе загруженного бандла

На этапе создания кластера через web-интерфейс ADCM генерируется только сам кластер. Это означает добавление записи о кластере в ADCM — сервисы на этом шаге еще не устанавливаются.

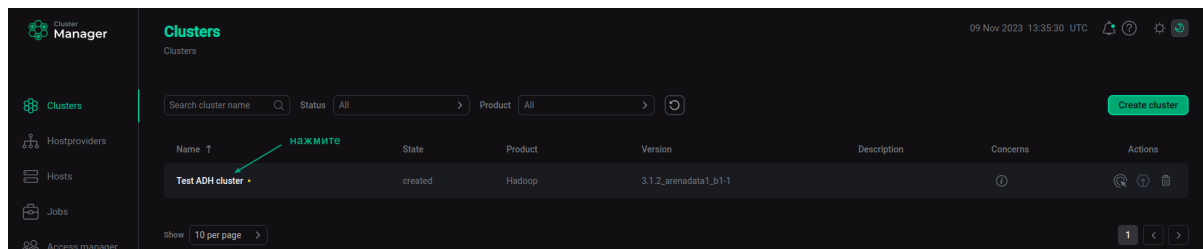
Для создания кластера необходимо выполнить шаги:

1. Выберите пункт *Clusters* в левом навигационном меню и нажмите *Create cluster*.
2. В открывшемся окне требуется описать новый кластер:
 - Выберите загруженный бандл в поле *Product*.
 - Выберите версию бандла в поле *Product version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа продуктов.
 - Введите имя кластера в поле *Cluster name*.
 - При необходимости добавьте описание кластера в поле *Description*.
 - Установите флаг *I accept Terms of Agreement* для подписания лицензионного соглашения. Флаг отображается только для бандлов версии Enterprise. Чтобы ознакомиться с текстом соглашения, перейдите по ссылке *Terms of Agreement*. Подписать соглашение также возможно на странице *Bundles*, нажав на соответствующий продуктовый бандл. Нажмите *Create*.
3. После успешного добавления кластер отображается на странице *Clusters*.

1.3.2 Добавление сервисов

В ADCM сервис означает программное обеспечение, выполняющее некоторую функцию. Примеры сервисов в кластерах ADH: HDFS, HBase, Hive и другие. Для добавления сервисов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.



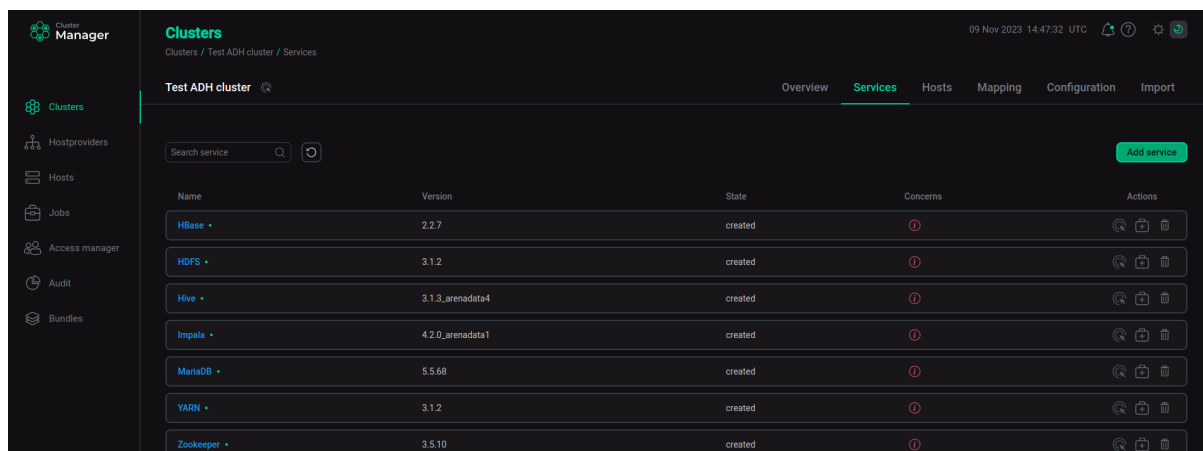
2. На вкладке *Services* на странице кластера нажмите *Add service*.
3. В открывшемся окне выберите сервисы для добавления в кластер и нажмите *Add*.

Минимальный набор сервисов, рекомендуемый для кластеров ADH:

- HDFS;
- YARN;
- Zookeeper (опционально для Community-версии ADH).

Эти сервисы составляют ядро Hadoop и являются достаточными для организации распределенного хранения и обработки данных. Полный перечень сервисов будет зависеть от требований конкретного проекта.

4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



Опциональные сервисы могут быть добавлены в кластер позднее. Процесс добавления сервисов в уже развернутый кластер не отличается от описанного выше.

1.3.3 Добавление хостов в кластер

Для добавления хостов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Hosts* на странице кластера и нажмите *Add hosts*.
3. В открывшемся окне выберите хосты для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные хосты отображаются на вкладке *Hosts*.

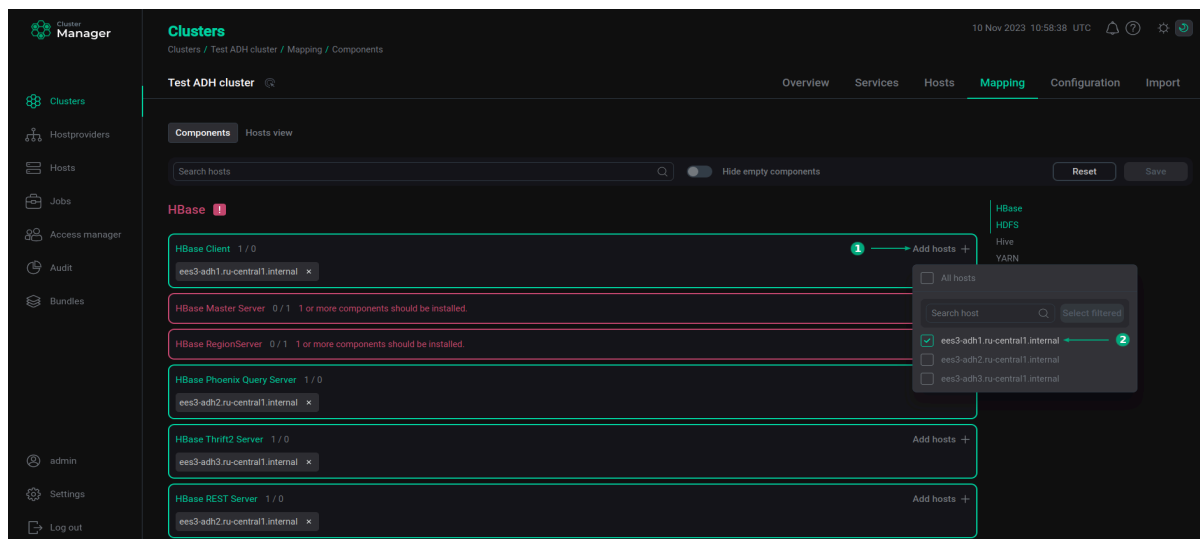
1.3.4 Добавление компонентов

В ADCM компонент означает часть сервиса, которая должна быть развернута на одном или нескольких хостах кластера. Например, сервис HDFS состоит из таких компонентов, как DataNode, NameNode, Client и другие.

Изначально на хостах кластера нет ни одного компонента — распределение компонентов между хостами должно быть проведено вручную.

Для добавления компонентов на хосты кластера:

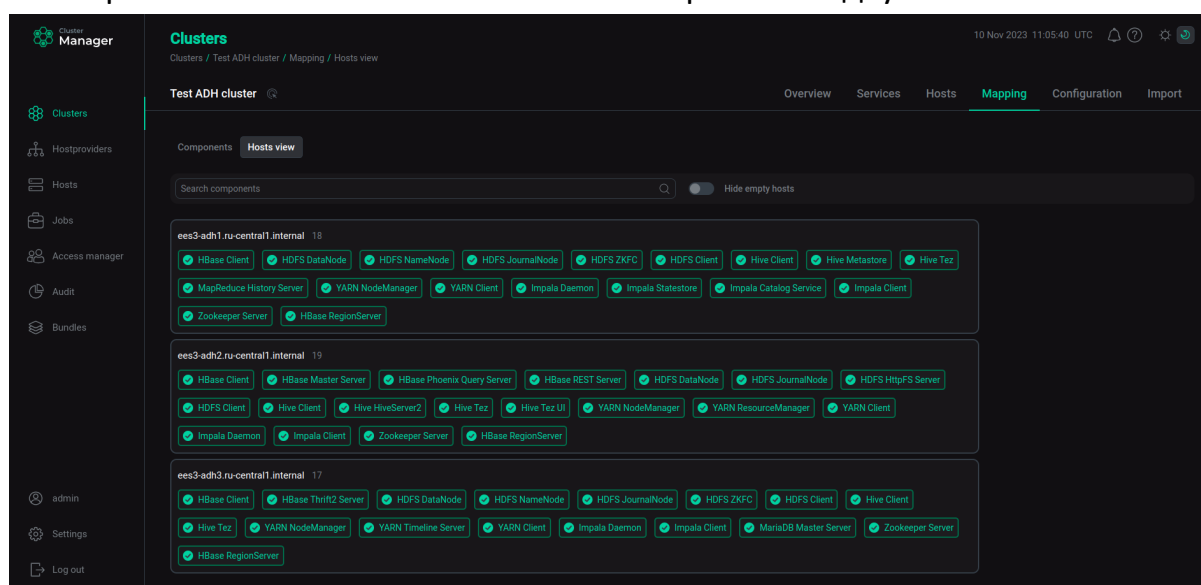
1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера.
3. Распределите все компоненты между хостами кластера. Для назначения хостов компоненту нажмите *Add hosts* и выберите хосты в открывшемся окне.



Каждый сервис может иметь обязательные и опциональные компоненты. Обязательные подсвечиваются красным цветом. Также в интерфейсе ADCM для них выводится информация о том, в каком количестве необходимо установить каждый компонент.

Без назначения нужного числа хостов обязательным компонентам сохранение общей схемы распределения невозможно.

4. После завершения распределения компонентов нажмите *Save*.
5. Для перехода к альтернативному варианту просмотра — с разбиением по хостам — можно открыть вкладку *Hosts view*.



1.3.5 Настройка сервисов

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера. Нажмите на имя сервиса в столбце *Name*.
3. В открывшемся окне:
 - Откройте вкладку *Primary Configuration*.
 - Переведите в активное состояние переключатель *Show advanced*.
 - Заполните конфигурационные параметры выбранного сервиса. Поля, подсвеченные красным, обязательны для заполнения.
 - Нажмите *Save*.

1.3.6 Настройка кластера

В большинстве случаев кастомная настройка кластера не требуется — значения параметров можно оставить по умолчанию.

После добавления нового кластера можно выполнить его настройку. Для этого выполните шаги:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Configuration* на странице кластера и переведите в активное состояние переключатель *Show advanced*.
3. Заполните все необходимые конфигурационные параметры и нажмите *Save*.

Вкладка *Configuration* для кластера включает следующие параметры:

- Раздел *Reliability control*. Отображается, если установлен флажок *Show advanced*. Он содержит следующие настройки:
 - *retries* — количество попыток для операций в кластере (по умолчанию 20);
 - *delay* — задержка между попытками в секундах (по умолчанию 10);
 - *timeout* — тайм-аут проверки сетевых портов в секундах (по умолчанию 60).
- Пути к репозиториям CentOS, Red Hat и ALT Linux, которые будут использованы в процессе установки в зависимости от операционной системы кластера:
 - *monitoring* — репозиторий для установки мониторинга;
 - *zookeeper_repo* — YUM-репозиторий Arenadata Zookeeper;
 - *arenadata* — YUM-репозиторий Arenadata;
 - *arenadata_postgres* — YUM-репозиторий Arenadata PostgreSQL;
 - *ranger* — репозиторий для установки Ranger client.
- Параметр *DockerRegistry*, указывающий адрес Docker Registry. Значение по умолчанию — *hub.adsw.io* — Docker Registry из репозитория Arenadata.
- Раздел *Kerberos*, содержащий параметры, перечисленные в таблице ниже. Некоторые параметры становятся видимыми только при установленном флажке *Show advanced*.
- Раздел *ssl_default_config* с настройками.

- Параметр `java.io.tmpdir`, указывающий переменную среды для управления `java.io.tmpdir` для YARN, HBase, Hive, Spark.

1.3.7 Установка кластера

1.3.7.1 Шаг 1. Запуск установки

Существует два способа установки:

1. Установка всего кластера. В этом методе все сервисы кластера устанавливаются автоматически друг за другом после применения действия *Install* к кластеру.
2. Установка отдельных сервисов. В этом методе каждый сервис устанавливается вручную путем применения к нему одного или нескольких действий. Этот способ также можно использовать для добавления новых сервисов в уже развернутый кластер.

В обоих методах каждый сервис запускается автоматически после установки.

Установка должна производиться на чистые хосты без предустановленного подобного ПО.

Установка всего кластера. Чтобы установить все сервисы кластера за одну операцию, выполните шаги:

1. Примените действие *Install* к кластеру, выбранному на странице *Clusters*, нажав на иконку в столбце *Actions*.
2. Подтвердите действие в открывшемся окне.
3. Дождитесь завершения установки.

Установка отдельных сервисов. Чтобы установить один или несколько сервисов вручную, необходимо предварительно добавить эти сервисы в кластер и затем применить к ним действия, описанные ниже.

Обратите внимание, что сервисы устанавливаются по очереди друг за другом. Не пытайтесь установить следующий сервис до окончания установки предыдущего.

- Откройте вкладку *Services* на странице кластера. Для сервиса, который должен быть установлен, нажмите на иконку в столбце *Actions* и выберите действие *Install* из выпадающего списка.
- Подтвердите действие в открывшемся окне.
- Дождитесь завершения установки и убедитесь, что статус сервиса изменился с `created` на `installed`.

- Повторите предыдущие шаги для всех сервисов, которые необходимо установить.

При установке сервисов вручную важно соблюдать следующий порядок:

1. Zookeeper
2. HDFS
3. YARN
4. HBase
5. Hive
6. Spark
7. Monitoring

1.3.7.2 Шаг 2. Просмотр результатов установки

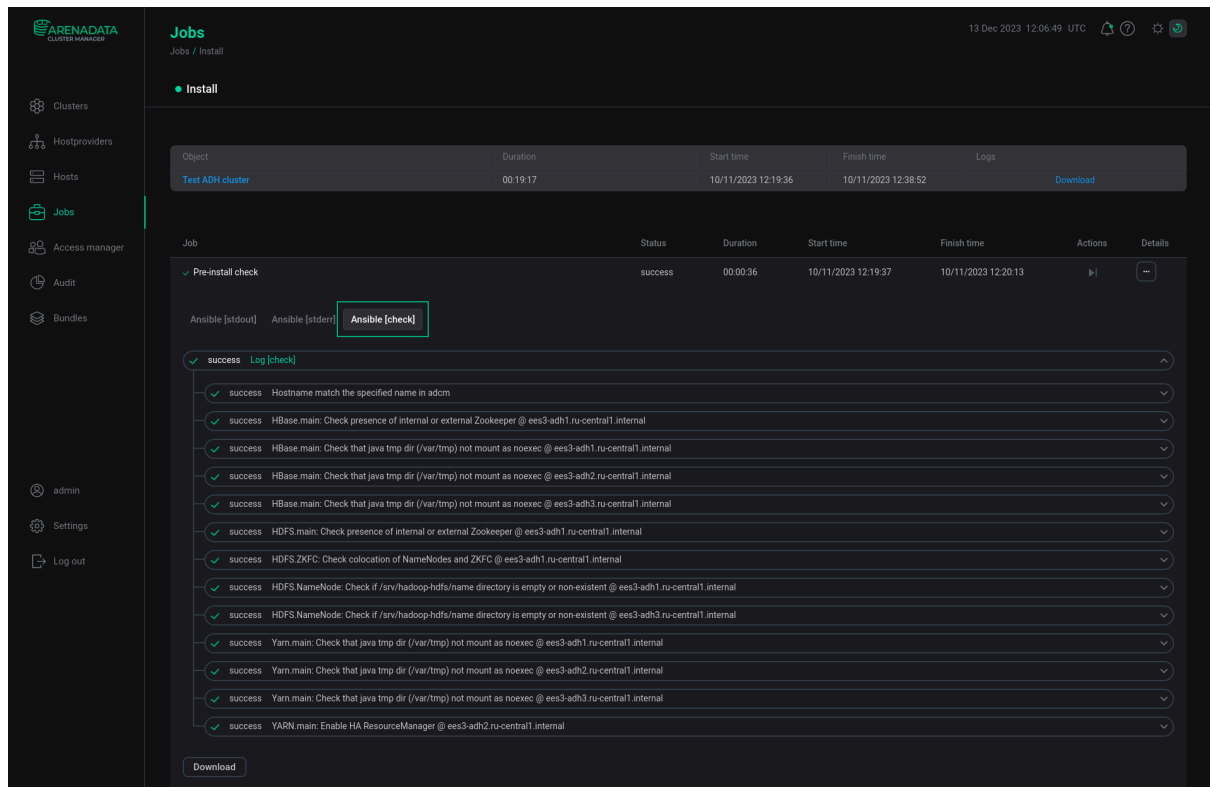
ADCM запускает отдельную задачу (job) для установки, как и для любого другого процесса. Для получения информации о статусе задач предназначена страница *Jobs*.

Если установка завершается успешно, соответствующая ей задача на странице *Jobs* переходит из статуса `running` в `success`; в противном случае — в статус `failed`. Чтобы получить более детальную информацию о возникших ошибках, можно перейти на страницу задачи, кликнув по ней в списке задач на странице *Jobs*. Откроется окно с шагами выполнения задачи. Кликните многоточие в столбце *Details*, чтобы получить дополнительную информацию об определенном шаге.

Страница задачи содержит два основных раздела: *ansible [stdout]* и *ansible [stderr]*. Это технические логи, содержащие информацию из стандартных I/O-поток stdout/stderr.

В дополнение к ним может выводиться третий опциональный раздел *ansible [check]*. Он содержит краткие описания наиболее часто возникающих ошибок. По сравнению с предыдущими разделами, эти описания, как правило, более простые и конкретные.

При возникновении ошибок в процессе установки рекомендуется изучить логи всех трех разделов.



1.3.7.3 Шаг 3. Проверка статуса кластера после установки

В результате успешной установки происходит обновление статуса кластера и его сервисов:

- Статус кластера, отображаемый в столбце *State* на странице *Clusters*, меняет свое значение с `created` на `installed`.
- Статус сервисов, отображаемый на вкладке *Services* страницы кластера, меняет свое значение с `created` на `installed`.

Особенности установки Hive. Компонент Metastore сервиса Hive требует предварительной установки базы данных MySQL/MariaDB и наличия доступов к ней. Сервис MariaDB входит в состав бандла ADH. Таким образом, можно установить и настроить MariaDB через ADCM.

Кроме установки и настройки базы данных MariaDB, необходимо настроить доступ к ней со стороны сервиса Hive. Простейший пример предоставления доступов с использованием команды `GRANT`:

```
GRANT ALL ON *.* TO 'root'@'<metastore_fqdn>' identified by '<password>' WITH GRANT OPTION;
```

— где `<metastore_fqdn>` и `<password>` должны быть заменены соответствующими данными сервера MariaDB (FQDN и пароль).

1.4 Установка мониторинга

Мониторинг является опциональным компонентом (однако, рекомендуемым). ADCM поддерживает настройку мониторинга на базе трех сервисов: Diamond, Graphite и Grafana.

Для процесса online-установки необходим отдельный кластер мониторинга. Шаги для установки мониторинга в этом случае приведены ниже. Их можно выполнить до и после установки кластера ADH, кроме заключительного шага Интеграция с кластером ADH, который может быть выполнен только на базе добавленного кластера ADH.

1.4.1 Создание кластера

В ADCM кластер означает набор сервисов, запущенных на группе хостов. Дистрибутивы кластеров для ADCM поставляются в виде бандлов. В общем случае бандл представляет обычный архив, описывающий логику развертывания кластера.

1.4.1.1 Шаг 1. Загрузка бандла

На сайте продукта Arenadata Cluster Manager нажмите *Скачать*.

Другой способ — перейдите на сайт <https://network.arenadata.io/> и выберите продукт Arenadata Cluster Manager.

На следующей странице загрузите необходимый бандл. Бандлы для кластера мониторинга называются Monitoring Bundle.

1.4.1.2 Шаг 2. Загрузка бандла в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.

1.4.1.3 Шаг 3. Создание кластера на базе загруженного бандла

На этапе создания кластера через web-интерфейс ADCM генерируется только сам кластер. Это означает добавление записи о кластере в ADCM — сервисы на этом шаге еще не устанавливаются.

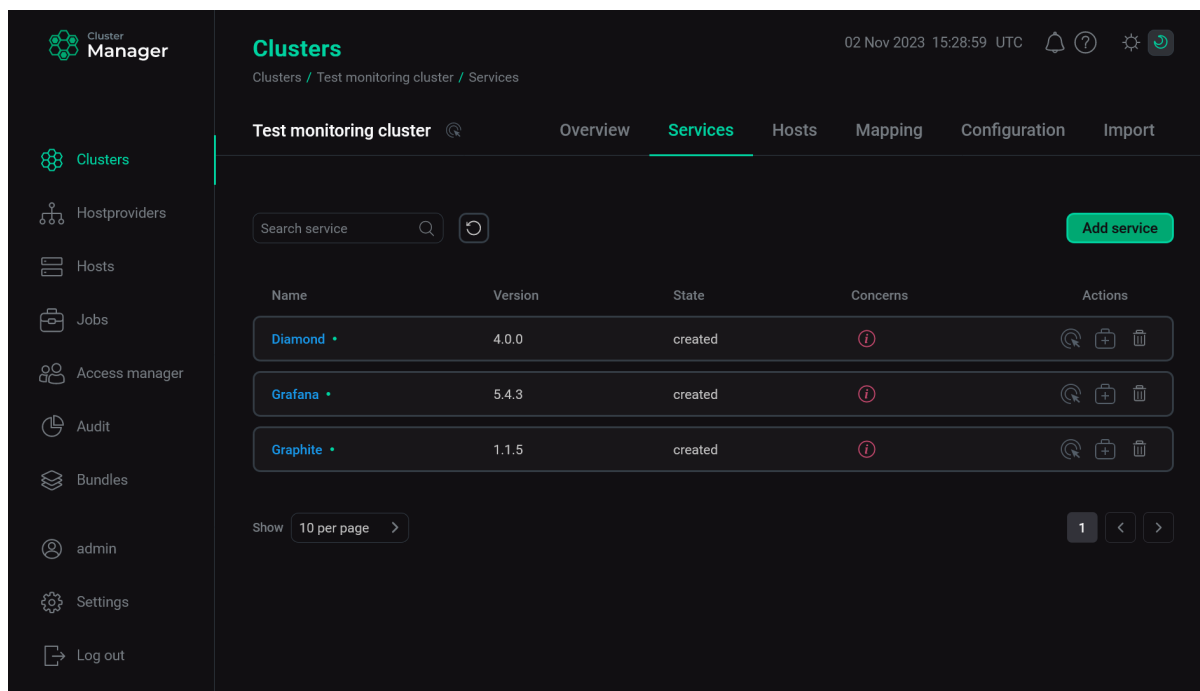
Для создания кластера необходимо выполнить шаги:

1. Выберите пункт *Clusters* в левом навигационном меню и нажмите *Create cluster*.
2. В открывшемся окне требуется описать новый кластер:
 - Выберите загруженный бандл в поле *Product*.
 - Выберите версию бандла в поле *Product version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа продуктов.
 - Введите имя кластера в поле *Cluster name*.
 - При необходимости добавьте описание кластера в поле *Description*.
 - Установите флаг *I accept Terms of Agreement* для подписания лицензионного соглашения. Флаг отображается только для бандлов версии Enterprise. Чтобы ознакомиться с текстом соглашения, перейдите по ссылке *Terms of Agreement*. Подписать соглашение также возможно на странице *Bundles*, нажав на соответствующий продуктовый бандл.
 - Нажмите *Create*.
3. После успешного добавления кластер отображается на странице *Clusters*.

1.4.2 Добавление сервисов

В ADCM сервис означает программное обеспечение, выполняющее некоторую функцию. Для добавления сервисов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. В открывшемся окне выберите сервисы для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



1.4.3 Добавление хостов в кластер

Для добавления хостов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Hosts* на странице кластера и нажмите *Add hosts*.
3. В открывшемся окне выберите хосты для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные хосты отображаются на вкладке *Hosts*.

1.4.4 Добавление компонентов

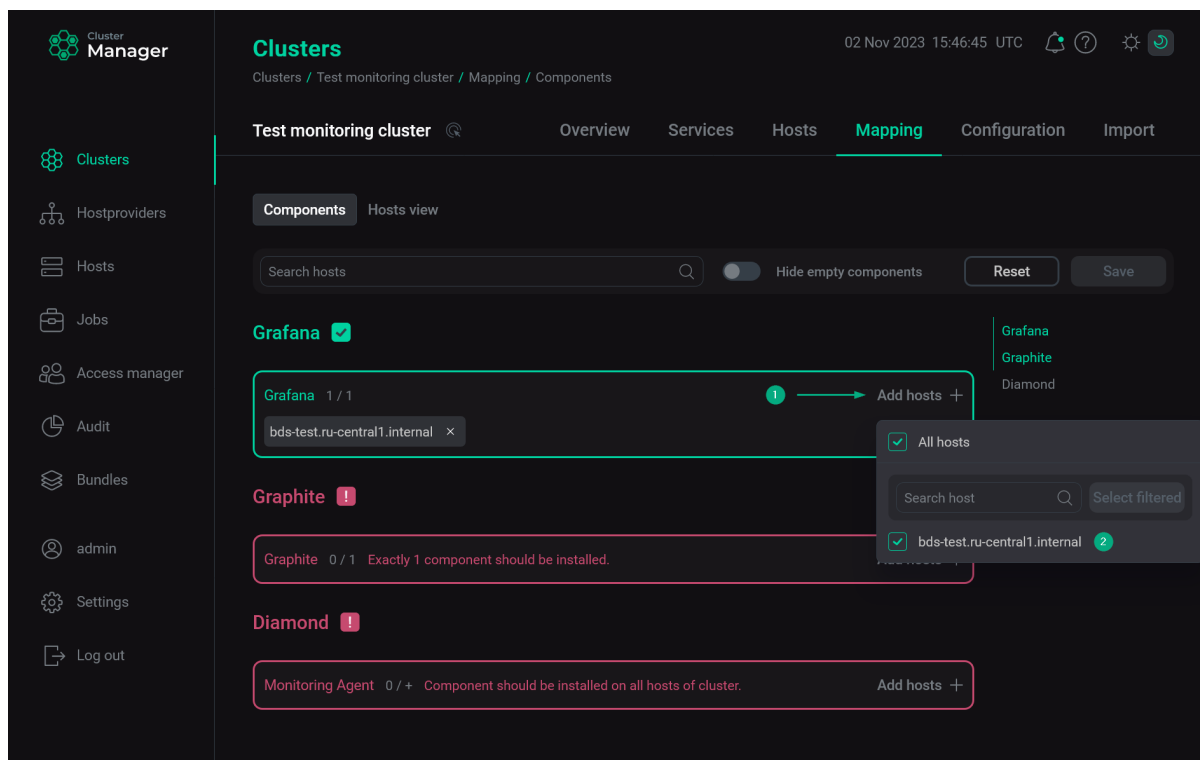
В ADCM компонент означает часть сервиса, которая должна быть развернута на одном или нескольких хостах кластера.

Изначально на хостах кластера нет ни одного компонента — распределение компонентов между хостами должно быть проведено вручную.

Для добавления компонентов на хосты кластера:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера.

3. Распределите все компоненты между хостами кластера. Для назначения хостов компоненту нажмите *Add hosts* и выберите хосты в открывшемся окне.



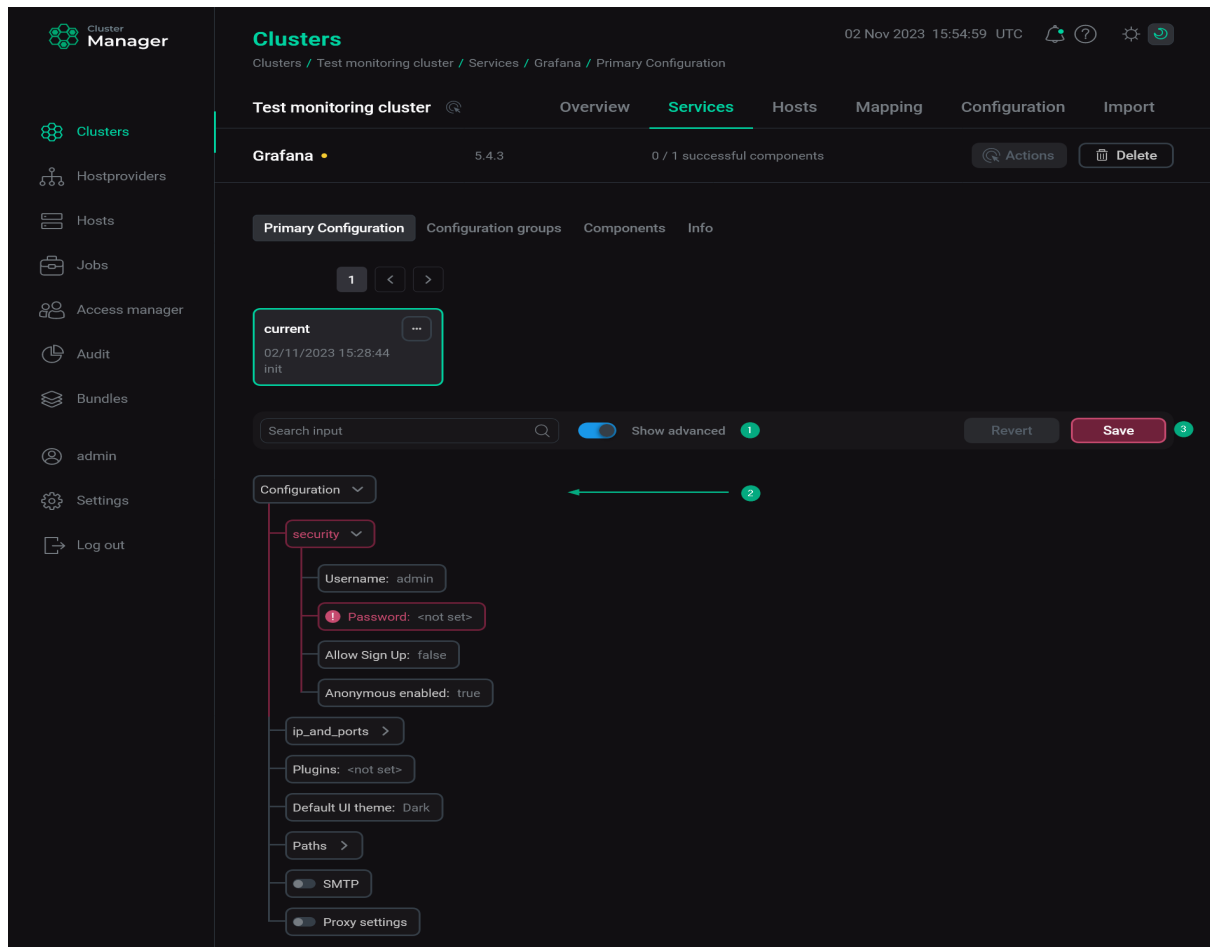
4. После завершения распределения компонентов нажмите *Save*.
5. Для перехода к альтернативному варианту просмотра — с разбиением по хостам — можно открыть вкладку *Hosts view*.

1.4.5 Настройка сервисов

Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.

Откройте вкладку *Services* на странице кластера. Нажмите на имя сервиса в столбце *Name*. В открывшемся окне:

- Откройте вкладку *Primary Configuration*.
- Переведите в активное состояние переключатель *Show advanced*.
- Заполните конфигурационные параметры выбранного сервиса. Поля, подсвеченные красным, обязательны для заполнения.
- Нажмите *Save*.



1.4.6 Установка кластера

1.4.6.1 Шаг 1. Запуск установки

1. Запустите действие *Install* для кластера, выбранного на странице *Clusters*, нажав на иконку в столбце *Actions*. Подтвердите операцию в открывшемся окне. Все сервисы устанавливаются по очереди.
2. Дождитесь завершения процесса установки.

1.4.6.2 Шаг 2. Проверка результатов

Graphite. Чтобы проверить доступность web-интерфейса Graphite, необходимо выполнить шаги:

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Web-interface TCP port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Graphite при

установке кластера мониторинга. Значение 80 (используемое по умолчанию) можно не указывать.

2. Разверните дерево в левой части открывшейся формы. Здесь отображаются метрики System метрики с группировкой по хостам. Можно выбрать любую из метрик для просмотра динамики ее изменения в виде графика в правой части экрана.

Grafana. Чтобы проверить доступность web-интерфейса Grafana, необходимо выполнить шаги:

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Grafana при установке кластера мониторинга (по умолчанию 3000).
2. Выберите пункт меню *Home* в открывшемся окне.
3. Откройте любой из доступных дашбордов (dashboards). Например, *Arenadata System metrics*.
4. Проанализируйте доступные графики.

1.4.7 Интеграция с кластером ADH

Для настройки мониторинга нового кластера необходимо интегрировать его с предварительно созданным кластером мониторинга и добавить отдельный сервис Monitoring.

Шаги для интеграции кластера мониторинга и кластера ADH описаны ниже. Все действия выполняются на базе кластера ADH.

1.4.7.1 Шаг 1. Импорт настроек мониторинга

- Выберите кластер на странице Clusters. Для этого нажмите на имя кластера в столбце Name.
- Откройте вкладку Import на странице кластера.
- Выберите сервисы Graphite и Grafana ранее добавленного кластера мониторинга.
- Нажмите Import.

1.4.7.2 Шаг 2. Добавление сервиса мониторинга

Для добавления сервисов в кластер:

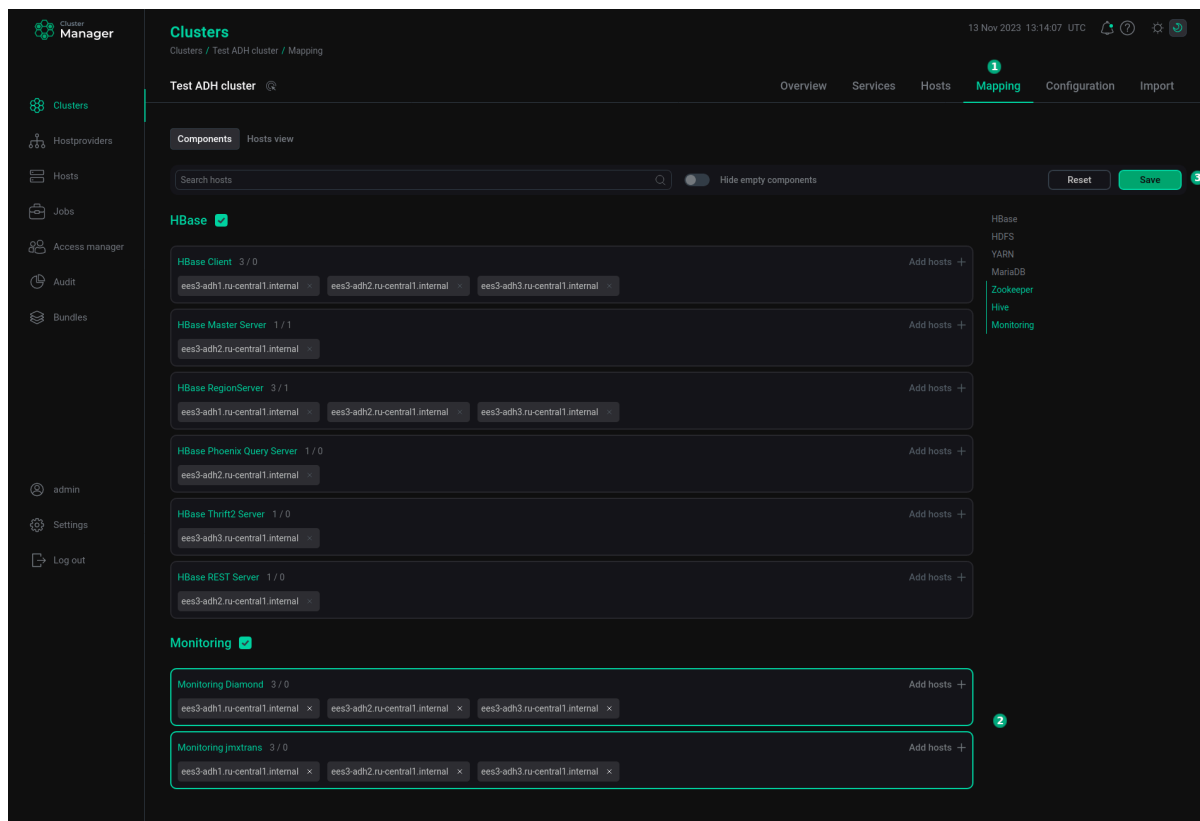
1. Выберите кластер на странице Clusters. Для этого нажмите на имя кластера в столбце Name.

2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. Выберите сервис *Monitoring* в открывшемся окне и нажмите *Add*.
4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.

1.4.7.3 Шаг 3. Распределение компонентов мониторинга

Для добавления компонентов на хосты кластера:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера, добавьте компоненты мониторинга на все хосты кластера и нажмите *Save*.



1.4.7.4 Шаг 4. Установка сервиса мониторинга

Откройте вкладку *Services* на странице кластера. Нажмите на иконку в столбце *Actions* и выберите действие *Install* в строке, содержащей сервис *Monitoring*; подтвердите действие в открывшемся окне.

Дождитесь завершения процесса установки. В результате статус сервиса мониторинга изменяет свое значение с `created` на `installed`.

После установки кластера мониторинга необходимо перезапустить кластер ADH, чтобы просматривать метрики.

1.4.7.5 Шаг 5. Проверка результатов

По завершении установки необходимо убедиться в следующем:

- Системные метрики собираются с хостов вашего кластера, а не только с хоста, на котором ранее был развернут кластер мониторинга.
- Наряду с системными метриками, с хостов кластера поступают сервисные метрики. Например, в кластерах ADH поддерживаются метрики для таких сервисов, как HDFS, YARN, HBase, Spark.

Оба пункта можно проверить в web-интерфейсе Graphite и Grafana.

Graphite

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Web-interface TCP port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Graphite при установке кластера мониторинга. Значение 80 (используемое по умолчанию) можно не указывать.
2. Разверните дерево в левой части открывшейся формы. Здесь отображаются метрики System и сервисные метрики с группировкой по хостам. Можно выбрать любую из метрик для просмотра динамики ее изменения в виде графика в правой части экрана.

Grafana

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Grafana при установке кластера мониторинга (по умолчанию 3000).
2. Выберите пункт меню *Home* в открывшемся окне.
3. Обратите внимание, что новые дашборды доступны. Можно выбрать один из них для просмотра сервисных метрик, поступающих из вашего кластера.
4. Проанализируйте доступные графики.

2 Offline-установка

Этот раздел посвящен offline-установке ADH. Данный способ установки подходит для хостов без доступа к Интернету.

Эта функция доступна только в Enterprise-версии ADH T4.

Процесс offline-установки через Arenadata Cluster Manager (ADCM) включает следующие шаги:

- Установка ADCM (пп. 2.1);
- Подготовка хостов (пп. 2.2);
- Установка кластера Enterprise Tools (пп. 2.3);
- Установка кластера ADH T4 (пп. 2.4);
- Установка мониторинга (пп. 2.5).

2.1 Установка ADCM

Для развертывания ADCM требуется отдельный хост.

2.1.1 Шаг 1. Предварительные требования

2.1.1.1 Требования к software и hardware

ADCM поставляется в виде Docker-образа. Для его установки необходимы:

- Операционная система версии CentOS 7/RHEL 7/Alt Linux SP 8/Astra Linux SE 1.7 Орел. Обратите внимание, что архивные версии не поддерживаются и не тестируются.
- Пользователь с правами sudo.
- Доступ к официальным репозиториям CentOS Extras и CentOS Base. По умолчанию, эти репозитории доступны в ОС. Однако, если они были отключены, потребуется их предварительное включение (например, путем редактирования конфигурационного файла в директории `/etc/yum.repos.d/`).
- Менеджер программных пакетов YUM/APT.
- Хост, отвечающий следующим требованиям:
 - CPU: 4 ядра CPU;
 - RAM: 16 ГБ;
 - HDD: 50 ГБ.

Приведенные системные требования являются минимальными. Целевой сайзинг необходимо рассчитывать исходя из требований конкретной организации.

Также перед началом установки вам нужно отредактировать файл `/etc/sudoers`, чтобы впоследствии пользователь, выполняющий установку ADCM, обладал необходимыми правами. Для этого выполните команду:

```
$ sudo vi /etc/sudoers
```

В этом файле необходимо удалить символ `#` в начале строки `WHEEL_USERS ALL=(ALL) ALL`. Чтобы сохранить это изменение и вернуться в терминал, нажмите `Esc`, после чего наберите `:wq` либо `:wq!`.

2.1.1.2 Установка Docker

Установка Docker с официального сайта не рекомендуется. Вместо этого используйте официальные репозитории дистрибутива Linux. Репозитории, предоставляемые разработчиками операционных систем, как правило, более стабильны и безопасны.

В зависимости от операционной системы и менеджера пакетов, которые вы используете, последовательность установки включает следующие шаги.

- **YUM**

1. Обновите все программные пакеты, установленные в системе:

```
$ sudo yum update -y
```

2. Установите пакеты, необходимые для Docker:

```
$ sudo yum install -y yum-utils docker device-mapper-persistent-data lvm2
```

3. Запустите Docker:

```
$ sudo systemctl start docker
```

4. Включите Docker в качестве системного сервиса:

```
$ sudo systemctl enable docker
```

- **APT**

1. Обновите все программные пакеты, установленные в системе:

```
$ sudo apt-get update  
$ sudo apt-get dist-upgrade
```

2. Установите пакеты, необходимые для Docker:

```
$ sudo apt-get install docker-engine
```

3. Запустите Docker:

```
$ sudo systemctl start docker
```

4. Включите Docker в качестве системного сервиса:

```
$ sudo systemctl enable docker
```

2.1.1.3 Отключение SELinux (опционально)

Для долговременного отключения SELinux следует прописать `SELINUX=disabled` в конфигурационном файле `/etc/selinux/config` и перезагрузить систему.

Файл может быть изменен с помощью редактора vi (или vim):

```
$ sudo vi /etc/selinux/config
```

Содержимое обновленного файла:

```
# This file controls the state of SELinux on the system.
# SELINUX= can take one of these three values:
#     enforcing - SELinux security policy is enforced.
#     permissive - SELinux prints warnings instead of enforcing.
#     disabled - SELinux is fully disabled.
SELINUX=disabled
# SELINUXTYPE= type of policy in use. Possible values are:
#     targeted - Only targeted network daemons are protected.
#     strict - Full SELinux protection.
SELINUXTYPE=targeted
```

Кроме этого, доступно временное отключение SELinux с помощью следующей команды:

```
$ sudo setenforce 0
```

Тем не менее, мы рекомендуем отключить SELinux на постоянной основе, чтобы он не перезапускался при каждой перезагрузке системы.

Не забудьте перезагрузить хост после отключения SELinux.

2.1.1.4 Остановка firewall

Если на хосте используется сервис `firewalld`, требуется его остановка. Можно использовать команду:

```
$ sudo systemctl stop firewalld
```

Альтернативный вариант — отключить применение правил к сетевым пакетам при помощи команды `disable`:

```
$ sudo systemctl disable firewalld
```

2.1.1.5 Настройка доступа к Docker CLI без прав root

Для настройки доступа к Docker CLI без использования прав `root` необходимо создать группу пользователей `docker` (если она еще не существует) и добавить текущего пользователя в эту группу с определенными правами:

```
$ sudo groupadd docker
$ sudo usermod -a -G docker $USER
```

После запуска команд требуется перезайти в систему под текущим пользователем.

2.1.2 Шаг 2. Установка ADCM

Offline-установка ADCM включает следующие шаги:

1. Запросите специальный установочный shell-скрипт у команды поддержки Arenadata.
2. Создайте директорию на хосте, где должен быть развернут ADCM, и разместите в ней полученный скрипт.
3. Запустите скрипт из созданной директории, используя следующую команду:

```
$ sudo bash <script_name> master_adcm
```

где `<script_name>` означает полное имя скрипта с расширением (например, `pack.sh`).

В продуктовой среде рекомендуется использовать PostgreSQL в качестве внешней базы данных вместо встроенной SQLite.

Offline-установка занимает некоторое время. Пожалуйста, дождитесь ее завершения. Результат обработки команды может выглядеть следующим образом:

```
Loaded image: hub.arenadata.io/adcm/adcm:2.0  
e97f2afb04e568386b2755ef8b1d245eb2abfab77c2a402f736e0c1cd99e7b11  
adcm
```

2.1.3 Шаг 3. Запуск ADCM

Поскольку ADCM представляет собой контейнер под управлением Docker, за его запуск и установку отвечают стандартные Docker-команды:

- Запуск ADCM:

```
$ sudo docker start adcm
```

- Остановка ADCM:

```
$ sudo docker stop adcm
```

- Автозапуск ADCM. Чтобы настроить автоматический запуск Docker-контейнера в случае непредвиденных ошибок, используйте следующую команду:

```
$ sudo docker update --restart=on-failure adcm
```

Дополнительная информация по политикам перезапуска Docker-контейнеров (`restart policies`) может быть получена из документации Docker.

Если для текущего пользователя настроен доступ к Docker CLI без использования прав `root`, слово `sudo` в предыдущих командах необязательно.

2.1.4 Шаг 4. Проверка установки

2.1.4.1 Проверка статуса Docker-контейнера

Чтобы убедиться в успешной установке Docker-контейнера `adcm`, выполните следующую команду:

```
$ sudo docker container ls
```

Результат команды приведен ниже. Убедитесь, что в столбце `STATUS` выводится значение `Up`:

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
74b5628146e5	hub.arenadata.io/adcm/adcm:2.0	"/etc/startup.sh"	4 days ago	Up 21 minutes	0.0.0.0:8000->8000/tcp	adcm

2.1.4.2 Проверка доступности порта 8000

После установки и запуска ADCM его web-интерфейс должен стать доступным на 8000 порте развернутого контейнера. Для проверки этого выполните команды:

- `netstat` — для вывода информации о состоянии порта.

```
$ sudo netstat -ntpl | grep 8000
```

Результат команды:

tcp6	0	0	:::8000	:::*	LISTEN
1514/docker-proxy-c					

- `curl` — для проверки URL-соединения.

```
$ curl http://localhost:8000
```

Результат команды:

```
<!doctype html>
<html lang="en">
<head>
  <meta charset="utf-8">
  <title>Arenadata Cluster Manager</title>
  <base href="/">
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <link rel="icon" type="image/x-icon" href="assets/favicon.ico">
<link rel="stylesheet" href="styles.10db6328264e0907c52f.css"></head>
<body>
  <app-root></app-root>
<script src="runtime-es2015.7eb406ed18bf0258cd35.js" type="module"></script><script
src="runtime-es5.7eb406ed18bf0258cd35.js" nomodule defer></script><script
src="polyfills-es5.2e224d70daec4412d3c2.js" nomodule defer></script><script
src="polyfills-es2015.37e0553ac06970d6a5b5.js" type="module"></script><script
src="main-es2015.39851da0ebf9ed6fec45.js" type="module"></script><script src="main-
es5.39851da0ebf9ed6fec45.js" nomodule defer></script></body>
```

</html>

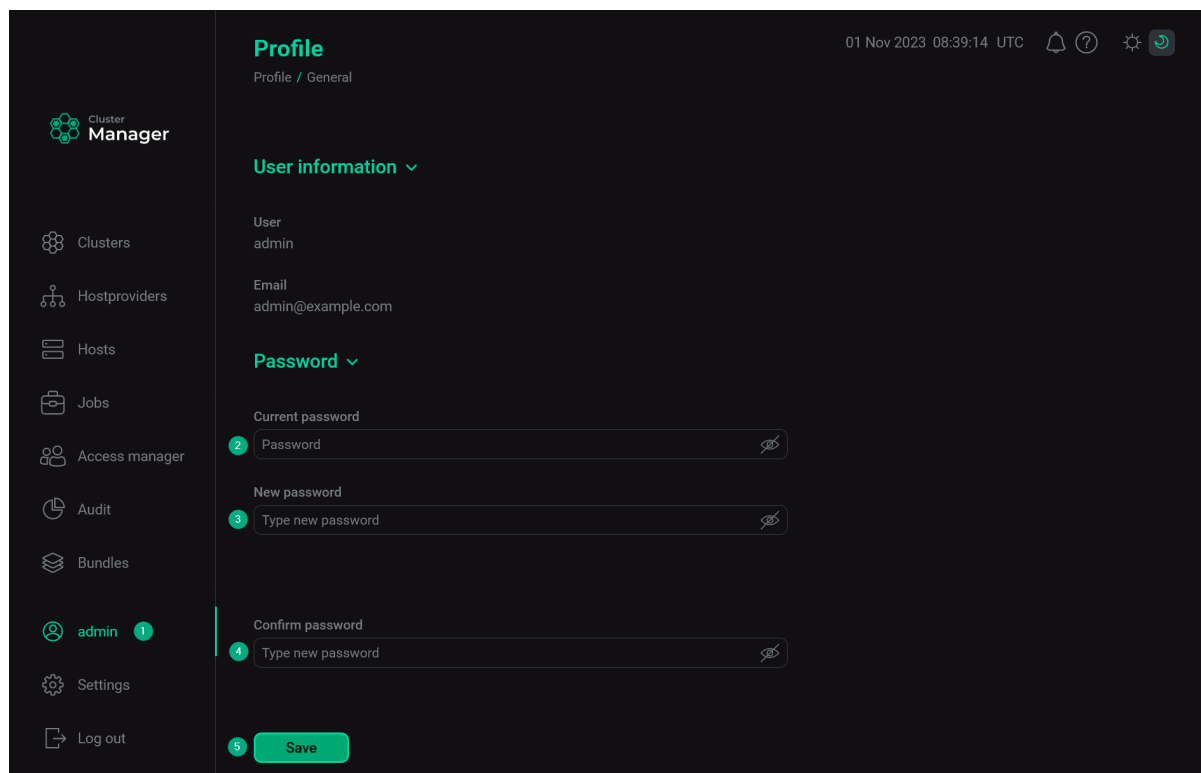
2.1.4.3 Проверка web-интерфейса ADCM

При установке ADCM автоматически создается пользователь с административными правами. Чтобы зайти в ADCM, используйте следующие учетные данные:

- *User*: admin
- *Password*: admin

Для изменения пароля требуется выполнить следующие действия:

1. Откройте профиль текущего пользователя в левом навигационном меню.
2. Введите текущий пароль в поле *Current password*.
3. Укажите новый пароль в поле *New password*.
4. Подтвердите новый пароль в поле *Confirm password*.
5. Нажмите *Save*.

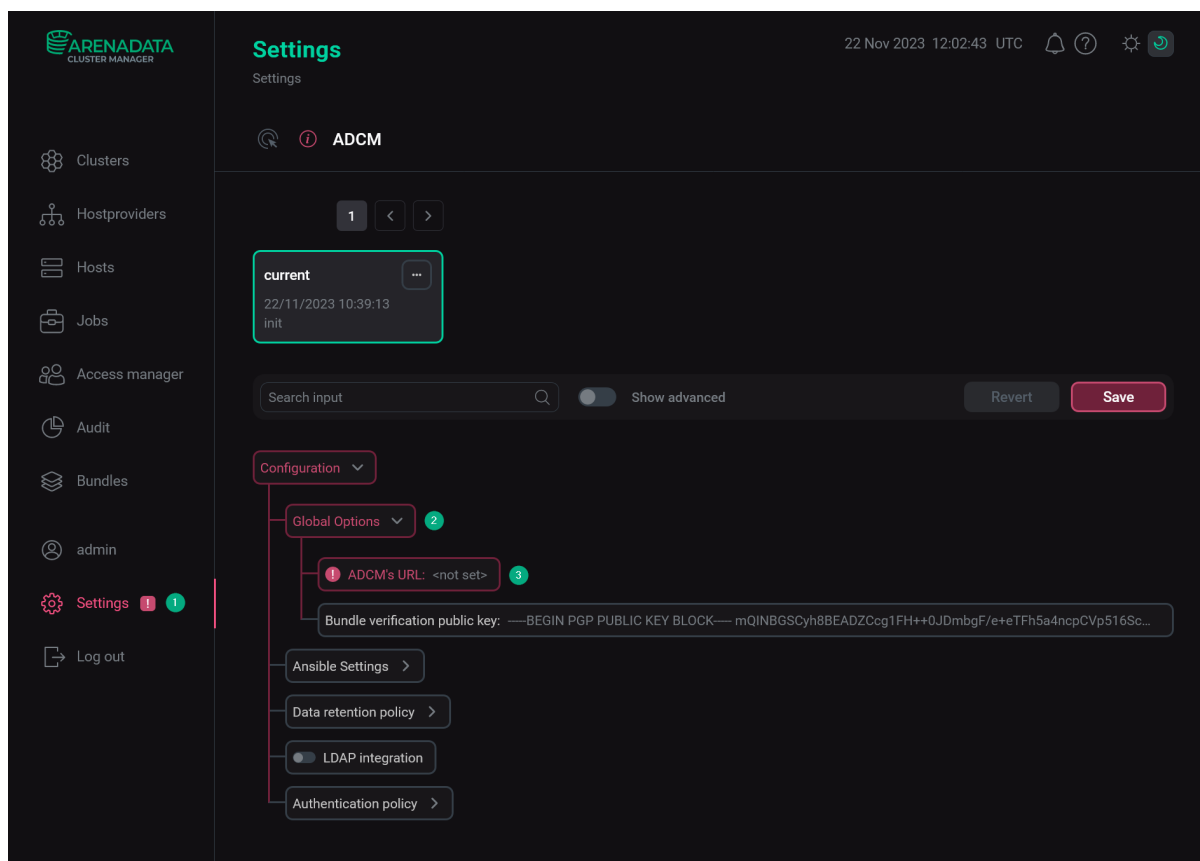


2.1.4.4 Установка URL ADCM

Для успешного обмена данными между ADCM и развернутыми с его помощью кластерами важно корректное определение внешнего адреса ADCM. В дальнейшем этот адрес будет использоваться для отправки в

ADCM статусов кластерных компонентов. Для указания URL ADCM используется переменная окружения `DEFAULT_ADCM_URL` при создании контейнера ADCM. Если переменная не была указана, необходимо установить URL вручную в web-интерфейсе ADCM. Для этого выполните шаги:

1. Выберите пункт *Settings* в левом навигационном меню.
2. Раскройте узел *Global Options* в дереве конфигурационных настроек.
3. Нажмите на значение поля *ADCM's URL*.



4. В открывшемся окне измените значение *ADCM's URL* и нажмите *Apply*.
5. Нажмите *Save* для сохранения внесенных изменений.

2.1.5 Шаг 5. Обновление ADCM

Перед выполнением обновления рекомендуется сделать бэкап данных ADCM. Для обновления ADCM до нужной версии необходимо:

1. Остановите Docker-контейнер `adcm`.

```
$ sudo docker stop adcm
```

2. Удалите Docker-контейнер `adcm`.

```
$ sudo docker container rm adcm
```

3. Загрузите из Arenadata Docker Registry необходимый Docker-образ.

```
$ sudo docker pull hub.arenadata.io/adcm/adcm:<version>
```

где `<version>` — это версия Docker-образа в одном из следующих форматов:

- `<major>.<minor>.<patch>` — если необходим конкретный патч ADCM. Пример: 2.0.0.
- `<major>.<minor>` — если необходим последний патч в рамках выбранной версии ADCM. Пример: 2.0.

Обратите внимание, что до выпуска ADCM 2.0.0 для версионирования использовался следующий формат: `YYYY.MM.DD.HH`.

4. Создайте новый Docker-контейнер на базе загруженного образа.

```
$ sudo docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data  
hub.arenadata.io/adcm/adcm:<version>
```

5. Запустите новый Docker-контейнер adcm.

```
$ sudo docker start adcm
```

Если вы используете ADCM с внешней базой данных, Docker-контейнер нужно запускать с соответствующими переменными окружения.

2.2 Подготовка хостов

2.2.1 Добавление хостпровайдера SSH

Перед установкой нового кластера необходимо добавить информацию о его будущих хостах в ADCM. Это задача хостпровайдеров — специальных плагинов, отвечающих за взаимодействие между хостами и ADCM.

ADCM поддерживает несколько типов хостпровайдеров. Эта глава описывает добавление хостпровайдера SSH, который позволяет подключаться к уже созданным (в облаке или on-premises) хостам с использованием SSH-протокола. Шаги для добавления хостпровайдера SSH в ADCM приведены ниже.

2.2.1.1 Шаг 1. Загрузка бандла

Дистрибутивы хостпровайдеров поставляются в виде бандлов. В случае хостпровайдера SSH бандл представляет собой обычный архив, описывающий логику установки SSH-соединения к заданному хосту.

Для добавления бандла выполните шаги:

1. На сайте продукта Arenadata Cluster Manager нажмите *Скачать*.

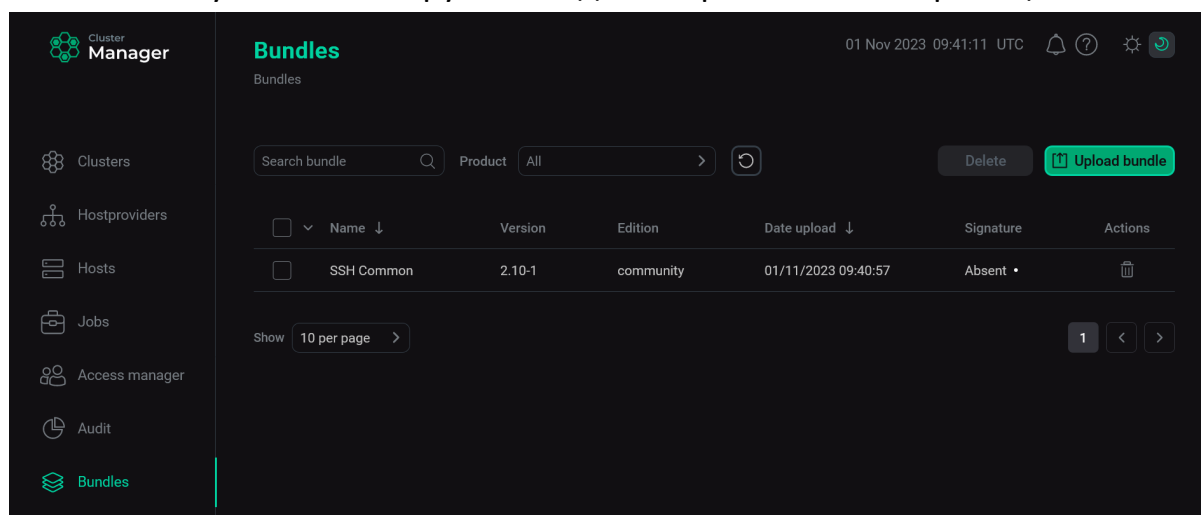
Другой способ — перейдите на сайт <https://network.arenadata.io/> и выберите продукт *Arenadata Cluster Manager*.

2. В открывшемся списке релизов нажмите *Download* в строке с необходимой версией ADCM.
3. На следующей странице загрузите необходимый бандл. Бандлы для хостпровайдера SSH называются *SSH Common Bundle*.

2.2.1.2 Шаг 2. Загрузка бандла хостпровайдера в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.



2.2.1.3 Шаг 3. Создание хостпровайдера на базе загруженного бандла

Для создания в ADCM нового хостпровайдера на базе загруженного бандла необходимо выполнить шаги:

1. Выберите пункт *Hostproviders* в левом навигационном меню и нажмите *Create provider*.
2. В открывшемся окне:
 - Выберите загруженный бандл в поле *Type*.
 - Выберите версию бандла в поле *Version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа хостпровайдеров.
 - Введите имя хостпровайдера в поле *Name*.

- При необходимости добавьте описание хостпровайдера в поле *Description*.
 - Нажмите *Create*.
3. После успешного добавления хостпровайдер отображается на странице *Hostproviders*.

2.2.2 Добавление хостов

В ADCM под хостом понимается любая физическая либо виртуальная машина. При этом не имеет значения, какая инфраструктура для нее используется — облако или bare metal.

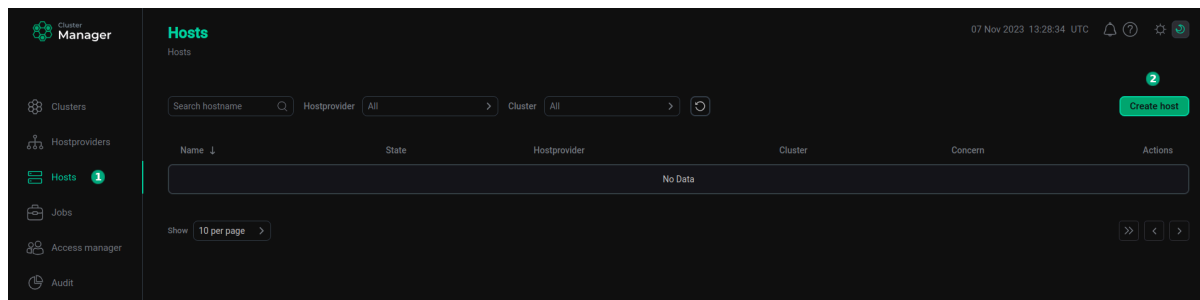
Процесс настройки хостов в ADCM отличается в зависимости от выбранного типа хостпровайдера. Эта статья описывает, как добавлять и настраивать хосты в ADCM на базе хостпровайдера SSH, который является простейшим коннектором к существующим хостам.

Перед добавлением хостов в ADCM на основе хостпровайдера SSH убедитесь в следующем:

- Хостпровайдер SSH установлен в ADCM.
- Хосты с возможностью подключения по протоколу SSH предварительно созданы в облаке или on-premises. Обратите внимание, что эта операция выполняется вручную и не поддерживается текущим хостпровайдером.

Шаги для добавления хостов в ADCM на базе хостпровайдера SSH:

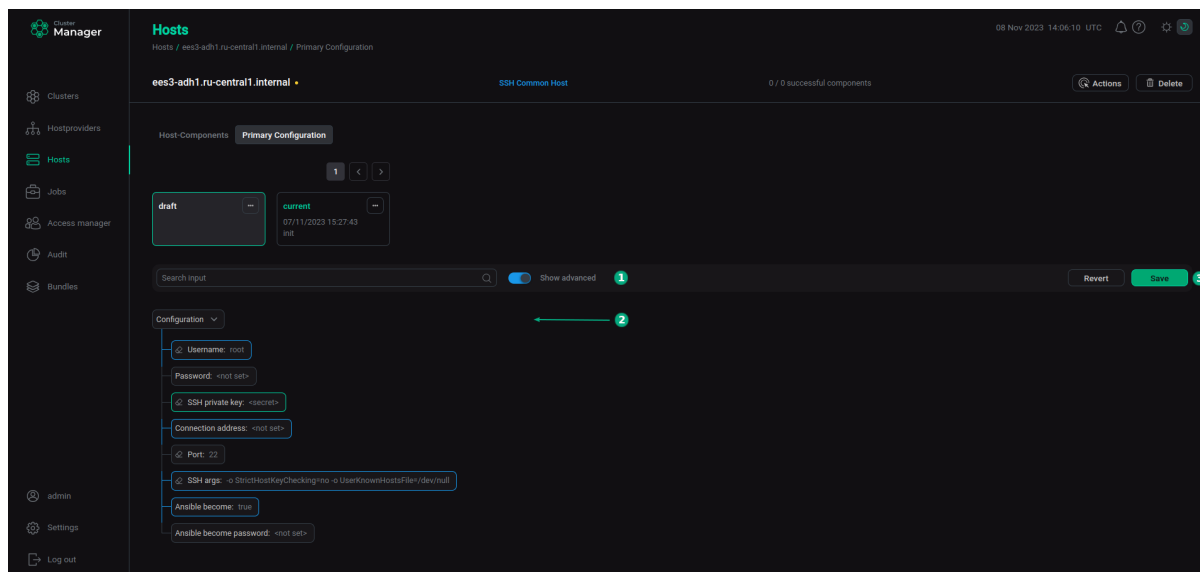
1. Выберите пункт левого навигационного меню *Hosts* и нажмите *Create host*.



2. В открывшемся окне:

- Выберите хостпровайдер в поле *Hostprovider*. Обратите внимание, что хостпровайдеры отображаются в списке с именами, присвоенными им на этапе добавления в ADCM.
- Введите FQDN хоста в поле *Name*. FQDN должен содержать имя хоста и домен.

- Нажмите *Create*.
3. Вернитесь на страницу *Hosts* и перейдите к настройке добавленного хоста. Для этого нажмите на имя хоста в столбце *Name*.
 4. В открывшемся окне выберите вкладку *Primary Configuration*, переведите в активное состояние переключатель *Show advanced*, заполните необходимые параметры и нажмите *Save*.



5. Вернитесь на страницу *Hosts*. В строке, содержащей добавленный хост, нажмите на иконку в столбце *Actions* и выберите пункт выпадающего меню *Check connection*. Затем подтвердите действие в открывшемся окне.
6. Дождитесь завершения проверки соединения. Если проверка выполнена успешно, это означает, что хост настроен корректно и может быть добавлен в новый кластер.
7. После завершения проверки соединения установите *statuschecker*. Это специальный демон, который будет периодически проверять состояние служб и компонентов, установленных на хосты кластера через ADCM. Чтобы установить этот демон, необходимо повторно нажать на иконку в столбце *Actions* и выбрать пункт выпадающего меню *Install statuschecker*. Затем подтвердите действие в открывшемся окне.
8. Дождитесь завершения установки *statuschecker*.

2.3 Установка кластера Enterprise Tools

2.3.1 Создание кластера

В ADCM кластер означает набор сервисов, запущенных на группе хостов. Дистрибутивы кластеров для ADCM поставляются в виде бандлов. В общем случае бандл представляет обычный архив, описывающий логику развертывания кластера.

Шаги для добавления нового кластера приведены ниже.

2.3.1.1 Шаг 1. Получение Enterprise-версии бандла

Пожалуйста, запросите бандл у команды поддержки Arenadata.

2.3.1.2 Шаг 2. Загрузка бандла в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.

2.3.1.3 Шаг 3. Создание кластера на базе загруженного бандла

На этапе создания кластера через web-интерфейс ADCM генерируется только сам кластер. Это означает добавление записи о кластере в ADCM — сервисы на этом шаге еще не устанавливаются.

Для создания кластера необходимо выполнить шаги:

1. Выберите пункт *Clusters* в левом навигационном меню и нажмите *Create cluster*.
2. В открывшемся окне требуется описать новый кластер:
 - Выберите загруженный бандл в поле *Product*.
 - Выберите версию бандла в поле *Product version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа продуктов.
 - Введите имя кластера в поле *Cluster name*.
 - При необходимости добавьте описание кластера в поле *Description*.
 - Установите флаг *I accept Terms of Agreement* для подписания лицензионного соглашения. Флаг отображается только для бандлов версии Enterprise. Чтобы ознакомиться с текстом соглашения,

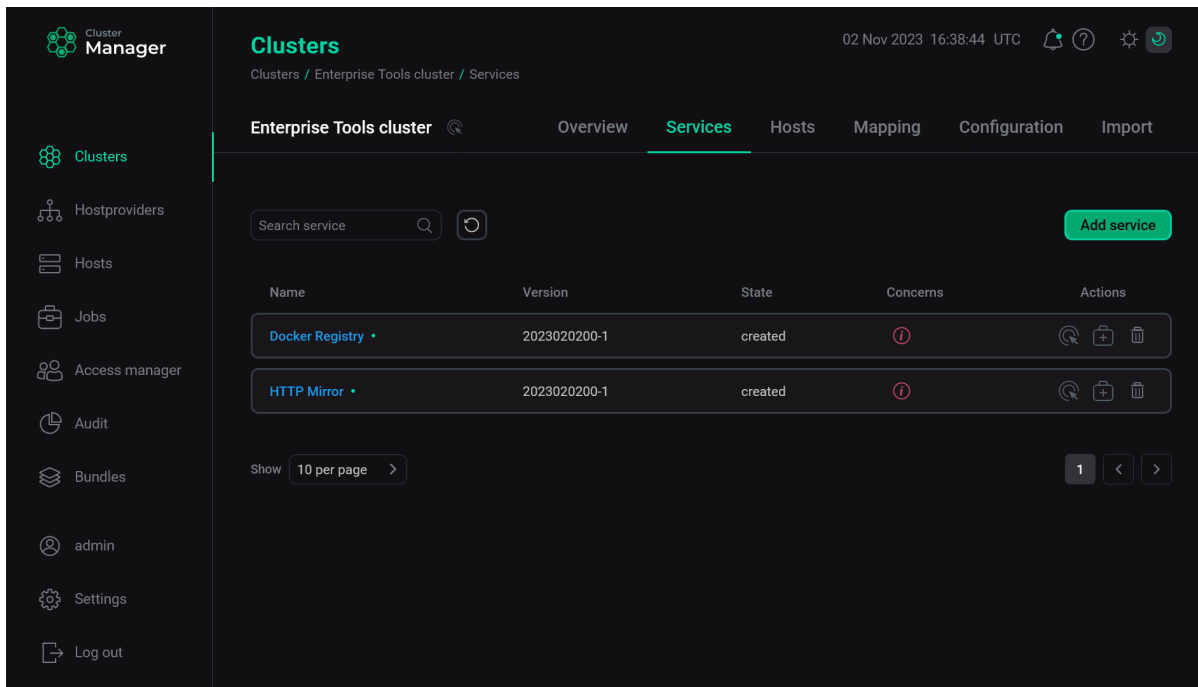
перейдите по ссылке *Terms of Agreement*. Подписать соглашение также возможно на странице *Bundles*, нажав на соответствующий продуктовый бандл.

- Нажмите *Create*.
3. После успешного добавления кластер отображается на странице *Clusters*.







2.3.2 Добавление сервисов

В ADCM сервис означает программное обеспечение, выполняющее некоторую функцию. Для добавления сервисов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. В открывшемся окне выберите сервисы для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



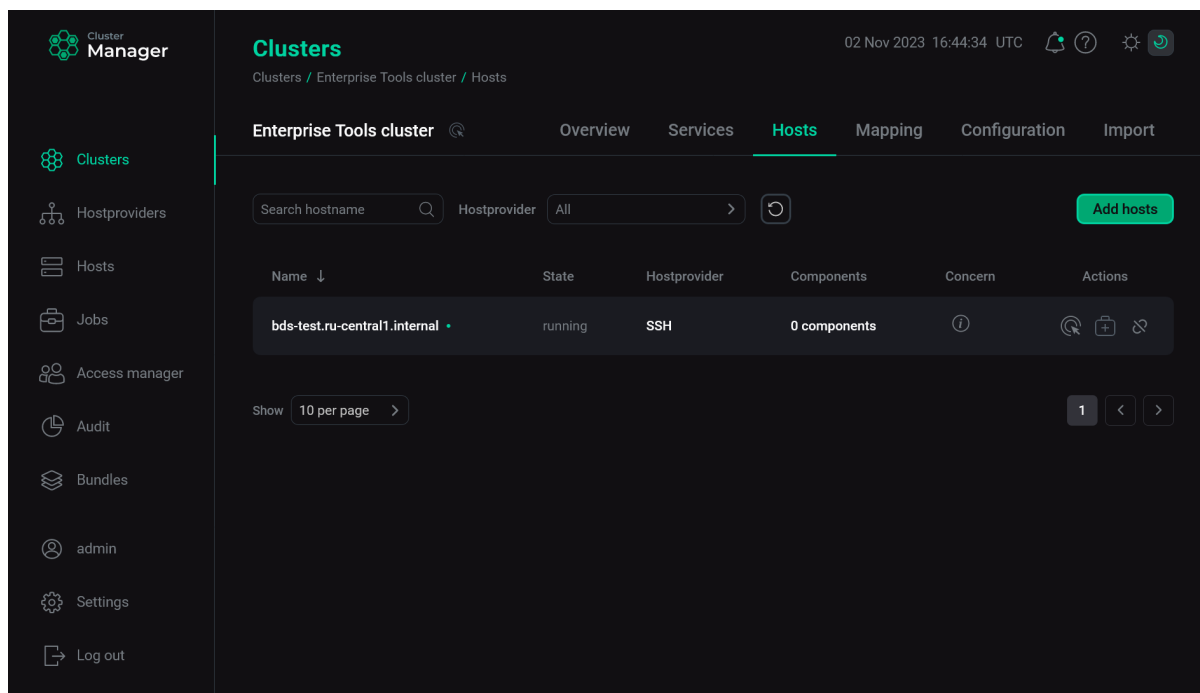
The screenshot shows the ADCM Cluster Manager interface. The left sidebar contains navigation options: Clusters, Hostproviders, Hosts, Jobs, Access manager, Audit, Bundles, admin, Settings, and Log out. The main content area is titled 'Clusters' and shows the 'Enterprise Tools cluster' selected. The 'Services' tab is active, displaying a table of services. The table has columns for Name, Version, State, Concerns, and Actions. Two services are listed: 'Docker Registry' and 'HTTP Mirror', both with version '2023020200-1' and state 'created'. Each service row has a refresh icon, an add icon, and a delete icon. A search bar and an 'Add service' button are at the top of the table. The bottom of the table shows a pagination control set to '10 per page' and a page indicator '1'.

Name	Version	State	Concerns	Actions
Docker Registry	2023020200-1	created	!	  
HTTP Mirror	2023020200-1	created	!	  

2.3.3 Добавление хостов в кластер

Для добавления хостов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Hosts* на странице кластера и нажмите *Add hosts*.
3. В открывшемся окне выберите хосты для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные хосты отображаются на вкладке *Hosts*.



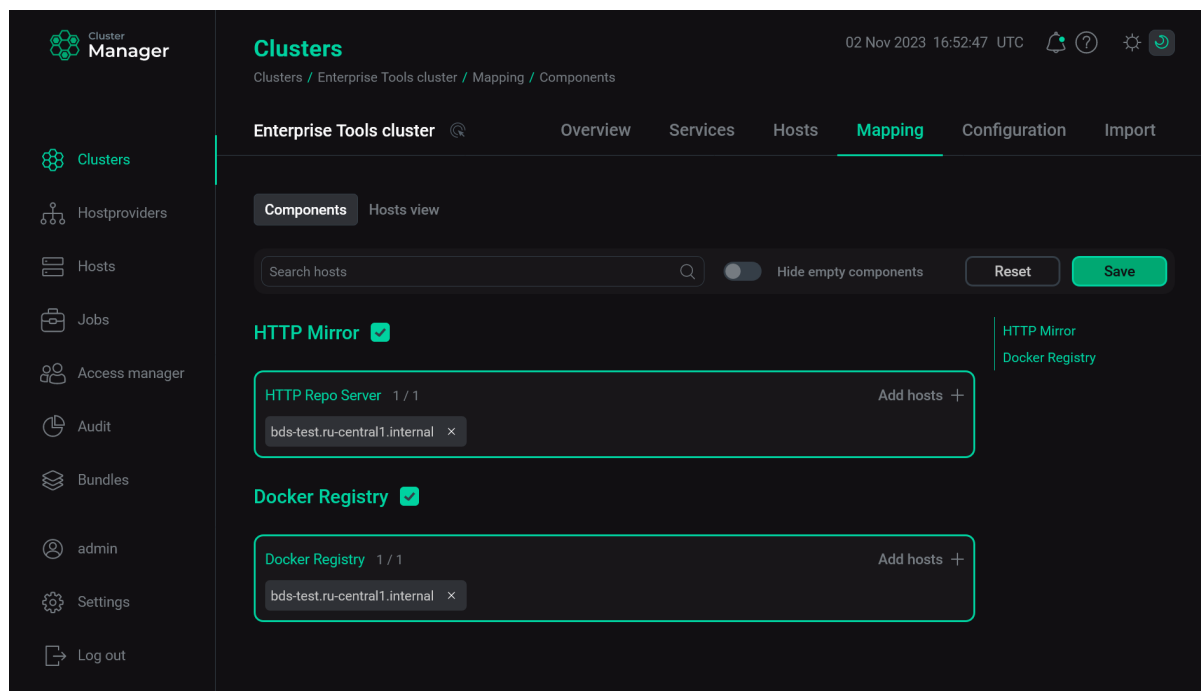
2.3.4 Добавление компонентов

В ADCM компонент означает часть сервиса, которая должна быть развернута на одном или нескольких хостах кластера.

Изначально на хостах кластера нет ни одного компонента — распределение компонентов между хостами должно быть проведено вручную.

Для добавления компонентов на хосты кластера:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера.
3. Распределите все компоненты между хостами кластера. Для назначения хостов компоненту нажмите *Add hosts* и выберите хосты в открывшемся окне.
4. После завершения распределения компонентов нажмите *Save*.



Для перехода к альтернативному варианту просмотра — с разбиением по хостам — можно открыть вкладку *Hosts view*.

2.3.5 Установка кластера

2.3.5.1 Шаг 1. Запуск установки

Для установки кластера Enterprise Tools выполните следующие шаги:

1. Запустите действие *Offline Install* для кластера, выбранного на странице *Clusters*, нажав на иконку в столбце *Actions*.
2. В открывшемся диалоговом окне введите полный путь к shell-скрипту, полученному от команды поддержки Arenadata, и нажмите *Run*.
3. Дождитесь завершения процесса установки.

2.3.5.2 Шаг 2. Загрузка пакетов для offline-установки

В дополнение к установке сервисов ET требуется загрузка всех установочных файлов и образов в инициализированный http-репозиторий и Docker Registry на хосте, где развернут кластер ET. Это важно для последующего offline-развертывания продуктов Arenadata. Для загрузки выполните шаги:

1. Запустите действие *Upload Pack* для кластера, выбранного на странице *Clusters*, нажав на иконку в столбце *Actions*.

2. В открывшемся диалоговом окне введите полный путь к shell-скрипту, полученному от команды поддержки Arenadata, и нажмите *Run*.
3. Дождитесь завершения процесса загрузки.

2.3.5.3 Шаг 3. Проверка результатов

Для проверки установки на хосте, где развернут кластер, выполните шаги:

1. Убедитесь, что порты 443 и 81 находятся в статусе LISTEN. Первый порт используется сервисом Docker Registry, второй порт — сервисом HTTP Mirror.

```
$ sudo netstat -ntlp|grep L|grep -E "81|443"
```

Результат команды:

tcp6	0	0	:::443	:::*	LISTEN	1694/docker-proxy-c
tcp6	0	0	:::81	:::*	LISTEN	1708/docker-proxy-c

2. Проверьте наличие директории `/opt/arenadata/etools/httprepo/arenadata-repo`. Она должна содержать файлы, необходимые для offline-установки продуктов, выбранных вами на этапе получения shell-скрипта от команды поддержки Arenadata.

```
$ ls -la /opt/arenadata/etools/httprepo/arenadata-repo
```

Следующий пример показывает данные, которые могут храниться в директории в случае установки ADH.

```
total 0
drwxr-xr-x 5 root root 45 Sep 16 13:35 .
drwxr-xr-x 3 root root 28 Sep 16 13:35 ..
drwxr-xr-x 4 root root 38 Sep 16 13:35 ADH
drwxr-xr-x 4 root root 28 Sep 16 13:35 ADM
drwxr-xr-x 3 root root 20 Sep 16 13:35 zookeeper
```

2.4 Установка кластера ADH T4

2.4.1 Создание кластера

В ADCM кластер означает набор сервисов, запущенных на группе хостов. Дистрибутивы кластеров для ADCM поставляются в виде бандлов. В общем случае бандл представляет обычный архив, описывающий логику развертывания кластера.

2.4.1.1 Шаг 1. Получение Enterprise-версии бандла

Пожалуйста, запросите бандл у команды поддержки Arenadata.

2.4.1.2 Шаг 2. Загрузка бандла в ADCM

Чтобы загрузить бандл в ADCM:

1. Выберите пункт *Bundles* в левом навигационном меню и нажмите *Upload bundle*.
2. Выберите бандл в открывшемся диалоговом окне.
3. После успешной загрузки бандл отображается на странице *Bundles*.

2.4.1.3 Шаг 3. Создание кластера на базе загруженного бандла

На этапе создания кластера через web-интерфейс ADCM генерируется только сам кластер. Это означает добавление записи о кластере в ADCM — сервисы на этом шаге еще не устанавливаются.

Для создания кластера необходимо выполнить шаги:

1. Выберите пункт *Clusters* в левом навигационном меню и нажмите *Create cluster*.
2. В открывшемся окне требуется описать новый кластер:
 - Выберите загруженный бандл в поле *Product*.
 - Выберите версию бандла в поле *Product version*. Несколько версий становятся доступны после загрузки нескольких бандлов для одного типа продуктов.
 - Введите имя кластера в поле *Cluster name*.
 - При необходимости добавьте описание кластера в поле *Description*.
 - Установите флаг *I accept Terms of Agreement* для подписания лицензионного соглашения. Флаг отображается только для бандлов версии Enterprise. Чтобы ознакомиться с текстом соглашения, перейдите по ссылке *Terms of Agreement*. Подписать соглашение также возможно на странице *Bundles*, нажав на соответствующий продуктовый бандл.
 - Нажмите *Create*.
3. После успешного добавления кластер отображается на странице *Clusters*.

2.4.2 Добавление сервисов

В ADCM сервис означает программное обеспечение, выполняющее некоторую функцию. Примеры сервисов в кластерах ADH: HDFS, HBase, Hive и другие. Для добавления сервисов в кластер:

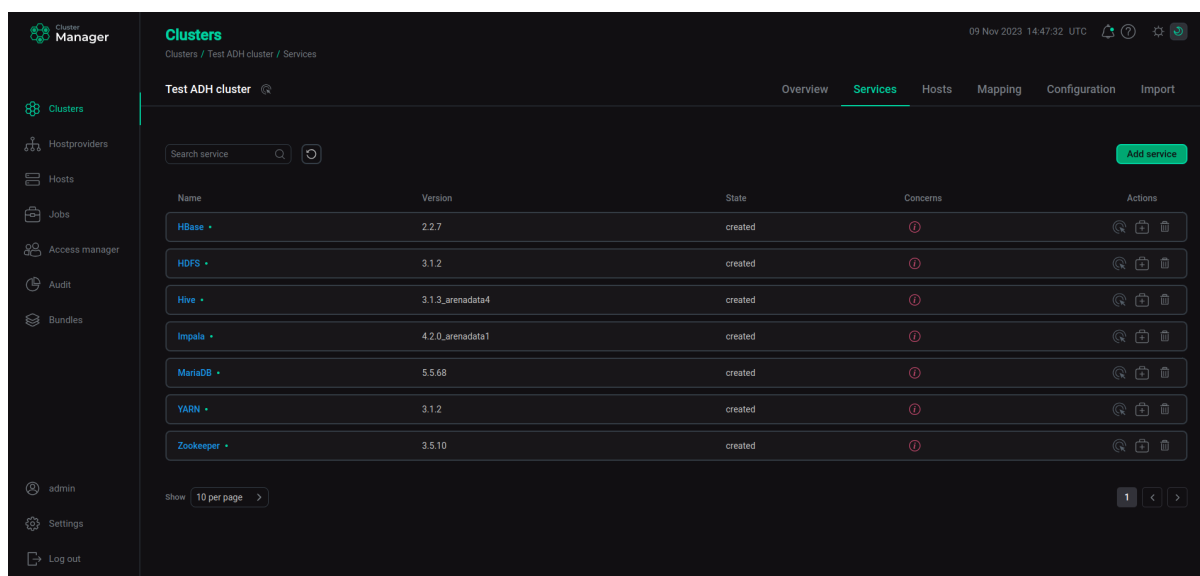
1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. В открывшемся окне выберите сервисы для добавления в кластер и нажмите *Add*.

Минимальный набор сервисов, рекомендуемый для кластеров ADH, приведен ниже:

- HDFS;
- YARN;
- Zookeeper (опционально для Community-версии ADH).

Эти сервисы составляют ядро Hadoop и являются достаточными для организации распределенного хранения и обработки данных. Полный перечень сервисов будет зависеть от требований конкретного проекта.

4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



Name	Version	State	Concerns	Actions
HBase	2.2.7	created		
HDFS	3.1.2	created		
Hive	3.13_arenadata4	created		
Impala	4.2.0_arenadsta1	created		
MariaDB	5.5.68	created		
YARN	3.1.2	created		
Zookeeper	3.5.10	created		

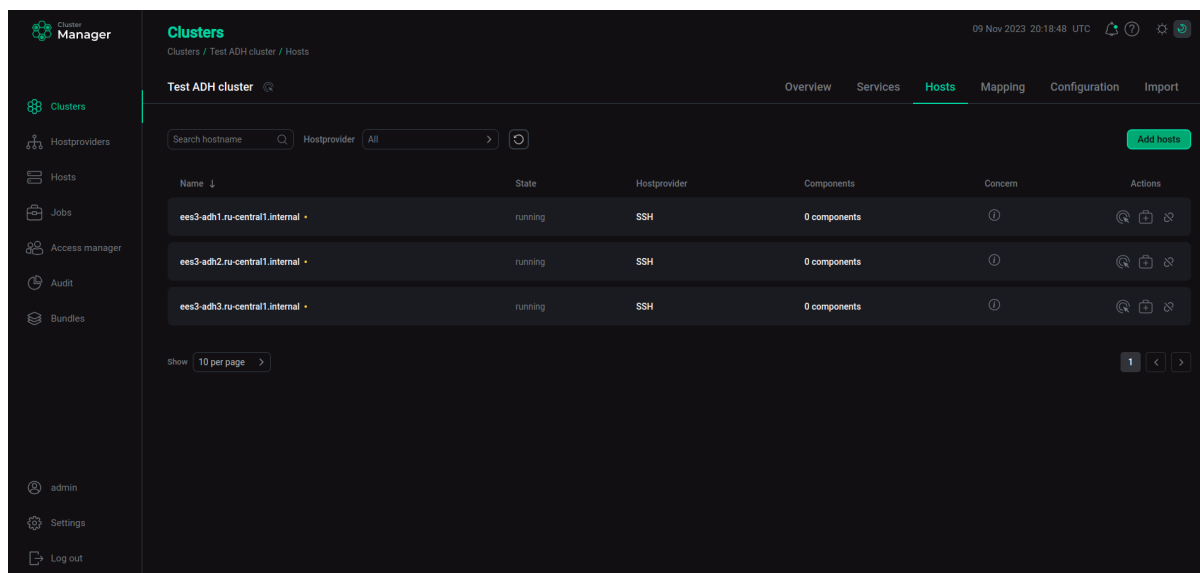
Оptionальные сервисы могут быть добавлены в кластер позднее. Процесс добавления сервисов в уже развернутый кластер не отличается от описанного выше.

2.4.3 Добавление хостов в кластер

Для добавления хостов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.

2. Откройте вкладку *Hosts* на странице кластера и нажмите *Add hosts*.
3. В открывшемся окне выберите хосты для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные хосты отображаются на вкладке *Hosts*.



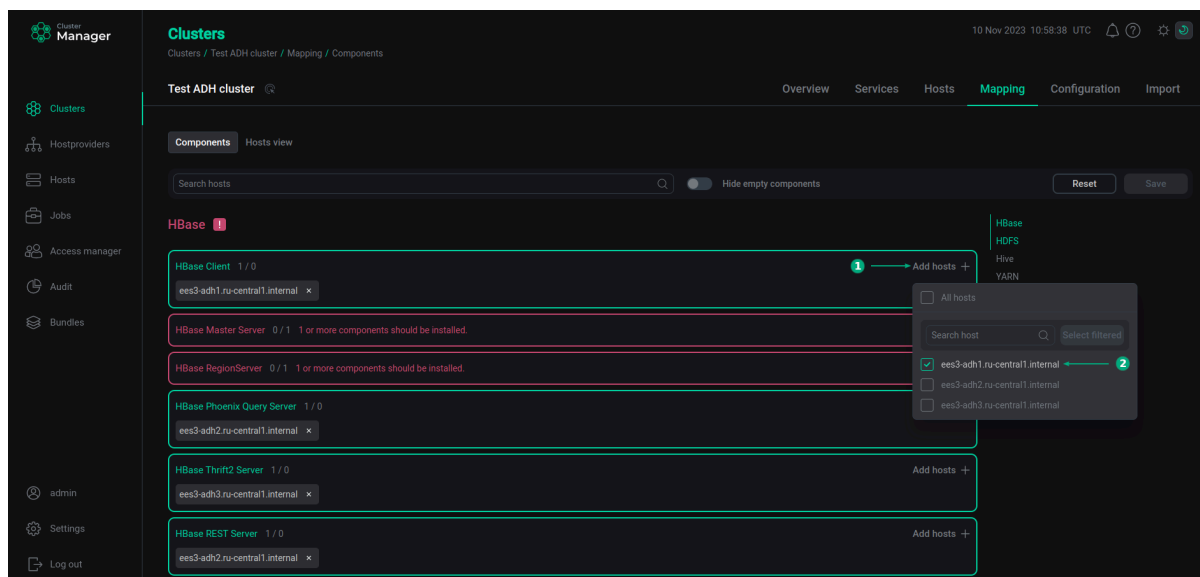
2.4.4 Добавление компонентов

В ADCM компонент означает часть сервиса, которая должна быть развернута на одном или нескольких хостах кластера. Например, сервис HDFS состоит из таких компонентов, как DataNode, NameNode, Client и другие.

Изначально на хостах кластера нет ни одного компонента — распределение компонентов между хостами должно быть проведено вручную.

Для добавления компонентов на хосты кластера:

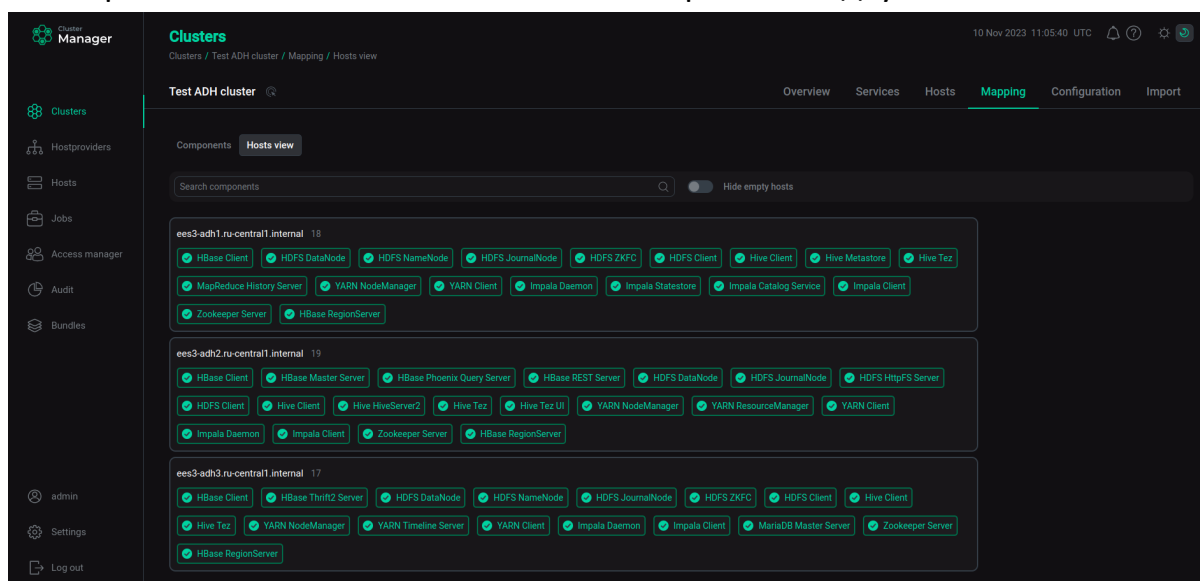
1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера.
3. Распределите все компоненты между хостами кластера. Для назначения хостов компоненту нажмите *Add hosts* и выберите хосты в открывшемся окне.



Каждый сервис может иметь обязательные и опциональные компоненты. Обязательные подсвечиваются красным цветом. Также в интерфейсе ADCM для них выводится информация о том, в каком количестве необходимо установить каждый компонент.

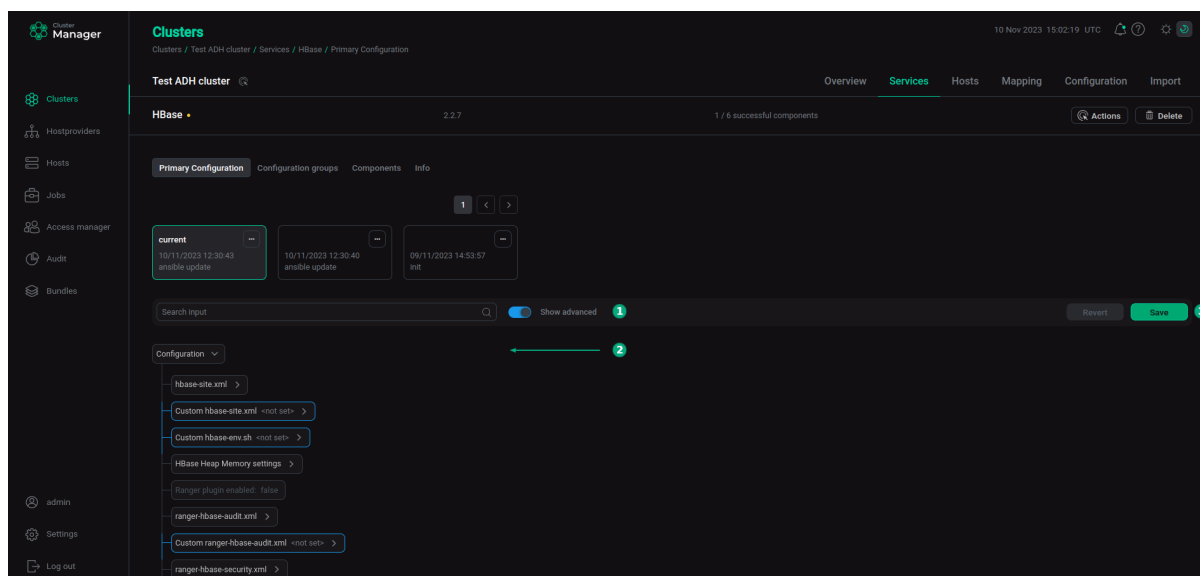
Без назначения нужного числа хостов обязательным компонентам сохранение общей схемы распределения невозможно.

4. После завершения распределения компонентов нажмите *Save*.
5. Для перехода к альтернативному варианту просмотра — с разбиением по хостам — можно открыть вкладку *Hosts view*.



2.4.5 Настройка сервисов

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера. Нажмите на имя сервиса в столбце *Name*.
3. В открывшемся окне:
 1. Откройте вкладку *Primary Configuration*.
 2. Переведите в активное состояние переключатель *Show advanced*.
 3. Заполните конфигурационные параметры выбранного сервиса. Поля, подсвеченные красным, обязательны для заполнения.
4. Нажмите *Save*.



2.4.6 Настройка кластера

В большинстве случаев кастомная настройка кластера не требуется — значения параметров можно оставить по умолчанию.

После добавления нового кластера можно выполнить его настройку. Для этого выполните шаги:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Configuration* на странице кластера и переведите в активное состояние переключатель *Show advanced*.
3. Заполните все необходимые конфигурационные параметры и нажмите *Save*.

Вкладка *Configuration* для кластера включает следующие параметры:

- Раздел *Reliability control*. Отображается, если установлен флажок *Show advanced*. Он содержит следующие настройки:
 - *retries* — количество попыток для операций в кластере (по умолчанию 20);
 - *delay* — задержка между попытками в секундах (по умолчанию 10);
 - *timeout* — тайм-аут проверки сетевых портов в секундах (по умолчанию 60).
- Пути к репозиториям CentOS, Red Hat и ALT Linux, которые будут использованы в процессе установки в зависимости от операционной системы кластера:
 - *monitoring* — репозиторий для установки мониторинга;
 - *zookeeper_repo* — YUM-репозиторий Arenadata Zookeeper;
 - *arenadata* — YUM-репозиторий Arenadata;
 - *arenadata_postgres* — YUM-репозиторий Arenadata PostgreSQL;
 - *ranger* — репозиторий для установки Ranger client.
- Параметр *DockerRegistry*, указывающий адрес Docker Registry. Значение по умолчанию — *hub.adsw.io* — Docker Registry из репозитория Arenadata.
- Раздел *Kerberos*, содержащий параметры, перечисленные в таблице на странице https://docs.arenadata.io/ru/ADH/current/get-started/offline_install/adh-install/config-cluster.html. Некоторые параметры становятся видимыми только при установленном флажке *Show advanced*.
- Параметр *java.io.tmpdir*, указывающий переменную среды для управления *java.io.tmpdir* для YARN, HBase, Hive, Spark.

2.4.7 Импорт настроек ET

Этот шаг необходим только в случае offline-установки.

Для установки кластера в режиме offline требуется интегрировать его с ранее созданным кластером Enterprise Tools (ET):

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Import* на странице кластера.

3. Выберите сервисы *Docker Registry* и *HTTP Mirror* предварительно созданного кластера ET.

4. Нажмите *Import*.

После успешного импорта данные, хранимые локально в кластере Enterprise Tools, могут быть использованы на этапе установки продуктового кластера. Таким образом, их загрузка через Интернет не требуется.

2.4.8 Установка кластера

2.4.8.1 Шаг 1. Запуск установки

Существует два способа установки:

1. Установка всего кластера. В этом методе все сервисы кластера устанавливаются автоматически друг за другом после применения действия *Install* к кластеру.
2. Установка отдельных сервисов. В этом методе каждый сервис устанавливается вручную путем применения к нему одного или нескольких действий. Этот способ также можно использовать для добавления новых сервисов в уже развернутый кластер.

В обоих методах каждый сервис запускается автоматически после установки.

Установка должна производиться на чистые хосты без предустановленного подобного ПО.

Установка всего кластера. Чтобы установить все сервисы кластера за одну операцию, выполните шаги:

1. Примените действие *Install* к кластеру, выбранному на странице *Clusters*, нажав на иконку в столбце *Actions*.
2. Подтвердите действие в открывшемся окне.
3. Дождитесь завершения установки.

Установка отдельных сервисов. Чтобы установить один или несколько сервисов вручную, необходимо предварительно добавить эти сервисы в кластер и затем применить к ним действия, описанные ниже.

Обратите внимание, что сервисы устанавливаются по очереди друг за другом. Не пытайтесь установить следующий сервис до окончания установки предыдущего.

1. Откройте вкладку *Services* на странице кластера. Для сервиса, который должен быть установлен, нажмите на иконку в столбце *Actions* и выберите действие *Install* из выпадающего списка.

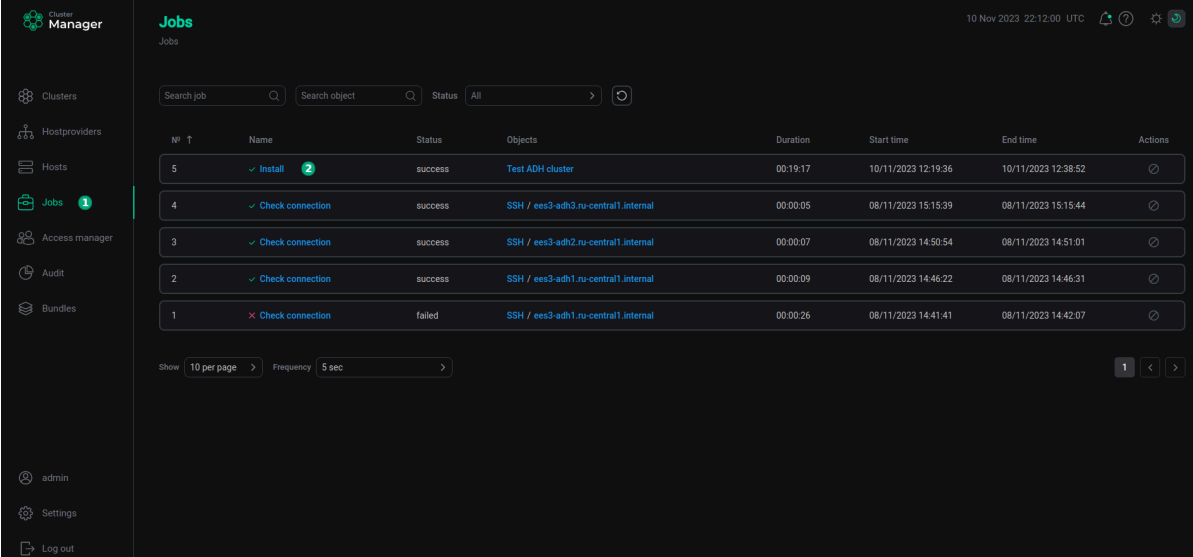
2. Подтвердите действие в открывшемся окне.
3. Дождитесь завершения установки и убедитесь, что статус сервиса изменился с `created` на `installed`.
4. Повторите предыдущие шаги для всех сервисов, которые необходимо установить.

При установке сервисов вручную важно соблюдать следующий порядок:

1. Zookeeper
2. HDFS
3. YARN
4. HBase
5. Hive
6. Spark
7. Monitoring

2.4.8.2 Шаг 2. Просмотр результатов установки

ADCM запускает отдельную задачу (job) для установки, как и для любого другого процесса. Для получения информации о статусе задач предназначена страница *Jobs*.



№	Name	Status	Objects	Duration	Start time	End time	Actions
5	✓ Install	success	Test ADH cluster	00:19:17	10/11/2023 12:19:36	10/11/2023 12:38:52	ⓘ
4	✓ Check connection	success	SSH / ees3-adh3.ru-central1.internal	00:00:05	08/11/2023 15:15:39	08/11/2023 15:15:44	ⓘ
3	✓ Check connection	success	SSH / ees3-adh2.ru-central1.internal	00:00:07	08/11/2023 14:50:54	08/11/2023 14:51:01	ⓘ
2	✓ Check connection	success	SSH / ees3-adh1.ru-central1.internal	00:00:09	08/11/2023 14:46:22	08/11/2023 14:46:31	ⓘ
1	✗ Check connection	failed	SSH / ees3-adh1.ru-central1.internal	00:00:26	08/11/2023 14:41:41	08/11/2023 14:42:07	ⓘ

Если установка завершается успешно, соответствующая ей задача на странице *Jobs* переходит из статуса `running` в `success`; в противном случае — в статус `failed`. Чтобы получить более детальную информацию о возникших ошибках, можно перейти на страницу задачи, кликнув по ней в списке задач на странице *Jobs*. Откроется окно с шагами выполнения

задачи. Кликните многоточие в столбце *Details*, чтобы получить дополнительную информацию об определенном шаге.

Страница задачи содержит два основных раздела: *ansible [stdout]* и *ansible [stderr]*. Это технические логи, содержащие информацию из стандартных I/O-поток stdout/stderr.

В дополнение к ним может выводиться третий опциональный раздел *ansible [check]*. Он содержит краткие описания наиболее часто возникающих ошибок. По сравнению с предыдущими разделами, эти описания, как правило, более простые и конкретные.

При возникновении ошибок в процессе установки рекомендуется изучить логи всех трех разделов.

The screenshot shows the 'Jobs' page in the ARENADATA interface. The 'Install' job for 'Test ADH cluster' is selected. The log output is displayed under the 'ansible [check]' tab, showing a list of successful pre-install checks. The checks include verifying hostnames, Zookeeper presence, Java tmp dir permissions, and HDFS/ZKFC configurations. A 'Download' button is visible at the bottom of the log list.

Object	Duration	Start time	Finish time	Logs
Test ADH cluster	00:19:17	10/11/2023 12:19:36	10/11/2023 12:38:52	Download

Job	Status	Duration	Start time	Finish time	Actions	Details
Pre-install check	success	00:00:36	10/11/2023 12:19:37	10/11/2023 12:20:13		

Log entries under 'ansible [check]':

- success Log [check]
- success Hostname match the specified name in adcm
- success HBase.main: Check presence of internal or external Zookeeper @ ees3-adh1.ru-central1.internal
- success HBase.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh1.ru-central1.internal
- success HBase.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh2.ru-central1.internal
- success HBase.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh3.ru-central1.internal
- success HDFS.main: Check presence of internal or external Zookeeper @ ees3-adh1.ru-central1.internal
- success HDFS.ZKFC: Check colocation of NameNodes and ZKFC @ ees3-adh1.ru-central1.internal
- success HDFS.NameNode: Check if /srv/hadoop-hdfs/name directory is empty or non-existent @ ees3-adh1.ru-central1.internal
- success HDFS.NameNode: Check if /srv/hadoop-hdfs/name directory is empty or non-existent @ ees3-adh3.ru-central1.internal
- success Yarn.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh1.ru-central1.internal
- success Yarn.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh2.ru-central1.internal
- success Yarn.main: Check that java tmp dir (/var/tmp) not mount as noexec @ ees3-adh3.ru-central1.internal
- success YARN.main: Enable HA ResourceManager @ ees3-adh2.ru-central1.internal

2.4.8.3 Шаг 3. Проверка статуса кластера после установки

В результате успешной установки происходит обновление статуса кластера и его сервисов:

- Статус кластера, отображаемый в столбце *State* на странице *Clusters*, меняет свое значение с `created` на `installed`.
- Статус сервисов, отображаемый на вкладке *Services* страницы кластера, меняет свое значение с `created` на `installed`.

2.4.8.4 Особенности установки Hive

Компонент Metastore сервиса Hive требует предварительной установки базы данных MySQL/MariaDB и наличия доступов к ней. Сервис MariaDB входит в состав бандла ADH. Таким образом, можно установить и настроить MariaDB через ADCM.

Кроме установки и настройки базы данных MariaDB, необходимо настроить доступ к ней со стороны сервиса Hive. Простейший пример предоставления доступов с использованием команды `GRANT` приведен ниже:

```
GRANT ALL ON *.* TO 'root'@'<metastore_fqdn>' identified by '<password>' WITH GRANT OPTION;
```

— где `<metastore_fqdn>` и `<password>` должны быть заменены соответствующими данными сервера MariaDB (FQDN и пароль).

Настроить адрес и учетные данные для подключения к базе данных можно на этапе конфигурирования сервиса Hive.

2.5 Установка мониторинга

Мониторинг является опциональным компонентом (однако, рекомендуемым). ADCM поддерживает настройку мониторинга на базе трех сервисов: Diamond, Graphite и Grafana.

Для процесса offline-установки необходимости в отдельном кластере мониторинга нет, так как его функции выполняет кластер Enterprise Tools (ET). Шаги для установки мониторинга в этом случае приведены ниже. Их можно выполнить до и после установки кластера ADH, кроме последнего шага Интеграция с кластером ADH, который может быть выполнен только на базе добавленного кластера ADH.

- Настройка кластера Enterprise Tools (пп. 2.5.1);
- Интеграция с кластером ADH (пп. 2.5.2).

2.5.1 Настройка кластера Enterprise Tools

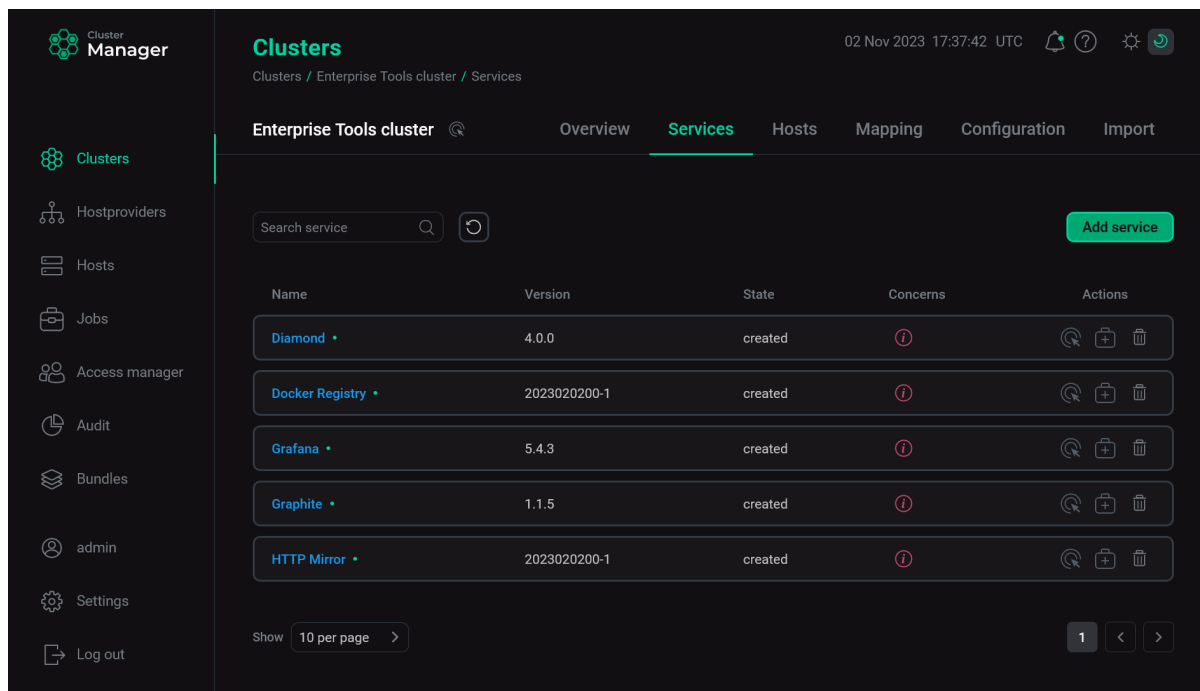
Перед настройкой мониторинга убедитесь, что кластер Enterprise Tools (ET) установлен в ADCM. Все последующие действия выполняются в нем.

2.5.1.1 Шаг 1. Добавление сервисов мониторинга в кластер ET

Для добавления сервисов в кластер:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.

2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. В открывшемся окне выберите сервисы для добавления в кластер и нажмите *Add*.
4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



Cluster Manager

Clusters

Enterprise Tools cluster

Overview Services Hosts Mapping Configuration Import

Search service

Add service

Name	Version	State	Concerns	Actions
Diamond	4.0.0	created		
Docker Registry	2023020200-1	created		
Grafana	5.4.3	created		
Graphite	1.1.5	created		
HTTP Mirror	2023020200-1	created		

Show 10 per page

1 < >

2.5.1.2 Шаг 2. Распределение компонентов мониторинга

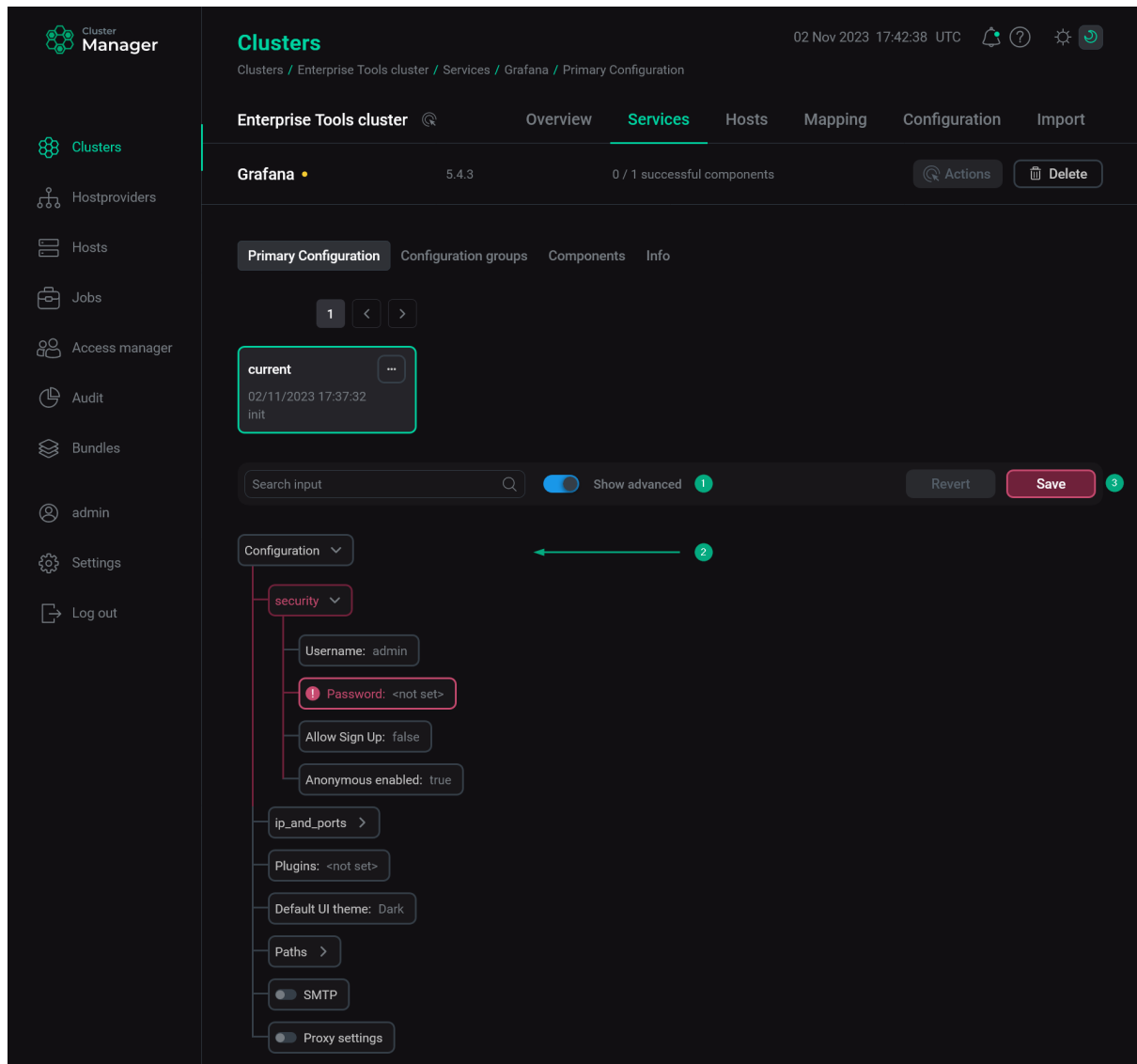
Для добавления компонентов на хосты кластера:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера, добавьте компоненты мониторинга на все хосты кластера и нажмите *Save*.

The screenshot shows the 'Clusters' management interface. The left sidebar contains navigation options: Clusters, Hostproviders, Hosts, Jobs, Access manager, Audit, Bundles, admin, Settings, and Log out. The main content area is titled 'Enterprise Tools cluster' and is currently on the 'Mapping' tab. It shows a list of services with their status (checked) and a search bar. The services listed are HTTP Mirror, Docker Registry, Grafana, Graphite, and Diamond. Each service has a configuration card with a search bar and an 'Add hosts' button. A red circle '2' is placed over the 'Add hosts' button of the Graphite service, with arrows pointing to it from the right.

2.5.1.3 Шаг 3. Настройка сервисов мониторинга

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера. Нажмите на имя сервиса в столбце *Name*.
3. В открывшемся окне:
 1. Откройте вкладку *Primary Configuration*.
 2. Переведите в активное состояние переключатель *Show advanced*.
 3. Заполните конфигурационные параметры выбранного сервиса. Поля, подсвеченные красным, обязательны для заполнения.
 4. Нажмите *Save*.



2.5.1.4 Шаг 4. Установка сервисов мониторинга

1. Запустите действие *Install Monitoring* для кластера, выбранного на странице *Clusters*, нажав на иконку в столбце *Actions*.
2. Нажмите *Run* в открывшемся окне.
3. Дождитесь завершения процесса установки.

2.5.1.5 Шаг 5. Проверка результатов

Graphite

Чтобы проверить доступность web-интерфейса Graphite, необходимо выполнить шаги:

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите

значение параметра *Web-interface TCP port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Graphite при установке кластера мониторинга. Значение 80 (используемое по умолчанию) можно не указывать.

В случае offline-установки этот адрес совпадает с адресом хоста, на котором развернут кластер Enterprise Tools.

2. Разверните дерево в левой части открывшейся формы. Здесь отображаются метрики System метрики с группировкой по хостам. Можно выбрать любую из метрик для просмотра динамики ее изменения в виде графика в правой части экрана.

Grafana

Чтобы проверить доступность web-интерфейса Grafana, необходимо выполнить шаги:

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Grafana при установке кластера мониторинга (по умолчанию 3000).
2. Выберите пункт меню *Home* в открывшемся окне.
3. Откройте любой из доступных дашбордов (dashboards). Например, *Arenadata System metrics*.
4. Проанализируйте доступные графики.

2.5.2 Интеграция с кластером ADH

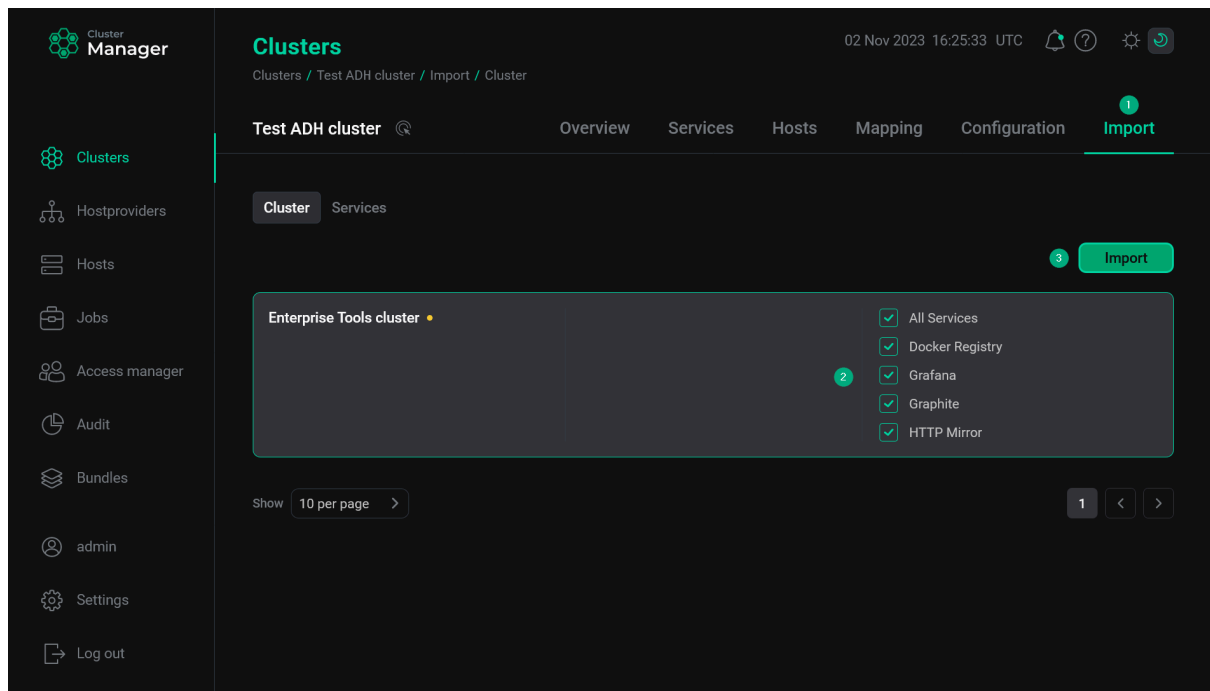
Для настройки мониторинга нового кластера необходимо интегрировать его с предварительно созданным кластером мониторинга и добавить отдельный сервис Monitoring .

Шаги для интеграции кластера мониторинга и кластера ADH описаны ниже. Все действия выполняются на базе кластера ADH.

2.5.2.1 Шаг 1. Импорт настроек мониторинга

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Import* на странице кластера.
3. Выберите сервисы *Graphite* и *Grafana* ранее добавленного кластера мониторинга.

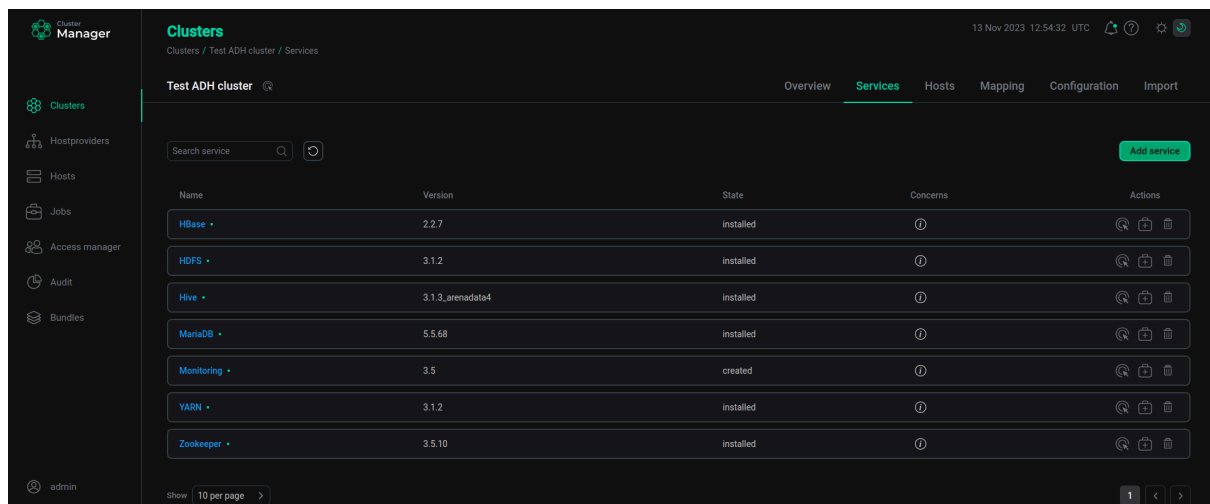
4. Нажмите *Import*.



2.5.2.2 Шаг 2. Добавление сервиса мониторинга

Для добавления сервисов в кластер:

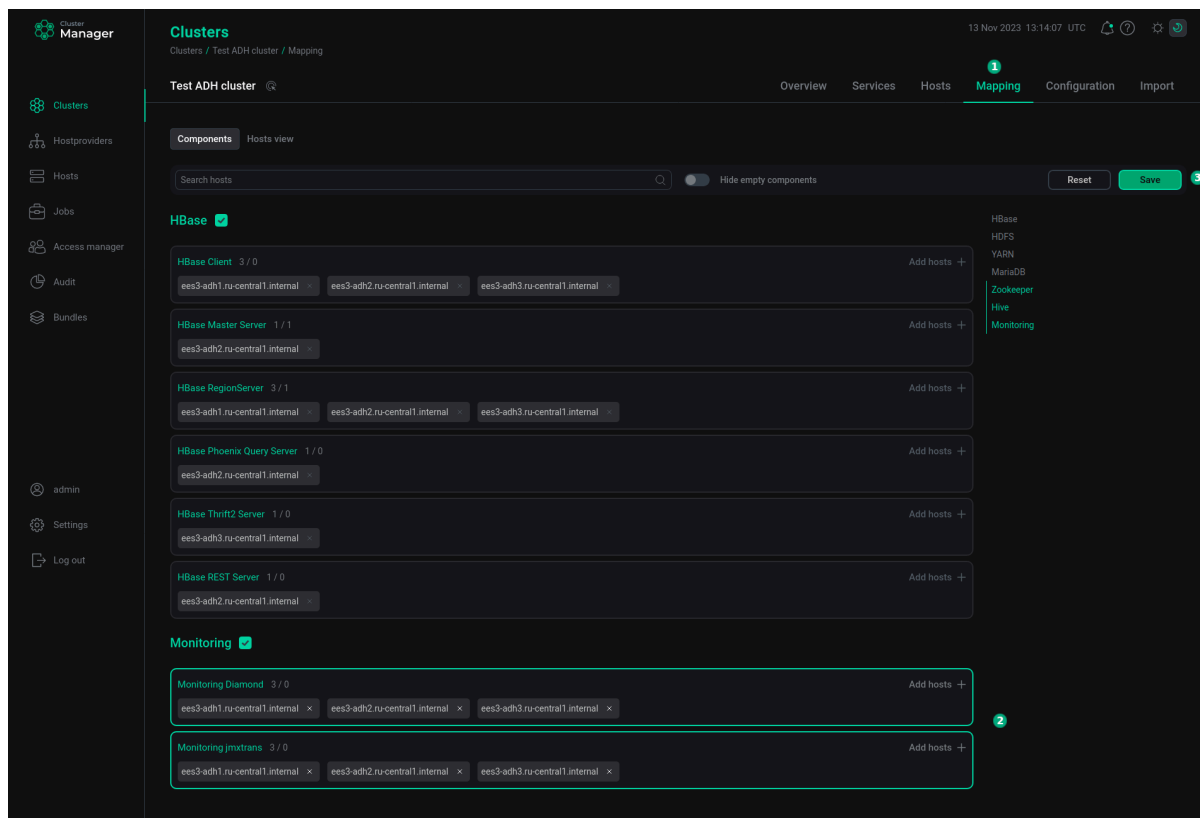
1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Services* на странице кластера и нажмите *Add service*.
3. Выберите сервис *Monitoring* в открывшемся окне и нажмите *Add*.
4. В результате успешно добавленные сервисы отображаются на вкладке *Services*.



2.5.2.3 Шаг 3. Распределение компонентов мониторинга

Для добавления компонентов на hosts кластера:

1. Выберите кластер на странице *Clusters*. Для этого нажмите на имя кластера в столбце *Name*.
2. Откройте вкладку *Mapping* на странице кластера, добавьте компоненты мониторинга на все hosts кластера и нажмите *Save*.



2.5.2.4 Шаг 4. Установка сервиса мониторинга

1. Откройте вкладку *Services* на странице кластера. Нажмите на иконку в столбце *Actions* и выберите действие *Install* в строке, содержащей сервис *Monitoring*; подтвердите действие в открывшемся окне.
2. Дождитесь завершения процесса установки. В результате статус сервиса мониторинга изменяет свое значение с `created` на `installed`.

После установки кластера мониторинга необходимо перезапустить кластер ADH, чтобы просматривать метрики.

2.5.2.5 Шаг 5. Проверка результатов

По завершении установки необходимо убедиться в следующем:

- Системные метрики собираются с хостов вашего кластера, а не только с хоста, на котором ранее был развернут кластер мониторинга.
- Наряду с системными метриками, с хостов кластера поступают сервисные метрики. Например, в кластерах ADH поддерживаются метрики для таких сервисов, как HDFS, YARN, HBase, Spark.

Оба пункта можно проверить в web-интерфейсе Graphite и Grafana.

Graphite

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Web-interface TCP port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Graphite при установке кластера мониторинга. Значение 80 (используемое по умолчанию) можно не указывать.
2. Разверните дерево в левой части открывшейся формы. Здесь отображаются метрики System и сервисные метрики с группировкой по хостам. Можно выбрать любую из метрик для просмотра динамики ее изменения в виде графика в правой части экрана.

Grafana

1. В адресной строке браузера введите адрес хоста, на котором развернут кластер мониторинга. В качестве номера порта укажите значение параметра *Port* из раздела *ip_and_ports*, определенное на этапе конфигурирования сервиса Grafana при установке кластера мониторинга (по умолчанию 3000).
2. Выберите пункт меню *Home* в открывшемся окне.
3. Обратите внимание, что новые дашборды доступны. Можно выбрать один из них для просмотра сервисных метрик, поступающих из вашего кластера.
4. Проанализируйте доступные графики.