Arenadata™ Cluster Manager

Версия - develop

User Guide

2023 Arenadata™

Оглавление

1 Understand the Basics **3**
- 1.1 Working with applications . . . 3
- 1.2 Current Supported Applications . . . 4
- 1.3 Supported Cloud Providers . . . 4
- 1.4 Working with infrastructure . . . 4

2 Pre-Requisites for ADCM Installation **5**
- 2.1 Requirements . . . 5
- 2.2 Install Docker . . . 5
- 2.3 Disable Selinux (optional) . . . 6

3 Install **7**

4 Start and Stop **8**

5 Backup and Restore **9**
- 5.1 Backup . . . 9
- 5.2 Restore . . . 9

6 Upgrade **10**

7 ADCM Objects **11**
- 7.1 Bundle . . . 11
- 7.2 Host . . . 11
- 7.3 Provider . . . 12
- 7.4 Service . . . 12
- 7.5 Component . . . 12
- 7.6 Cluster . . . 12

8 Quick Start **13**
- 8.1 First Bundles . . . 13

9 Maintenance mode **15**

10 Proxying **17**

11 Enable HTTPS **18**

12 RBAC **19**
- 12.1 Built-in roles . . . 19

12.2 Permissions and roles 20
12.3 Creating a role 21
12.4 Creating a policy 22

The user guide provides information for work with the ADCM platform. The guide can be useful to administrators, programmers, developers and employees of the information technology departments that support the platform.

Important: Support – e-mail: info@arenadata.io

Глава 1

Understand the Basics

Arenadata Cluster Manager (**ADCM**) is a Data Platform as a Service solution (DPaaS), that simplifies deploying and management of data services on premise, in private or public clouds. By splitting data landscape into two independent management layers – infrastructure layer and application layer – your platform becomes flexible, predictable and repeatable.

ADCM is a modular system – you can add specific modules to support new applications, containers and cloud providers. You can develop your own modules to support your specific applications and proprietary systems.

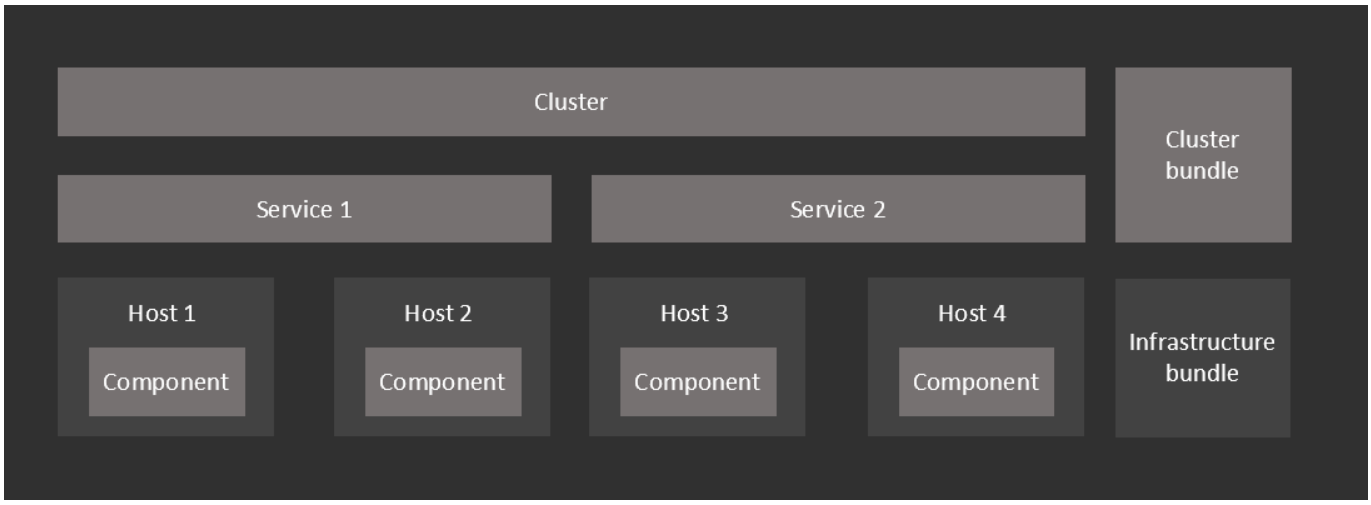


Рис.1.1.: Platform concept

ADCM comes with open powerful REST API – you can easily integrate **ADCM** with your infrastructure. Along with this, **ADCM** is run in two minutes – just pull the docker image and run the container. You can find additional information on Docker Hub.

Important: Please ensure that all pre_requisites_ are matched with your system

1.1 Working with applications

Application Functionality:

- Deploying;
- Upgrading;
- Management;
- Monitoring;
- Integration.

1.2 Current Supported Applications

Current Supported Applications are:

- Arenadata DB/Greenplum;
- Apache Hadoop;
- Apache Spark;
- Apache Hive;
- Apache HBase;
- Apache Zeppelin;
- Apache Solr;
- Apache Kafka;
- Apache NiFi;
- NiFi Registry;
- Schema Registry;
- Kafka Manager.

1.3 Supported Cloud Providers

Supported Cloud Providers are:

- Google Cloud Platform;
- Yandex Cloud;
- DataFort Cloud;
- DigitalEnergy.

1.4 Working with infrastructure

Infrastructure Functionality:

- Deploying / Removing VMs;
- Deploying physical hosts;
- Hosts preparing;
- Monitoring;
- User management;
- Access control.

Глава 2

Pre-Requisites for ADCM Installation

ADCM requires installed Docker package on your system. Currently **ADCM** Docker image tested on **CentOS/RHEL 7 OS**.

Important: Please check all requirements before Docker and ADCM installation

2.1 Requirements

2.1.1 OS requirements

To install Docker, you need a maintained version of **CentOS 7** or **Alt 8 SP**. Archived versions are not supported or tested.

The centos-extras repository must be enabled. This repository is enabled by default, but if you have disabled it, you need to re-enable it.

The **overlay2 storage driver** is recommended.

2.1.2 Hardware requirements

- CPU: 2-4 CPU Cores
- RAM: 8-16G
- HDD: 8-50G

2.2 Install Docker

Important: Do not install upstream Docker from *docker.com*. Use one from official repositories of your Linux distribution. Upstream Docker tend to be unstable and less secure, than provided by your OS developer

You can install Docker in different ways, depending on your needs.

Most users set up Docker's repositories and install from them, to ease the installation and upgrade tasks. This is the recommended approach.

Some users download the RPM-package and install it manually and manage upgrades completely manually. This is useful in situations such as installing Docker on air-gapped systems with no access to the internet.

In testing and development environments, some users choose to use automated convenience scripts to install Docker.

2.2.1 Install using the repository

Before you install Docker for the first time on a new host machine, you need to set up the Docker repository. Afterward, you can install and update Docker from the repository.

2.2.2 Install Necessary Packages

Install required packages. Yum-utils provides the **yum-config-manager** utility, and the **device-mapper-persistent-data** and **lvm2** are required by the devicemapper storage driver.

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

2.2.3 Start docker

To start Docker:

```
sudo systemctl start docker
```

2.2.4 Enable docker as system service

To enable Docker as a system service it is necessary:

```
sudo systemctl enable docker
```

2.3 Disable Selinux (optional)

It is recommended **SELinux** to be disabled for **ADCM**. To temporarily disable it, run the following command on each host in your cluster:

```
setenforce 0
```

Permanently disabling **SELinux** so that it does not restart on system reboot is strongly recommended. For achieving this:

```
vi /etc/selinux/config

# 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
```

Глава 3

Install

ADCM is a software distributed in a form of Docker image.

You have to have some Docker or compatible container engine to run **ADCM** as described in pre_requisites_.

To install software you could simply run the following command:

```
docker pull hub.arenadata.io/adcm/adcm:latest
```

Then you have to create container base on that image. **ADCM** has some mandatory persistent information which is stored in */adcm/data* dir inside container. That means that we have to map volume to that dir and care about backup of that volume:

```
docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data hub.arenadata.io/adcm/adcm:latest
```

In the command above we made a container on port 8000 with **ADCM** and all data will be stored on host machine in directory */opt/adcm/*.

If you want to run ADCM with **SELinux** enabled, a docker must be ctreated with `:Z` option in data catalog parameter:

```
docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data:Z hub.arenadata.io/adcm/adcm:latest
```

Глава 4

Start and Stop

ADCM is just a container operated by Docker. So you are able to use simple Docker commands to start and stop **ADCM**.

To start Docker you need to use the following:

```
docker start adcm
```

After starting you will be able to login to **WebUI** over browser with url `http://<ip_address_of_server>:8000`.

To stop the **ADCM** docker-container you need to issue the following:

```
docker stop adcm
```

Глава 5

Backup and Restore

5.1 Backup

To perform Backup, you need to:

1. Stop ADCM container:

```
docker stop adcm
```

2. Create Backup volume (by default /opt/adcm):

```
tar czvf ~/backup.tar.gz /opt/adcm
```

3. Start ADCM container:

```
docker start adcm
```

5.2 Restore

To perform Restore, you need to:

1. Create Restore volume (by default /opt/adcm):

```
cd /
tar xvf ~/backup.tar.gz
```

2. Install ADCM:

```
docker pull hub.arenadata.io/adcm/adcm:latest
```

3. Start ADCM:

```
docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data hub.arenadata.io/adcm/adcm:latest
```

Глава 6

Upgrade

To upgrade **ADCM** you have to do the following:

1. Pull new image:

```
docker pull hub.arenadata.io/adcm/adcm:latest
```

2. Stop and remove current container:

```
docker stop adcm
docker rm adcm
```

3. Create a new one in the same way it was done in install_:

```
docker create --name adcm -p 8000:8000 -v /opt/adcm:/adcm/data hub.arenadata.io/adcm/adcm:latest
```

Глава 7

ADCM Objects

Before using **ADCM** we recommend you to take a glance at ADCM's object model. At first let us show you what are the internal objects and their roles.

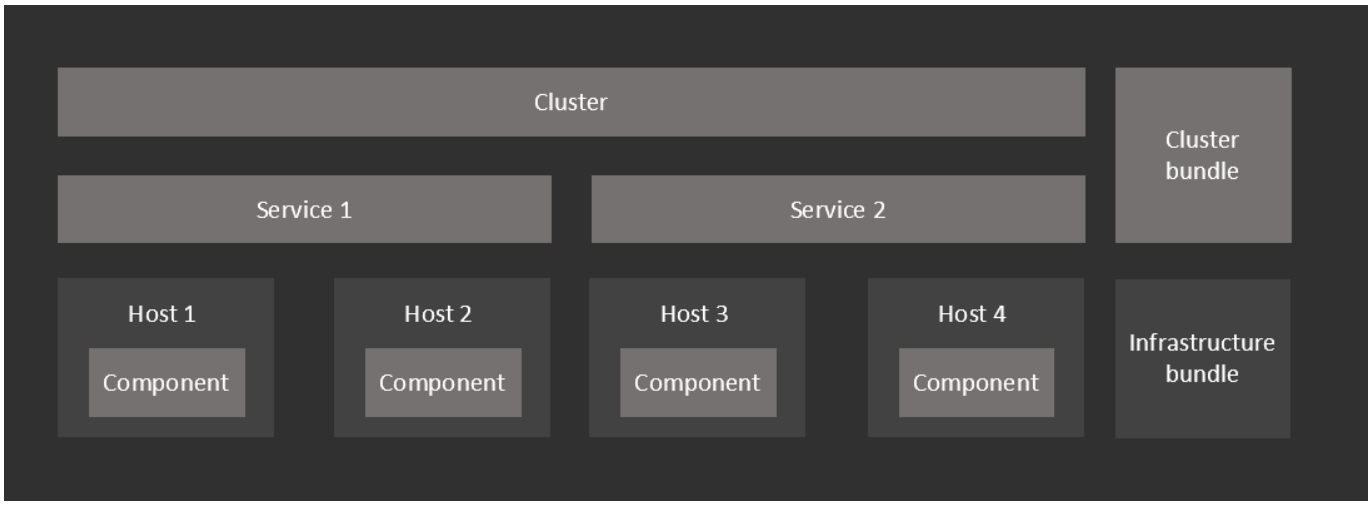


Рис.7.1.: Platform concept

7.1 Bundle

Bundle is a delivery unit in **ADCM**. New extensions and parts are coming over in bundles. In short, it is just a tarball with the information for **ADCM**, explaining how to do everything.

There are two kinds of bundle: **Cluster Bundle** and **Infrastructure Bundle**.

Cluster Bundle allows you to install *Cluster* (take a look at a definition bellow), while Infrastructure Bundle allows you to add support of different host providers (like **Google Cloud**, **Yandex Cloud** or regular ssh host).

7.2 Host

It is a representation of real hardware host or virtual host. It doesn't matter whether it comes in (a cloud or bare-metal) but it should be accessible over ssh. Every host could be added to created cluster only.

7.3 Provider

Provider is a complicated idea. Provider is something that produces a host. For example you have a **Google Cloud** account. In that case you create a **GCE Provider**, place credentials to it, and it allows you to create a host in **GCE**. Or create a lot of hosts and all of them will have the same Provider and consequently have the same credentials.

7.4 Service

Service is a software that does some function. For example **Hadoop** is a service, and **Hive** is a service, as well as **Postgres** and so on.

7.5 Component

Component is a part of the service. For example **Hadoop** service has DataNode component and ServiceNode component. It is a way of service decomposition which is native for complicated cluster software.

7.6 Cluster

In short, cluster is a set of services running on the set of hosts. With **ADCM** you are able to create a lot of clusters with different configuration and distribution of services over clusters.

But one cluster allows restricted set of service kinds only. So if you create a cluster **Hadoop** from bundle *Hadoop*, there is no way to install services from cluster **ADB** from bundle *ADB*.

Глава 8

Quick Start

By default **ADCM** start on port `8000`, so you just have to know the ip-address of a server with installed **ADCM** to start using it.

Open your browser and go to `http://<ip_of_your_server>:8000`. Default user is *admin* and password is *admin* (Рис.8.1.).

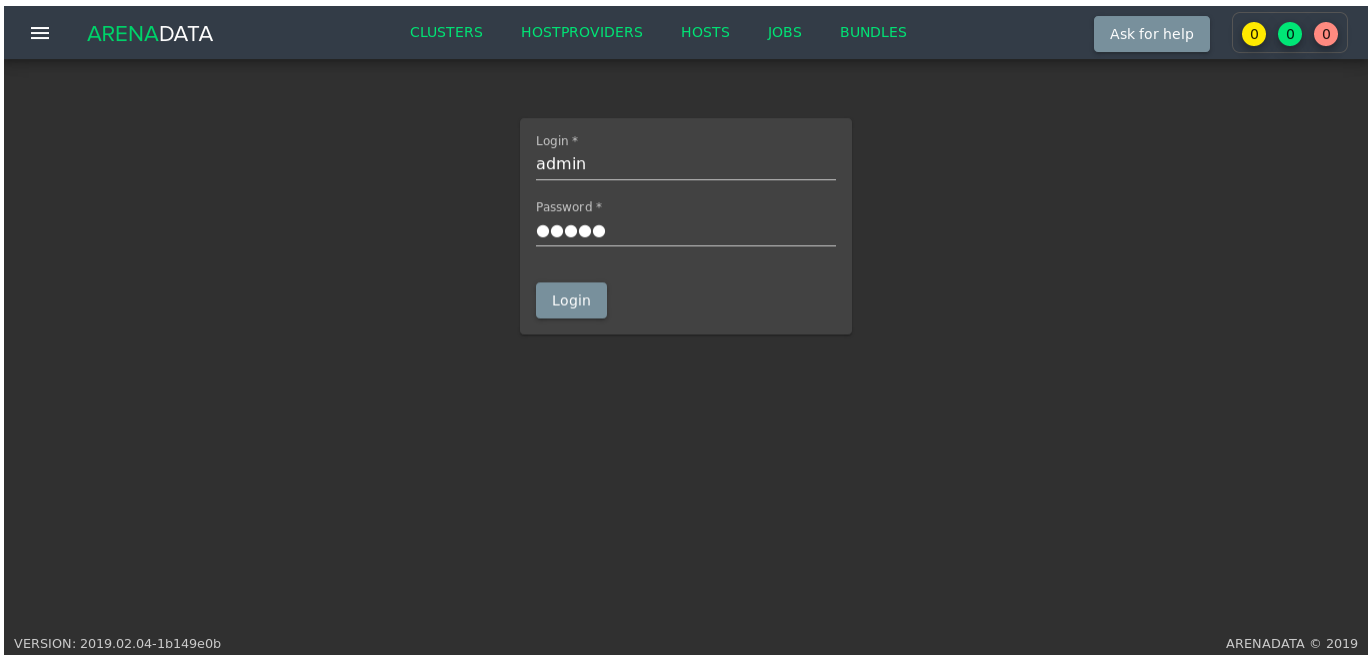


Рис.8.1.: Authorization

8.1 First Bundles

To continue consider to upload some of infrastructure bundles go to store.arenadata.io and download some of them.

Then click a “BUNDLES” in the head menu of **ADCM** and upload bundles with the “Upload bundle” button (Рис.8.2.).

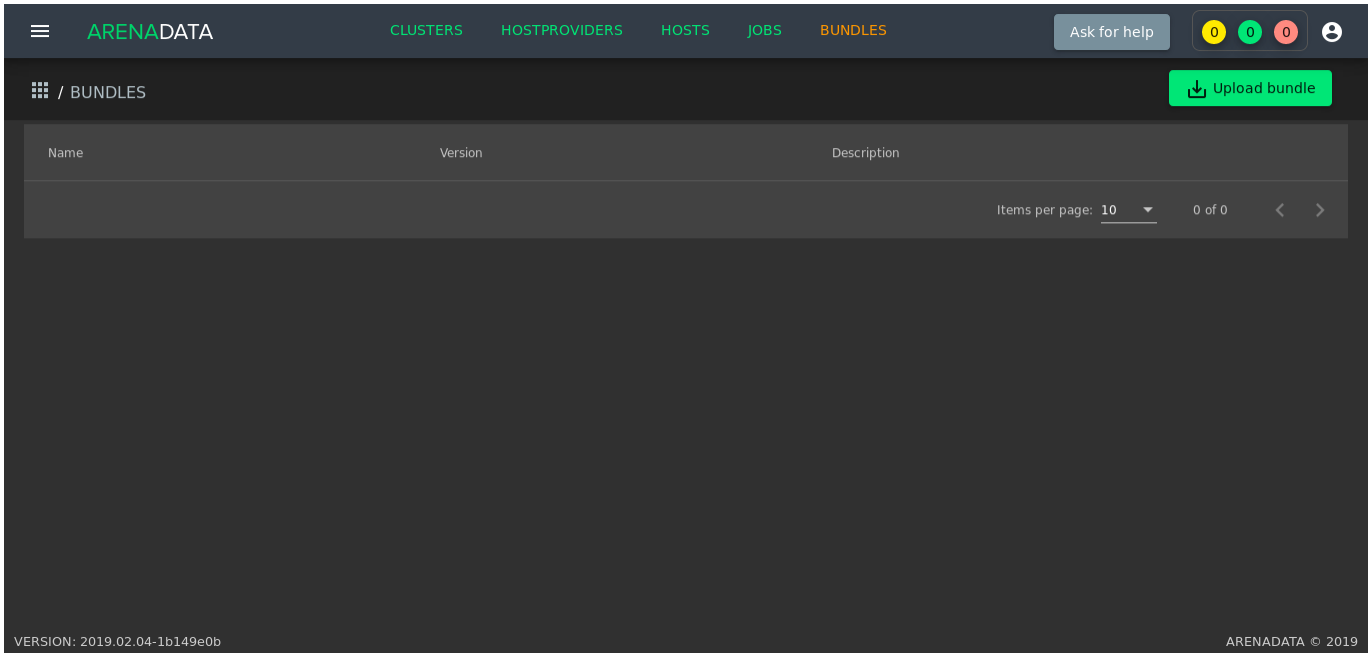


Рис.8.2.: Upload bundle

Глава 9

Maintenance mode

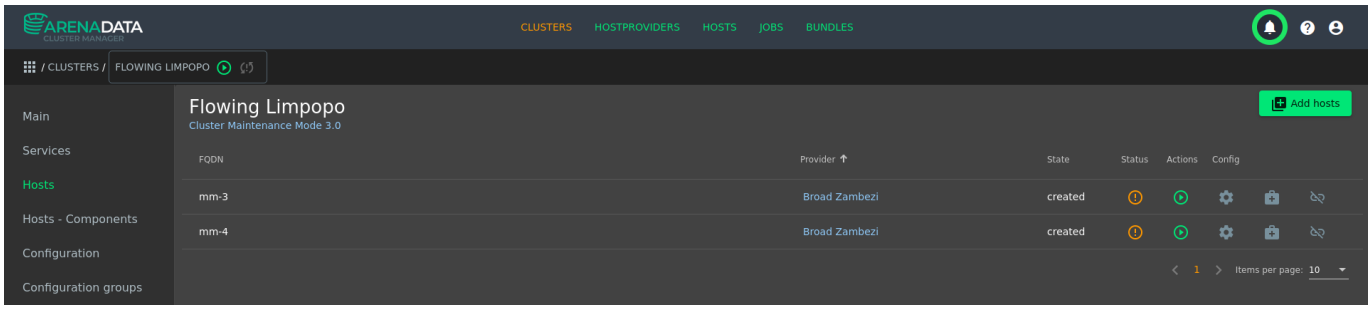
A maintenance mode for hosts is implemented in ADCM. This mode is designed to stop the implementation of any cluster- or service-related operations for a host. You also can't add any components to the host that is in maintenance mode. This is useful if the host is turned off, has malfunctioned, or is not ready to participate in ADCM actions for some reason.

Consider using maintenance mode when you perform hardware or software maintenance, change configuration settings, perform troubleshooting, decommissioning, or removing cluster nodes.

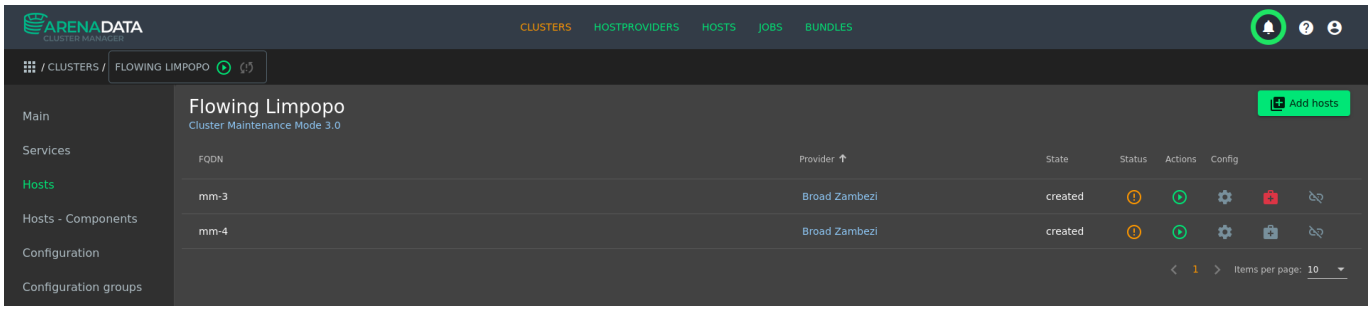
Important: Not every cluster, service, or component action is possible if one of cluster hosts is in maintenance mode. This depends on if the action affects all cluster hosts, as well as on action's logic.

To switch a host to maintenance mode, you should perform the following steps:

1. Locate the host you wish to switch to maintenace mode in ADCM interface. For instance, you can use the **Hosts** subsection of the **Clusters** tab.



2. Click the maintenance mode icon. The icon turns red. The host is now in maintenance mode.



3. To switch off the maintenace mode, click the maintenance mode icon again.

While you can't perform any operations with components on a host that is in maintenance mode, you can still remove a component from this host on the logical level. Generally speaking, the maintenance mode prevents bulk operations being performed on a host, yet you may explicitly start and stop a component on a host while in maintenance mode.

Глава 10

Proxying

A reverse proxy server is a type of proxy server that typically sits behind the firewall in a private network and directs client requests to the appropriate backend server. A reverse proxy provides an additional level of abstraction and control to ensure the smooth flow of network traffic between clients and servers.

You may want to use reverse proxy for TLS-termination, cache and compress content or some other purpose.

Here is an example of Nginx configuration:

```
server {
  listen 443 ssl;
  server_name <your_adcm_domain>;
  ssl_certificate      example.crt;
  ssl_certificate_key  example.key;
  ssl_protocols        TLSv1 TLSv1.1 TLSv1.2;
  ssl_ciphers          HIGH:!aNULL:!MD5;

  access_log /var/log/nginx/adcm.access.log;
  client_max_body_size 60M;  # used to increase size of an uploaded file. Size by default is 1MB

  location / {
    proxy_pass http://<ip_adcm>:<port_adcm>;
    proxy_read_timeout 90;
    proxy_set_header X-Real-IP $remote_addr;
    proxy_set_header Host $host;
    proxy_set_header X-Forwarded-For $proxy_add_x_forwarded_for;
    proxy_set_header X-Forwarded-Proto https;
  }

  location /ws {  # used to proxying websocket API
    proxy_pass http://<ip_adcm>:<port_adcm>;
    proxy_http_version 1.1;
    proxy_set_header Upgrade $http_upgrade;
    proxy_set_header Connection "upgrade";
    proxy_set_header X-Real-IP $remote_addr;
    proxy_set_header Host $host;
    proxy_set_header X-Forwarded-For $proxy_add_x_forwarded_for;
  }
}
```

Глава 11

Enable HTTPS

- Prepare server private key in pem format and place it to <data_volume>/conf/ssl/key.pem
- Prepare certificate chain and place it to <data_volume>/conf/ssl/cert.pem in PEM format.
- Drop current ADCM container and create it with “ssl” port

```
docker rm adcm
docker create --name adcm -p 8000:8000 -p 8443:8443 -v /opt/adcm:/adcm/data hub.arenadata.io/adcm/
↪adcm:latest
```

Note: Order of certs in cert.pem should be in sync with Nginx documentation.

Note: Do not disable HTTP port. We still need this for interconnection between ADCM and hosts. This should expected to be fixed in future.

Глава 12

RBAC

A role model is implemented in ADCM. Specific sets of permissions for operations with various ADCM objects can be grouped into roles. The roles are used to create policies to grant roles's permissions to users or user groups.

Permission is a basic element of the role model. **Role** is a set of permissions. Permissions can not be assigned directly to the users or groups. Permissions can be only added to the roles.

Some permissions are necessary to work with any product via ADCM. Such permissions are always present in ADCM. Other permissions are necessary to work only with a particular product. Those permissions appear in ADCM when a bundle is uploaded.

The role is assigned to a user or to a user group via policy. Assignment of multiple roles is supported. The user inherits the roles of the group to which he belongs. Inheritance occurs even if the user is assigned a separate role. Only the ADCM administrator (= superuser) can assign roles to the users. In particular, only the superuser can grant a superuser role to another user.

Policy is a triad that consists of the following elements:

1. User or user group
2. Role
3. Object, to which the role's permissions are applied.

Object can have one of the following types:

- Cluster
- Service
- Host
- Provider

While creating a policy the user is prompted to select one or more objects of the type that corresponds to the chosen user role.

A policy is applied to object as well as to all of its child objects. For example, if a policy is given to a cluster, then that policy also affects all services, components, and hosts of that cluster.

12.1 Built-in roles

ADCM comes with a few built-in roles. Changing the permissions of built-in roles is prohibited.

Таблица12.1.: Built-in ADCM roles

Role name	Description	Object type
ADCM User	View-only role that provides ability to view all objects information	None
Service Administrator	Role provides ability to configure and control the life cycle of the relevant service	Service
Cluster Administrator	Role provides ability to configure and control the relevant cluster, its hosts, and its services	Cluster
Provider Administrator	Role gives full control over the relevant provider	Provider
ADCM Administrator	Role gives full control over all aspects of ADCM	None

12.2 Permissions and roles

In addition to the built-in roles, ADCM administrator can create custom roles with permissions. Available permissions and their relations to the built-in roles are listed in the table below (scroll right to see all columns).

Таблица12.2.: Permissions and roles

Permission name	Description	ADCM User	Service Admin	Provider Admin	Cluster Admin	ADCM Admin
View any object configurations	The ability to view any object configurations	Yes	No	No	No	Yes
View cluster configurations	The ability to view cluster configurations	No	No	No	Yes	No
View service configurations	The ability to view service configurations	No	Yes	No	Yes	No
View component configurations	The ability to view component configurations	No	Yes	No	Yes	No
View provider configurations	The ability to view provider configurations	No	No	Yes	No	No
View host configurations	The ability to view host configurations	No	Yes	Yes	Yes	No
Edit cluster configurations	The ability to edit cluster configurations	No	No	No	Yes	Yes
Edit service configurations	The ability to edit service configurations	No	Yes	No	Yes	Yes
Edit component configurations	The ability to edit component configurations	No	Yes	No	Yes	Yes
Edit provider configurations	The ability to edite provider configurations	No	No	Yes	No	Yes
Edit host configurations	The ability to edit host configurations	No	No	Yes	Yes	Yes
View any object imports	The ability to view the import page	Yes	No	No	No	Yes
View imports	The ability to view the import page	No	Yes	No	Yes	No
Manage Imports	The ability to manage imports	No	Yes	No	Yes	Yes
View any cluster host-components	The ability to view any cluster hosts and components mappings	Yes	No	No	No	Yes
View Host-Components	The ability to view hosts and components mapping	No	Yes	No	Yes	No
Manage Host-Components	The ability to change hosts and components mapping	No	No	No	Yes	Yes
Add service	The ability to add service to cluster	No	No	No	Yes	Yes
Remove hosts	The ability to remove host (in Hosts page)	No	No	Yes	Yes	Yes
Add hosts to the cluster	The ability to add hosts to the cluster	No	No	No	Yes	Yes
Remove hosts from thecCluster	The ability to remove hosts from the cluster	No	No	No	Yes	Yes
Upgrade cluster bundle	The ability to upgrade the cluster bundle	No	No	No	Yes	Yes

Permission name	Description	ADCM User	Service Admin	Provider Admin	Cluster Admin	ADCM Admin
Upgrade provider bundle	The ability to upgrade the provider bundle	No	No	Yes	No	Yes
Create hostprovider	The ability to create hostproviders	No	No	No	No	Yes
Create host	The ability to create hosts	No	No	Yes	Yes	Yes
Remove hostprovider	The ability to remove hostprovider	No	No	No	No	Yes
Create cluster	The ability to create clusters	No	No	No	No	Yes
Remove cluster	The ability to remove cluster	No	No	No	No	Yes
Upload bundle	The ability to upload bundles	No	No	Yes	Yes	Yes
Remove bundle	The ability to remove bundles	No	No	Yes	Yes	Yes
View ADCM settings	The ability to view settings.	No	No	No	No	Yes
Edit ADCM settings	The ability to edit settings.	No	No	No	No	Yes
View users	The ability to view users	No	No	No	No	Yes
Add new user	The ability to create users	No	No	No	No	Yes
Delete user	The ability to remove users	No	No	No	No	Yes
Update user	The ability to update users	No	No	No	No	Yes
View roles	The ability to view roles	No	No	No	No	Yes
Add new role	The ability to create roles	No	No	No	No	Yes
Delete role	The ability to remove roles (only for custom roles)	No	No	No	No	Yes
Update role	The ability to update roles (only for custom roles)	No	No	No	No	Yes
View groups	The ability to view groups	No	No	No	No	Yes
Add new group	The ability to create groups	No	No	No	No	Yes
Delete group	The ability to remove groups	No	No	No	No	Yes
Update group	The ability to update groups	No	No	No	No	Yes
View policies	The ability to view policies	No	No	No	No	Yes
Add new policy	The ability to create policies	No	No	No	No	Yes
Delete policy	The ability to remove policies	No	No	No	No	Yes
Update policy	The ability to update policies	No	No	No	No	Yes
Cluster Action: <*Action name*>	The ability to perform the <*Action name*> action	No	No	No	Yes	Yes
Host Action: <*Action name*>	The ability to perform the <*Action name*> action	No	No	Yes	Yes	Yes
Service Action: <*Action name*>	The ability to perform the <*Action name*> action	No	Yes	No	Yes	Yes
Component Action: <*Action name*>	The ability to perform the <*Action name*> action	No	Yes	No	Yes	Yes
Provider Action: <*Action name*>	The ability to perform the <*Action name*> action	No	No	Yes	No	Yes

12.3 Creating a role

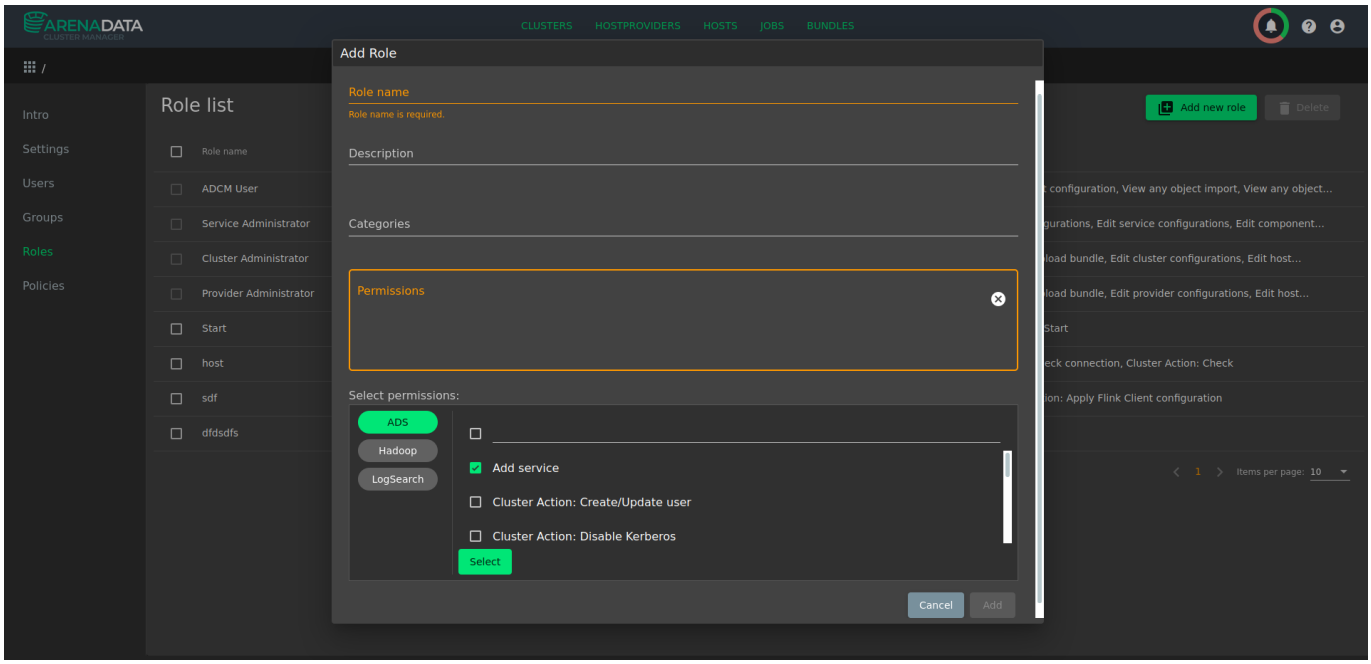
This functional is available only to the user with appropriate permissions.

In order to create a new role:

1. Log into ADCM console.

2. Navigate to the Roles tab. Roles list opens. 2. Click Add new role button. Add Role window opens. 3. Fill in the fields. Role name is required. 4. Add permissions. Use buttons with product names as filters to select permissions. 5. Click Add button.

New role is added.



12.4 Creating a policy

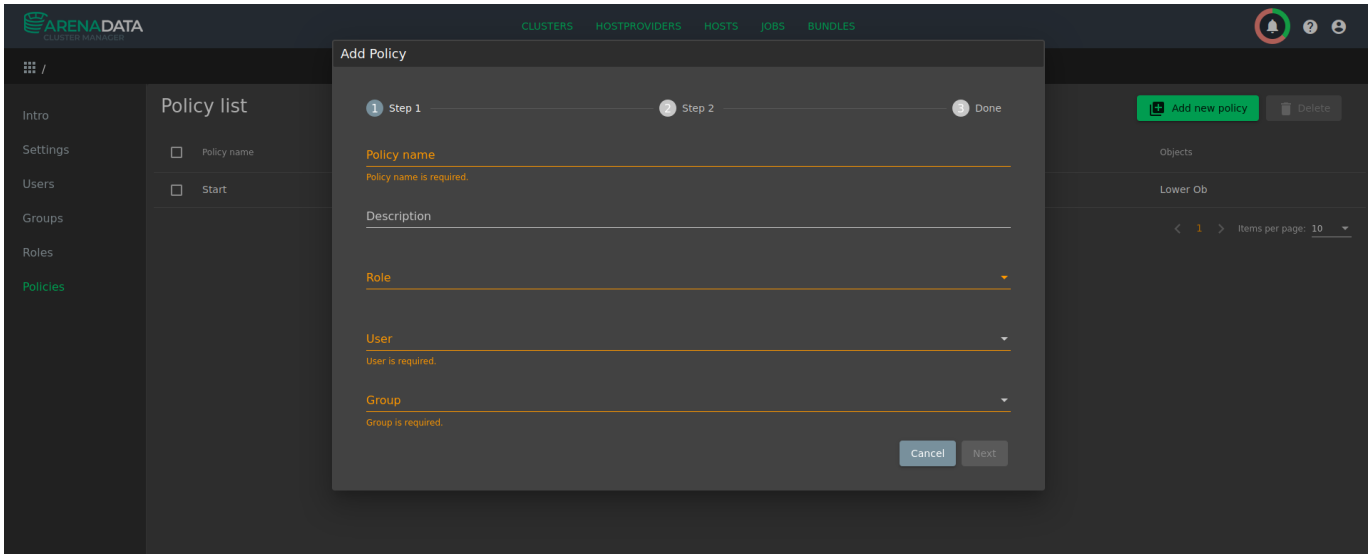
This functional is available only to the user with appropriate permissions.

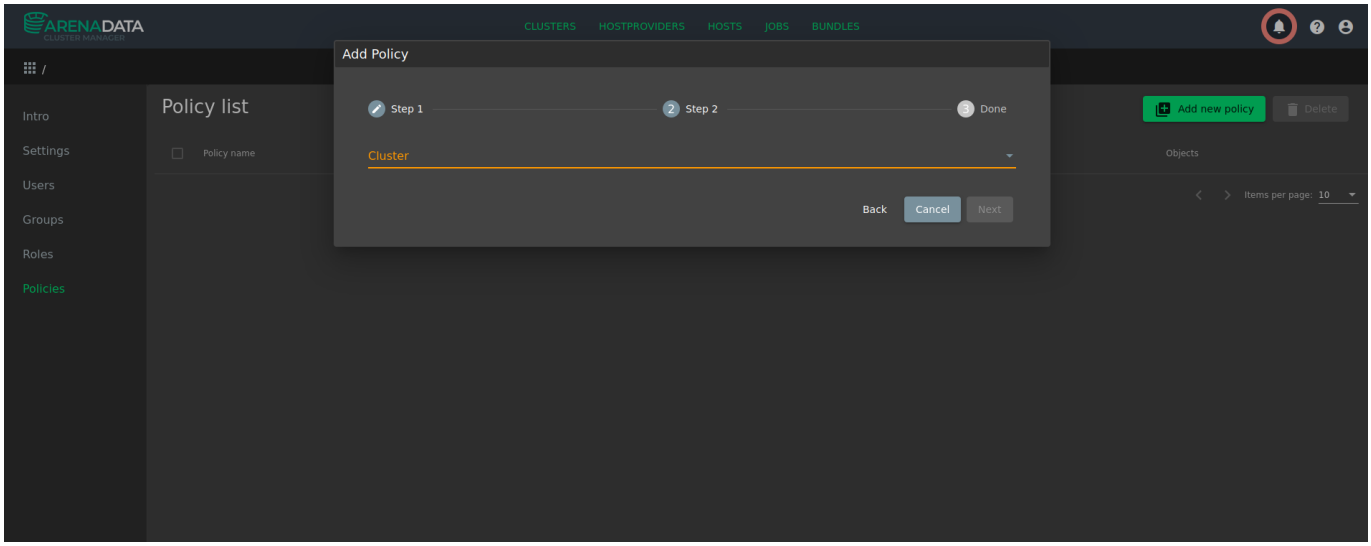
In order to create a new policy:

1. Log into ADCM console.

2. Navigate to the Policies tab. Policy list opens. 2. Click Add new policy button. Add Policy window opens. 3. Fill in the fields at Step 1. Policy name, Role, User, and Group are required. 4. Specify the objects at Step 2. 5. Click Add button at Step 3.

New policy is added.





Important: You can not mix permissions related to objects of different hierarchy.

Cluster hierarchy objects:

- Cluster
- Service
- Component
- Host

Provider hierarchy objects:

- Provider
- Host