

Arenadata™ Hadoop

Версия - v1.6.1

**Руководство администратора по работе с кластером
Arenadata Hadoop**

Оглавление

1	Учетная запись администратора Ambari	3
1.1	Авторизация в Ambari	3
1.2	Экран настроек Ambari	3
1.3	Изменение пароля учетной записи администратора	5
2	Управление кластером	7
2.1	Права доступа	7
2.2	Переименование кластера	8
2.3	Добавление узлов в существующий кластер	9
2.4	Переход на главный экран Ambari	9
3	Пользователи и группы	11
3.1	Типы пользователей и групп	11
3.2	Права администратора Ambari	12
3.3	Создание локального пользователя	12
3.4	Настройка статуса пользователя	12
3.5	Установка права администратора Ambari	14
3.6	Изменение пароля локального пользователя	15
3.7	Удаление локального пользователя	15
3.8	Создание локальной группы	15
3.9	Управление пользователями в группах	15
3.10	Удаление локальной группы	17
4	Карта портов	18
4.1	Порты HDFS	18
4.2	Порты MapReduce	19
4.3	Порты YARN	20
4.4	Порты Hive	22
4.5	Порты WebHCat	23
4.6	Порты HBase	24
4.7	Порты Oozie	24
4.8	Порты NiFi	25
4.9	Порты Zookeeper	26
5	Использование Ambari View	28
5.1	Files View	28
5.2	Capacity Scheduler View	29
5.3	Tez View	30
5.4	Hive View	30

5.5 Pig View 32

Apache Ambari – это система, которая помогает в управлении и контролировании кластеров **Hadoop**. Данное руководство предназначено для системных администраторов, ответственных за установку и обслуживание кластеров **Ambari** и **Hadoop**, управляемых **Ambari**.

При установке **Ambari** по умолчанию создается пользователь с правами администратора:

- Имя пользователя: *admin*;
- Пароль: *admin*.

При авторизации в системе администратору **Ambari** доступны следующие функции:

- Управление учетной записью администратора Ambari;
- Управление кластером;
- Управление пользователями и группами пользователей.

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

Глава 1

Учетная запись администратора Ambari

1.1 Авторизация в Ambari

После установки **Ambari** (см. [Инструкция по установке кластера](#)) для входа в систему необходимо ввести в веб-браузер следующий URL-адрес:

```
http://<your.ambari.server>:8080
```

Где указываются параметры:

- *<your.ambari.server>* – имя хоста установленного сервера Ambari;
- *8080* – HTTP-порт по умолчанию.

Далее следует ввести учетные данные администратора **Ambari**, созданного автоматически во время установки кластера:

- Имя пользователя: *admin*;
- Пароль: *admin*.

При корректном вводе данных открывается главная страница **Ambari**, на которой отображаются компоненты **Ambari** и их метрики ([Рис.1.1.](#)).

1.2 Экран настроек Ambari

Для перехода на экран настроек **Ambari** необходимо на главной странице перейти по пунктам меню панели управления “*admin* → *Manage Ambari*” (см. [Рис.1.1.](#)). При этом открывается экран настроек **Ambari** ([Рис.1.2.](#)).

Экран настроек **Ambari** состоит из панелей управления “Clusters”, “Views” и “User + Group Management”.

Панель “Clusters” отображает ссылки на имеющиеся кластеры и позволяет управлять их правами доступа. Панель “Views” позволяет создавать и редактировать представления системы и управлять правами доступа для этих представлений. Панель “User + Group Management” позволяет управлять, создавать и редактировать пользователей и группы пользователей.

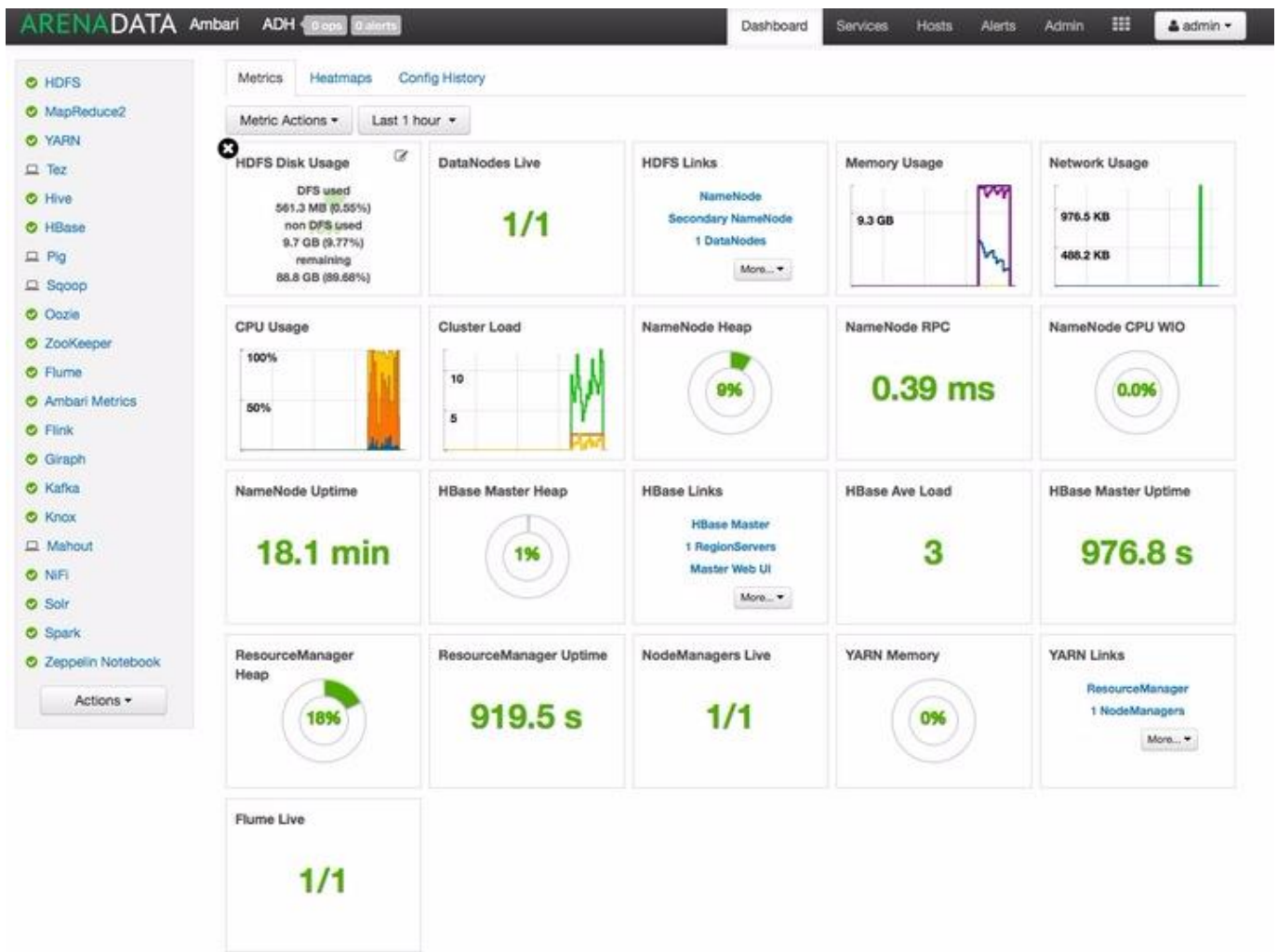


Рис.1.1.: Главная страница Амбари

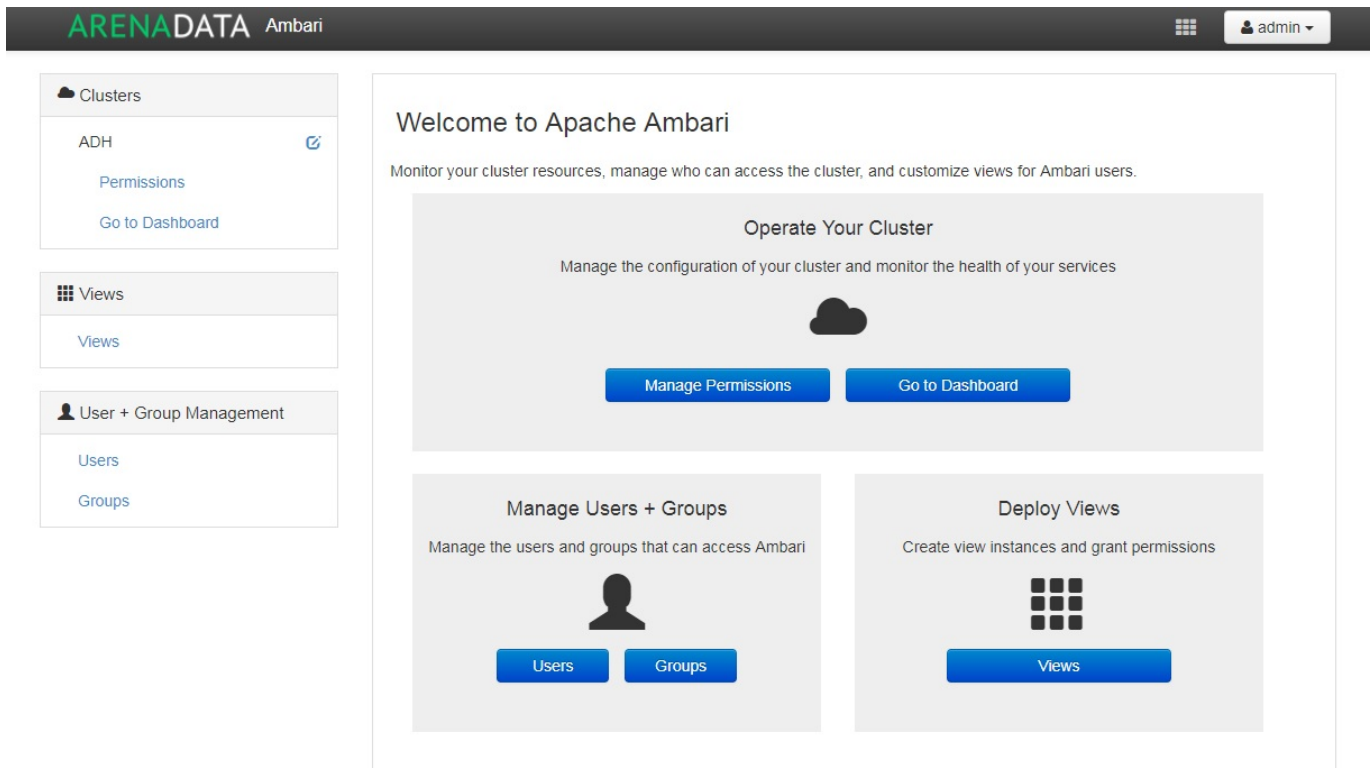


Рис.1.2.: Экран настроек Ambari

1.3 Изменение пароля учетной записи администратора

Во время установки и настройки кластера мастер установки автоматически создает пользователя с правами администратора **Ambari**. Для данного пользователя рекомендуется сменить пароль, заданный по умолчанию. Для этого на экране настроек **Ambari** необходимо (см. Рис.1.2.):

- В блоке “User + Group Management” перейти по ссылке “Users” и выбрать пользователя *admin*. При этом отображается окно параметров администратора (см. Рис.1.3.);
- В поле “Password” нажать кнопку *Change Password*;
- В открывшемся диалоговом окне ввести текущий пароль администратора и дважды указать новый пароль;
- Нажать кнопку *OK* для сохранения нового пароля.

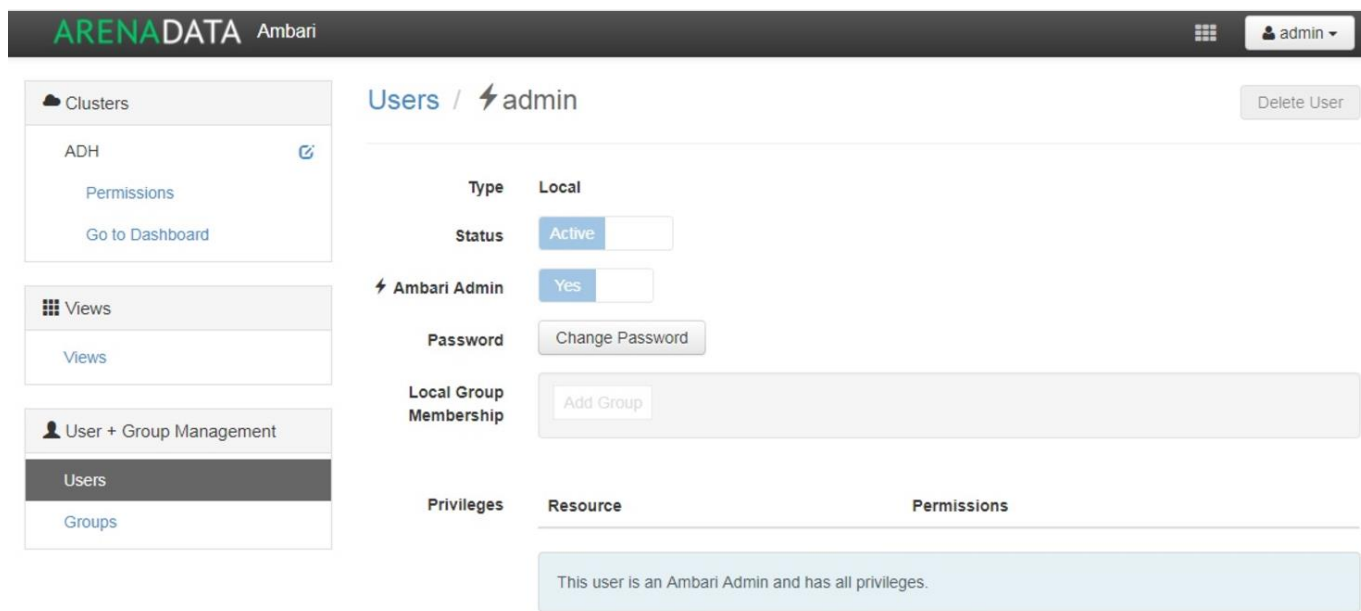


Рис.1.3.: Параметры администратора

Глава 2

Управление кластером

2.1 Права доступа

В **Ambari** существует два типа прав доступа к кластеру: “Operator” и “Read-Only”.

После создания кластера администраторы **Ambari** автоматически получают право доступа “Operator”. По умолчанию другие пользователи не имеют доступа к кластеру. Права доступа к кластеру для пользователей и групп пользователей назначаются администратором в экране настроек **Ambari** (см. Рис.1.2.).

Пользователям и группам пользователей с правом “Operator” предоставляется доступ к кластеру и обеспечивается контроль над следующими функциями:

- Запуск;
- Остановка;
- Перезапуск;
- Создание нового кластера.

Право доступа “Operator” обеспечивает контроль над следующими конфигурациями кластера:

- Изменение;
- Возврат.

Пользователям и группам пользователей с правом “Read-Only” предоставляется доступ только к просмотру кластера.

Для изменения прав доступа к кластеру для пользователей и групп пользователей администратору **Ambari** необходимо выполнить следующие действия:

- На панели “Clusters” в экране настроек **Ambari** нажать кнопку *Permissions*. При этом отображается табличная форма с правами доступа “Operator” и “Read-Only” с возможностью назначения их пользователям и группам пользователей (см. Рис.2.1.);
- Для назначения прав необходимо в блоках “Grant permissions to these users” и “Grant permissions to these groups” указать соответствующих пользователей и группы, нажав кнопку *Add User* или *Add Group*;
- Для снятия прав с пользователей или групп необходимо нажать значок “X”, расположенный рядом с именем пользователя или названием группы;
- Для сохранения изменений нажать кнопку с пиктограммой галочки.

Для группы без пользователей так же возможна установка прав.

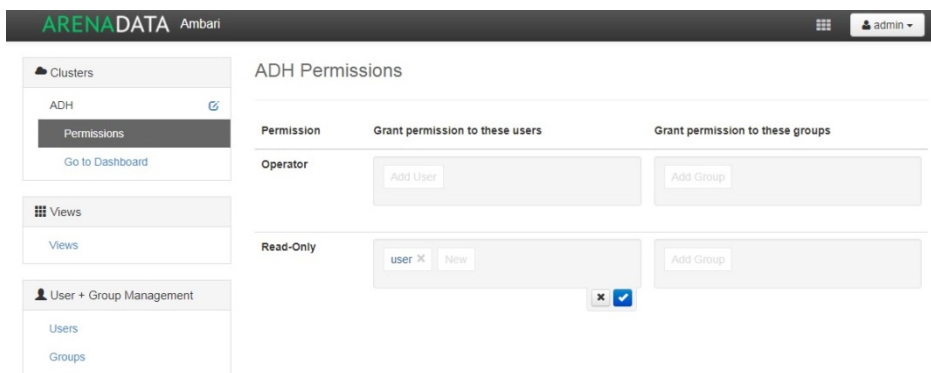


Рис.2.1.: Права доступа к кластеру

Important: Необходимо проверить наличие прав доступа у каждого пользователя, каждой группы и каждого пользователя в группе, чтобы убедиться, что у всех есть соответствующие разрешения

Дополнительные сведения об управлении пользователями и группами приведены в разделе [Пользователи и группы пользователей](#).

2.2 Переименование кластера

Для изменения названия кластера необходимо на панели “Clusters” кликнуть по значку *Rename Cluster* (Рис.2.2.).

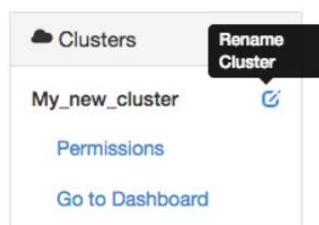


Рис.2.2.: Значок “Rename Cluster”

При этом название кластера становится доступным для редактирования. Необходимо вручную ввести новое название, нажать галочку и подтвердить действие.

Important: Чтобы изменения вступили в силу, необходимо перезапустить Ambari Server и Ambari Agents

При смене названия кластера также меняется ресурс **REST API**.

Important: Для использования кластера с новым названием необходимо настроить все применимые API

2.3 Добавление узлов в существующий кластер

Для добавления новых узлов в существующий кластер **Arenadata Hadoop** администратору **Ambari** необходимо:

1. В **Ambari** перейти на вкладку “Hosts” (см. [Рис.1.1.](#)) и в меню “Actions” выбрать пункт *Add New Hosts* ([Рис.2.3.](#)).

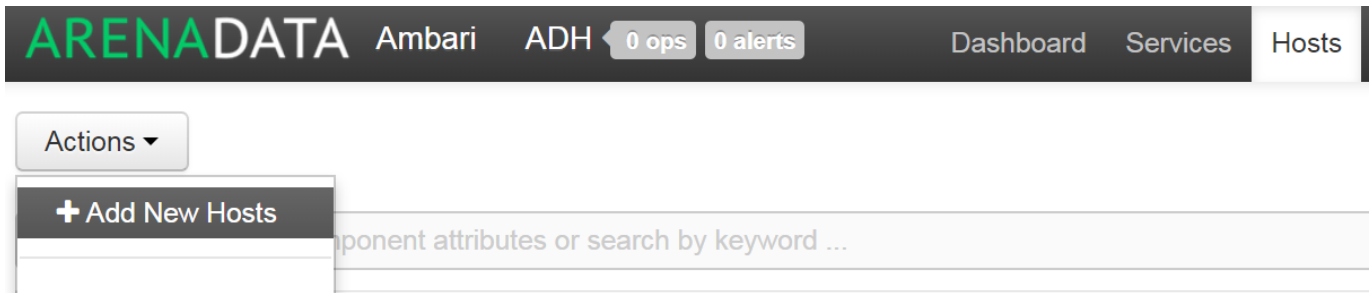


Рис.2.3.: Добавление нового узла

2. При этом запускается мастер добавления узла, в котором необходимо указать адрес нового узла, компоненты кластера, которые необходимо развернуть на новом узле, параметры этих компонентов и т.п. ([Рис.2.4.](#)).

Последовательность выполняемых действий и настроек в целом повторяет процесс развертывания нового кластера, за исключением некоторых пунктов.

Примечания:

- На добавляемом узле должны быть выполнены все настройки, описанные в разделе [Подготовка к установке Hadoop](#) документа [Инструкция по установке кластера](#);
- В случае добавления узла, аналогичного по функционалу уже имеющимся в кластере (например, дополнительного узла HDFS DataNode), рекомендуется использовать узлы с конфигурацией, аналогичной уже существующим узлам с той же ролью (аппаратные ресурсы, разметка дисков, параметры ОС и т.д.). В противном случае для добавляемого узла может потребоваться использование отдельной группы конфигурации **Ambari**.

2.4 Переход на главный экран Ambari

Для возвращения на главный экран **Ambari** необходимо на панели “Clusters” нажать кнопку *Go to Dashboard* (см. [Рис.1.2.](#)). При этом происходит переход на главный экран **Ambari** (см. [Рис.1.1.](#)).

Add Host Wizard

ADD HOST WIZARD

- Install Options**
- Confirm Hosts
- Assign Slaves and Clients
- Configurations
- Review
- Install, Start and Test
- Summary

Install Options

Enter the list of hosts to be included in the cluster and provide your S

Target Hosts

Enter a list of hosts using the Fully Qualified Domain Name (FQDN) Expressions

Host Registration Information

Provide your [SSH Private Key](#) to automatically register hosts

Файл не выбран

Рис.2.4.: Мастер добавления узла

Глава 3

Пользователи и группы

Администратор **Ambari** создает и управляет пользователями и группами пользователей, доступными для **Ambari**, поддерживающий два типа пользователей – локальные и LDAP.

- *Типы пользователей и групп;*
- *Права администратора Ambari.*

Администратор **Ambari** может импортировать информацию о пользователях и группах в **Ambari** из внешних систем **LDAP** и управлять ими:

- *Создание локального пользователя;*
- *Настройка статуса пользователя;*
- *Установка права администратора Ambari;*
- *Изменение пароля локального пользователя;*
- *Удаление локального пользователя;*
- *Создание локальной группы;*
- *Управление пользователями в группах;*
- *Удаление локальной группы.*

3.1 Типы пользователей и групп

Ambari поддерживает два типа пользователей и групп: локальные и LDAP.

Локальные пользователи хранятся и аутентифицируются в базе данных **Ambari**. Пользователи LDAP имеют базовую информацию учетной записи, хранящуюся в **Ambari**. В отличие от локальных, пользователи LDAP аутентифицируются по внешней системе **LDAP**.

Локальные группы хранятся в базе данных **Ambari**. Группы LDAP имеют базовую информацию, хранящуюся в **Ambari**, включая информацию о пользователях в группе. В отличие от локальных, группы LDAP импортируются и синхронизируются из внешней системы **LDAP**.

Необходимо настроить **Ambari** для аутентификации с внешней системой **LDAP** для того, чтобы использовать пользователей и группы LDAP с **Ambari**.

Important: Новый пользователь (или группа пользователей) Ambari, созданный либо локально, либо путем синхронизации с LDAP, по умолчанию не получает никаких прав. Администратору Ambari необходимо предоставлять каждому пользователю права на доступ к кластерам

3.2 Права администратора Ambari

Администратор **Ambari** может создавать и удалять пользователей, изменять их пароли и редактировать настройки, а так же управлять определенными привилегиями. В таблице перечислены доступные администратору **Ambari** права для локальных и LDAP пользователей и групп.

Таблица 3.1.: Права администратора для локальных и LDAP пользователей и групп

Права администратора	Локальный пользователь	Пользователь LDAP
Смена пароля	Доступно	Недоступно
Установка пользователю права администратора Ambari	Доступно	Доступно
Управление количеством пользователей в группе	Доступно	Недоступно
Удаление пользователя	Доступно	Недоступно
Управление статусом пользователя	Доступно	Доступно

3.3 Создание локального пользователя

Для создания локального пользователя администратору **Ambari** необходимо:

- В блоке “User + Group Management” перейти по ссылке “Users” (Рис.3.1).
- При этом открывается окно с перечнем пользователей Ambari (Рис.3.2);
- Нажать кнопку *Create Local User*. При этом открывается окно создания локального пользователя (Рис.3.3);
- В поле “Username” ввести уникальное имя пользователя (все имена пользователей автоматически преобразуются в нижний регистр);
- В поле “Password” ввести пароль и подтвердить его;
- Сохранить изменения, нажав кнопку *Save*.

3.4 Настройка статуса пользователя

У каждого пользователя **Ambari** есть статус, который может принимать следующие значения:

- *Active* – пользователь активен и ему разрешено входить в Ambari;
- *Inactive* – пользователь неактивен и лишен возможности входа в систему.

Установив флаг состояния как активный или неактивный, администратор может подключать или отключать доступ учетной записи пользователя к **Ambari**, сохраняя при этом информацию об учетной записи и о правах пользователя.

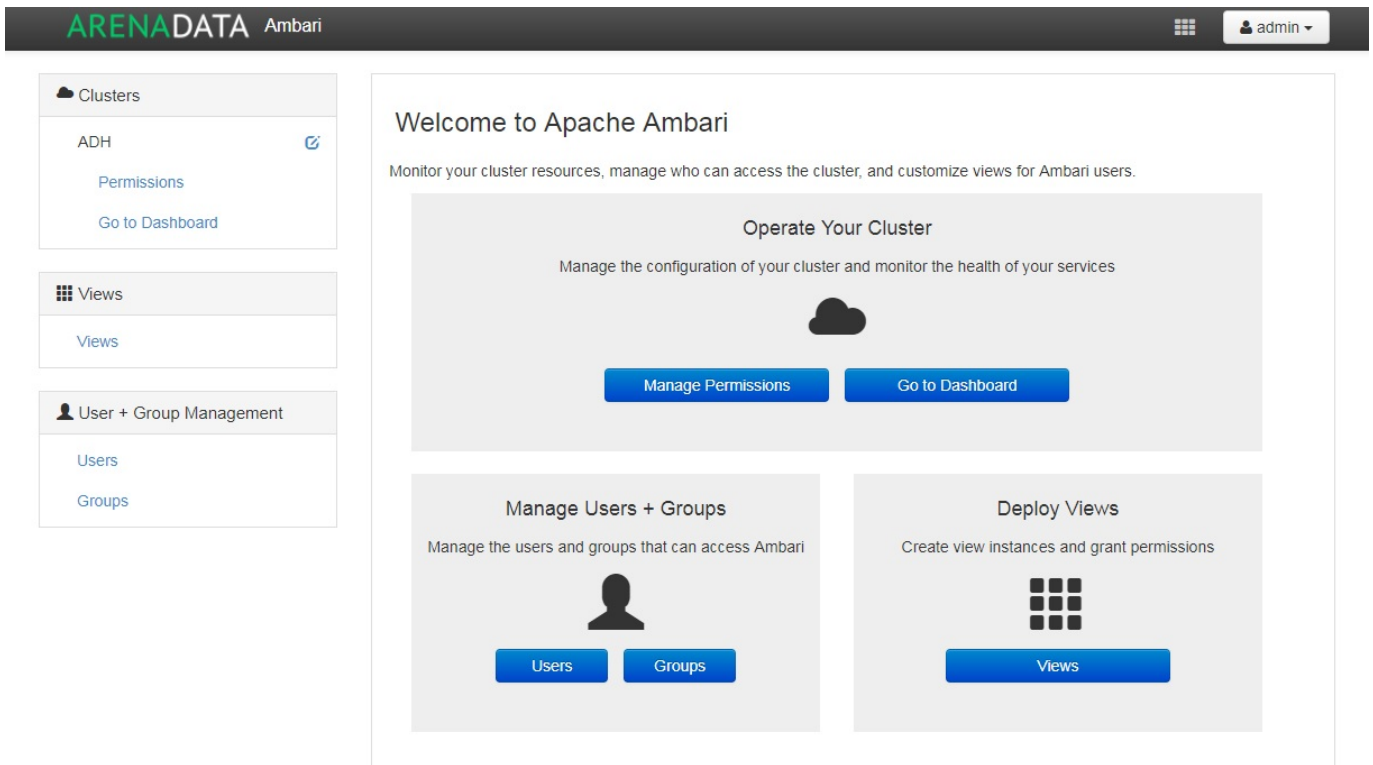


Рис.3.1.: Экран настроек Ambari

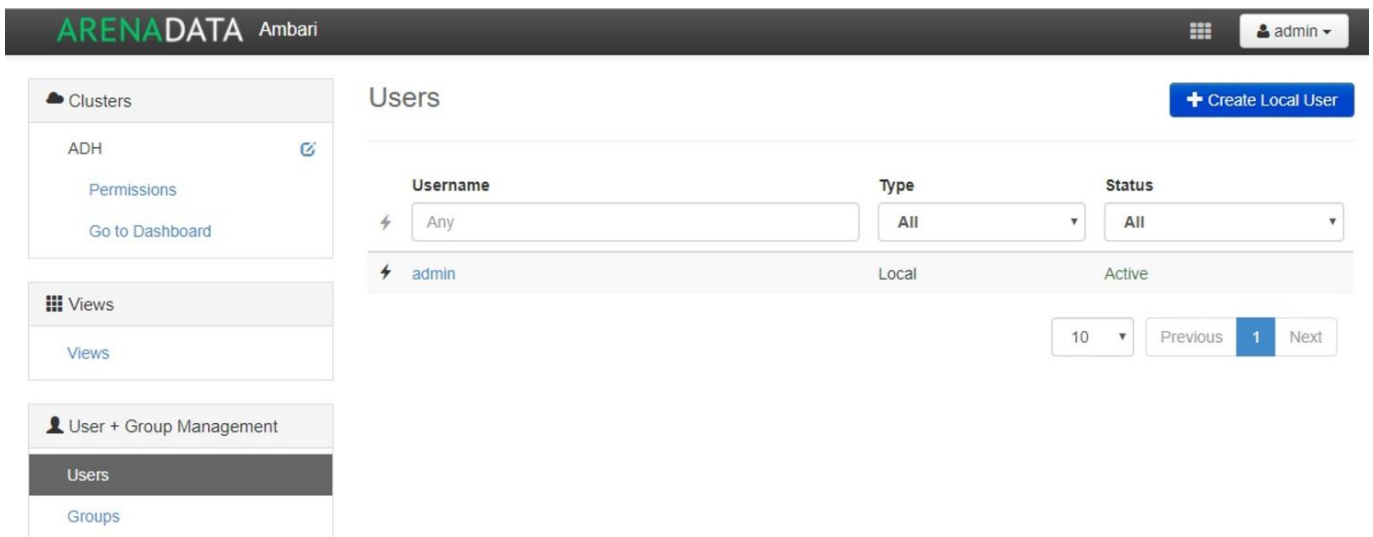


Рис.3.2.: Пользователи Ambari

The screenshot shows the 'Create Local User' page in the Ambari interface. On the left, there is a navigation sidebar with 'Users' selected under 'User + Group Management'. The main content area contains the following form fields:

- Username:** A text input field with the placeholder 'User name'.
- Type:** A dropdown menu set to 'Local'.
- Status:** A dropdown menu set to 'Active'.
- Ambari Admin:** A radio button set to 'No'.
- Password:** A text input field.
- Password confirmation:** A text input field.

At the bottom right, there are 'Cancel' and 'Save' buttons.

Рис.3.3.: Создание локального пользователя

Для установки пользователю статуса необходимо, чтобы администратор **Ambari** выполнил следующие действия:

- Выбрать пользователя в окне “Users” (см. Рис.3.2.). При этом открывается окно параметров (Рис.3.4.);

The screenshot shows the 'user' parameter page in the Ambari interface. The page includes a 'Delete User' button at the top right. The main content area displays the following details:

- Type:** Local
- Status:** Active
- Ambari Admin:** No
- Password:** Change Password
- Local Group Membership:** Add Group

Below these details is a table of Privileges:

Resource	Permissions
ADH	Read-Only

Рис.3.4.: Параметры пользователя

- В поле “Status” установить флаг в положение *Active* или *Inactive*;
- Нажать кнопку *OK* для подтверждения. Изменение статуса фиксируется незамедлительно.

3.5 Установка права администратора Ambari

Одному или нескольким пользователям системы можно установить права администратора **Ambari**. Для этого администратору **Ambari** необходимо в окне параметров пользователя в поле “Ambari Admin” установить флаг в положение *Yes* или *No* (см. Рис.3.4.).

Important: Для предотвращения блокировки прав администратора у самого администратора снятие флага “Ambari Admin” у данной учетной записи невозможно

3.6 Изменение пароля локального пользователя

Администратор **Ambari** может менять пароли локальных пользователей системы.

Important: Пароли LDAP не управляются Ambari и не могут быть изменены с Ambari, поскольку пользователи LDAP аутентифицируются по внешнему LDAP

Для изменения пароля локального пользователя администратору **Ambari** необходимо в окне параметров пользователя в поле “Password” нажать кнопку *Change password* и ввести пароль администратора для подтверждения наличия прав на смену пароля пользователю (см. Рис.3.4.). Далее следует ввести новый пароль для локального пользователя, подтвердить его и нажать кнопку *Save*.

3.7 Удаление локального пользователя

Important: Удаление локального пользователя полностью удаляет учетную запись пользователя из системы, включая все связанные с ним права

Для удаления локального пользователя администратору **Ambari** необходимо в окне параметров пользователя нажать кнопку *Delete User* и подтвердить действие (см. Рис.3.4.).

После удаления локального пользователя его имя можно повторно использовать.

Important: С целью предотвращения входа в систему локального пользователя можно установить ему статус “Inactive” (*Настройка статуса пользователя*)

3.8 Создание локальной группы

Для создания локальной группы администратору **Ambari** необходимо выполнить следующие действия:

- В блоке “User + Group Management” перейти по ссылке “Groups” (см. Рис.3.1.). При этом открывается окно групп пользователей (Рис.3.5.);
- Нажать кнопку *Create Local Group*;
- В поле “Group name” ввести уникальное имя группы пользователей;
- Сохранить изменения, нажав кнопку *Save*.

3.9 Управление пользователями в группах

Администратор **Ambari** может управлять составом групп пользователей. Для этого необходимо в окне со списком групп выбрать необходимую для изменения группу (см. Рис.3.5.). При этом открывается окно параметров группы пользователей (Рис.3.6.).

Для добавления локального пользователя в группу необходимо в поле “Local Members” в ячейку *New* ввести имя пользователя, которого необходимо добавить в группу. Для сохранения изменений нажать кнопку с пиктограммой галочки.

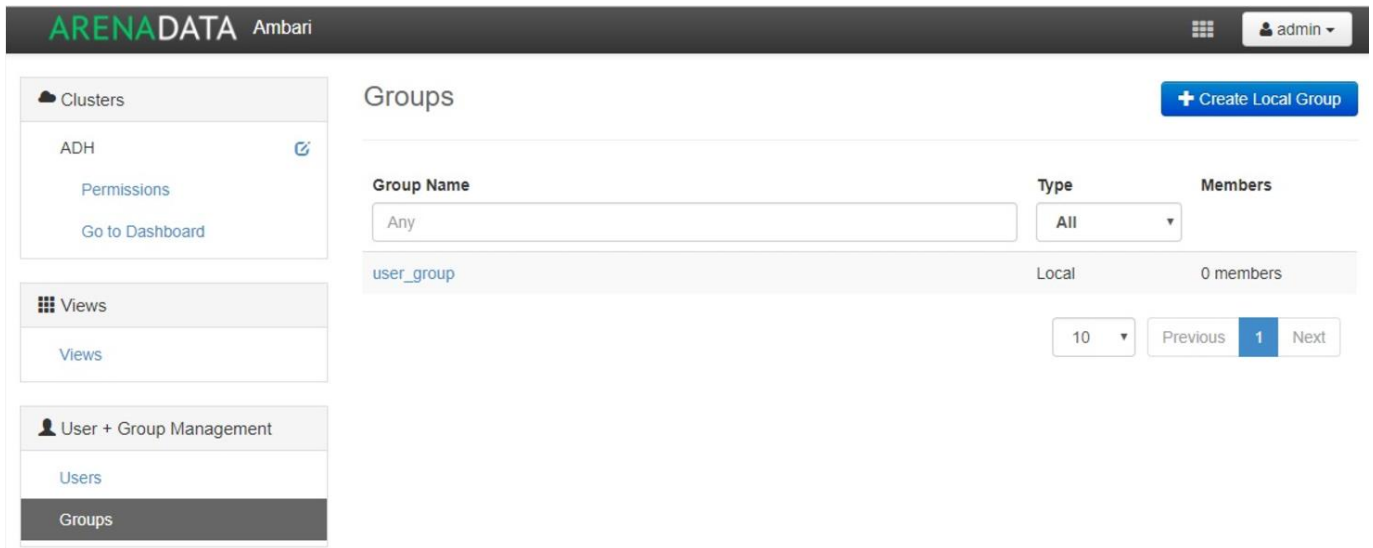


Рис.3.5.: Группы пользователей

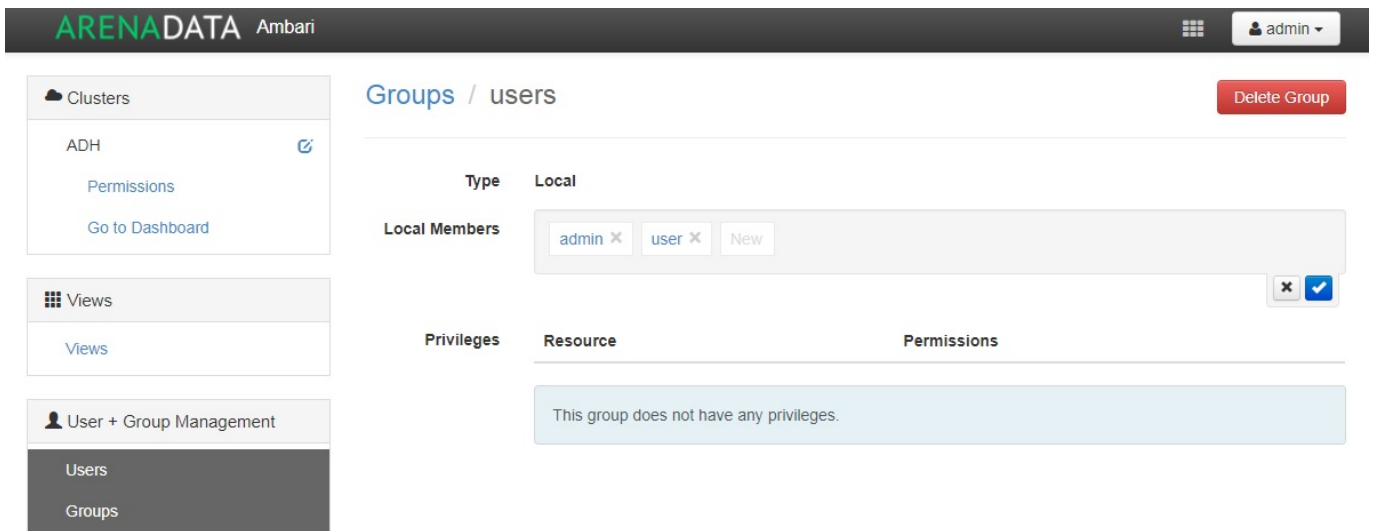


Рис.3.6.: Параметры группы пользователей

Для удаления локального пользователя из группы необходимо в поле “Local Members” нажать значок “X”, расположенный рядом с именем пользователя, которого необходимо исключить из группы. Для сохранения изменений нажать кнопку с пиктограммой галочки.

3.10 Удаление локальной группы

Important: Удаление локальной группы приводит к удалению всех прав, связанных с группой

Для удаления локальной группы администратору **Ambari** необходимо в окне параметров группы пользователей нажать кнопку *Delete Group* и подтвердить действие (см. [Рис.3.6](#)). При этом группа и соответствующая информация о ее пользователях удаляется.

Глава 4

Карта портов

4.1 Порты HDFS

Список портов, используемых по умолчанию для различных **HDFS**-сервисов, представлен в таблице.

Таблица 4.1.: Порты HDFS

Сервис	Характеристика
NameNode WebUI	Сервер – Master Nodes <ul style="list-style-type: none">• Порт по умолчанию – 50070<ul style="list-style-type: none">– Протокол: http– Описание: Web-интерфейс для работы с файловой системой HDFS– Доступ пользователям: Да (администратор, разработчик, поддержка)– Параметр конфигурации: dfs.http.address• Порт по умолчанию – 50470<ul style="list-style-type: none">– Протокол: https– Описание: Безопасный http-сервис– Доступ пользователям: Да (администратор, разработчик, поддержка)– Параметр конфигурации: dfs.https.address
NameNode metadata service	Сервер – Master Nodes <ul style="list-style-type: none">• Порт по умолчанию – 8020/9000<ul style="list-style-type: none">– Протокол: IPC– Описание: Взаимодействие с метаданными файловой системы– Доступ пользователям: Да (все клиенты для прямого взаимодействия с HDFS)– Параметр конфигурации: fs.default.name

Сервис	Характеристика
DataNode	<p>Сервер – Все Slave Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 50075 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для доступа к логам, статусам и пр. – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.datanode.http.address • Порт по умолчанию – 50475 <ul style="list-style-type: none"> – Протокол: https – Описание: Безопасный http-сервис – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.datanode.https.address • Порт по умолчанию – 50010 <ul style="list-style-type: none"> – Описание: Передача данных – Параметр конфигурации: dfs.datanode.address • Порт по умолчанию – 50020 <ul style="list-style-type: none"> – Протокол: IPC – Описание: Операции с метаданными – Доступ пользователям: Нет – Параметр конфигурации: dfs.datanode.ipc.address
Secondary NameNode	<p>Сервер – Secondary NameNodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 50090 <ul style="list-style-type: none"> – Протокол: http – Описание: Чекпоинт для метаданных NameNode – Доступ пользователям: Нет – Параметр конфигурации: dfs.secondary.http.address
HDFS HFTP	<p>Сервер – Master NameNodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 50470 <ul style="list-style-type: none"> – Протокол: https – Описание: Порт для доступа к HFTP файловой системы – Параметр конфигурации: dfs.https.port
JournalNode Web UI	<p>Сервер – Master NameNodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8480 <ul style="list-style-type: none"> – Протокол: http – Описание: Порт для доступа к HFTP файловой системы – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.journalnode.http-address • Порт по умолчанию – 8481 <ul style="list-style-type: none"> – Протокол: https – Описание: Безопасный http-сервис – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.journalnode.https-address

4.2 Порты MapReduce

Список портов, используемых по умолчанию для различных **MapReduce**-сервисов, представлен в таблице.

Таблица 4.2.: Порты MapReduce

Сервис	Характеристика
JobTracker WebUI	Сервер – Master Nodes <ul style="list-style-type: none"> • Порт по умолчанию – 50030 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для работы JobTracker – Доступ пользователям: Да – Параметр конфигурации: mapred.job.tracker.http.address
JobTracker	Сервер – Master Nodes <ul style="list-style-type: none"> • Порт по умолчанию – 8021 <ul style="list-style-type: none"> – Протокол: IPС – Описание: Для публикации заданий – Доступ пользователям: Да (все клиенты, которым требуется запуск MR, Hive, Pig и т.д.) – Параметр конфигурации: mapred.job.tracker
TaskTracker Web UI and Shuffle	Сервер – Все Slave Nodes <ul style="list-style-type: none"> • Порт по умолчанию – 50060 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для DataNode (логи, статус) – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: mapred.task.tracker.http.address
History Server WebUI	<ul style="list-style-type: none"> • Порт по умолчанию – 51111 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для истории заданий – Доступ пользователям: Да – Параметр конфигурации: mapreduce.history.server.http.address
MapReduce Shuffle Port	<ul style="list-style-type: none"> • Порт по умолчанию – 13562 <ul style="list-style-type: none"> – Описание: Порт, на котором работает ShuffleHandler – Доступ пользователям: Нет – Параметр конфигурации: mapreduce.shuffle.port

4.3 Порты YARN

Список портов, используемых по умолчанию для различных YARN-сервисов, представлен в таблице.

Таблица 4.3.: Порты YARN

Сервис	Характеристика
ResourceManager WebUI	<p>Сервер – Master Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8088 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для Resource Manager – Доступ пользователям: Да – Параметр конфигурации: yarn.resourcemanager.webapp.address
ResourceManager	<p>Сервер – Master Nodes (ResourceManager Node)</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8050 <ul style="list-style-type: none"> – Протокол: IPC – Описание: Для публикации заданий – Доступ пользователям: Да (все клиенты, которым требуется запуск YARN-приложений) – Параметр конфигурации: yarn.resourcemanager.address • Порт по умолчанию – 8025 <ul style="list-style-type: none"> – Протокол: http – Описание: Web-интерфейс для DataNode (логи, статус) – Доступ пользователям: Да (все клиенты, которым требуется запуск YARN-приложений) – Параметр конфигурации: mapred.task.tracker.http.address • Порт по умолчанию – 9099 <ul style="list-style-type: none"> – Протокол: http – Описание: Прокси для Resource Manager – Доступ пользователям: Да – Параметр конфигурации: yarn.web-proxy.address • Порт по умолчанию – 8141 <ul style="list-style-type: none"> – Протокол: http – Описание: Адрес планировщика – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.resourcemanager.admin.address
Scheduler	<p>Сервер – Master Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8030 <ul style="list-style-type: none"> – Описание: Адрес планировщика – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.resourcemanager.scheduler.address

Сервис	Характеристика
NodeManager	<p>Сервер – Master Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 45454 <ul style="list-style-type: none"> – Протокол: http – Описание: Адрес NodeManager – Параметр конфигурации: yarn.nodemanager.address <p>Сервер – Slave Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8040 <ul style="list-style-type: none"> – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.localizer.address • Порт по умолчанию – 8042 <ul style="list-style-type: none"> – Протокол: http – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.webapp.address • Порт по умолчанию – 8044 <ul style="list-style-type: none"> – Протокол: https – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.webapp.https.address
Timeline Server	<p>Сервер – Master Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 10200 <ul style="list-style-type: none"> – Протокол: http – Описание: Адрес Timeline Server – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.timeline-service.address • Порт по умолчанию – 8188 <ul style="list-style-type: none"> – Протокол: http – Описание: Адрес Timeline Server Webapp – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.timeline-service.webapp.address • Порт по умолчанию – 8190 <ul style="list-style-type: none"> – Протокол: https – Описание: Адрес Timeline Server Webapp https – Доступ пользователям: Да (администратор, разработчик, поддержка)

4.4 Порты Hive

Список портов, используемых по умолчанию для различных **Hive**-сервисов, представлен в таблице.

Таблица 4.4.: Порты Hive

Сервис	Характеристика
Hive Server2	<p>Сервер – Hive Server машина</p> <ul style="list-style-type: none"> • Порт по умолчанию – 10000 <ul style="list-style-type: none"> – Протокол: thrift – Описание: Сервис для подключения к Hive (Thrift/JDBC) – Доступ пользователям: Да (все клиенты, которым требуется подключение к Hive) – Параметр конфигурации: hive.server2.thrift.port • Порт по умолчанию – 10001 <ul style="list-style-type: none"> – Протокол: http – Описание: Сервис для подключения к Hive (http) – Доступ пользователям: Да (все клиенты, которым требуется подключение к Hive) – Параметр конфигурации: hive.server2.transport.mode
JobTracker	<p>Сервер – Master Nodes</p> <ul style="list-style-type: none"> • Порт по умолчанию – 8021 <ul style="list-style-type: none"> – Протокол: IPC – Описание: Для публикации заданий – Доступ пользователям: Да (все клиенты, которым требуется запуск MR, Hive, Pig. Задачи, использующие HCatalog)
Hive Web UI	<p>Сервер – Hive Server машина</p> <ul style="list-style-type: none"> • Порт по умолчанию – 9999 <ul style="list-style-type: none"> – Протокол: thrift – Описание: WebUI для Hive – Доступ пользователям: Да – Параметр конфигурации: hive.hwi.listen.port • Порт по умолчанию – 9933 <ul style="list-style-type: none"> – Протокол: http – Доступ пользователям: Да (все клиенты, которым требуется запуск MR, Hive, Pig) – Параметр конфигурации: hive.metastore.uris

4.5 Порты WebHCat

Список портов, используемых по умолчанию для различных **WebHCat**-сервисов, представлен в таблице.

Таблица 4.5.: Порты WebHCat

Сервис	Характеристика
WebHCat Server	<p>Сервер – WebHCat Server машина</p> <ul style="list-style-type: none"> • Порт по умолчанию – 50111 <ul style="list-style-type: none"> – Протокол: http – Описание: Web API для доступа к HCatalog и к другим сервисам Hadoop – Доступ пользователям: Да – Параметр конфигурации: templeton.port

4.6 Порты HBase

Список портов, используемых по умолчанию для различных **HBase**-сервисов, представлен в таблице.

Таблица 4.6.: Порты HBase

Сервис	Характеристика
HMaster	Сервер – Master Nodes (HBase Master Node и back-up HBase Master node) <ul style="list-style-type: none"> • Порт по умолчанию – 60000 <ul style="list-style-type: none"> – Доступ пользователям: Да – Параметр конфигурации: hbase.master.port
HMaster Info Web UI	Сервер – Master Nodes (HBase Master Node и back-up HBase Master node) <ul style="list-style-type: none"> • Порт по умолчанию – 60010 <ul style="list-style-type: none"> – Протокол: http – Описание: Порт для HBase Master UI – Доступ пользователям: Да – Параметр конфигурации: hbase.master.info.port
Region Server	Сервер – Все Slave Nodes <ul style="list-style-type: none"> • Порт по умолчанию – 60020 <ul style="list-style-type: none"> – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: hbase.regionserver.port • Порт по умолчанию – 60030 <ul style="list-style-type: none"> – Протокол: http – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: hbase.regionserver.info.port
HBase Thrift Server	Сервер – Все Thrift Server <ul style="list-style-type: none"> • Порт по умолчанию – 9090 <ul style="list-style-type: none"> – Описание: Порт, используемый HBase Thrift-сервером – Доступ пользователям: Да
HBase Thrift Server Web UI	Сервер – Все Thrift Server <ul style="list-style-type: none"> • Порт по умолчанию – 9090 <ul style="list-style-type: none"> – Описание: Web-интерфейс для HBase Thrift-сервера – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: hbase.thrift.info.port

4.7 Порты Oozie

Список портов, используемых по умолчанию для различных **Oozie**-сервисов, представлен в таблице.

Таблица 4.7.: Порты Oozie

Сервис	Характеристика
Oozie	Сервер – Oozie Server <ul style="list-style-type: none"> • Порт по умолчанию – 11000 <ul style="list-style-type: none"> – Протокол: TCP – Описание: Порт, используемый для запуска Oozie Server – Доступ пользователям: Да – Параметр конфигурации: OOOZIE_HTTP_PORT in oozie_env.sh • Порт по умолчанию – 11001 <ul style="list-style-type: none"> – Протокол: TCP – Описание: Порт, используемый админ-консолью для запуска Oozie Server – Доступ пользователям: Нет – Параметр конфигурации: OOOZIE_ADMIN_PORT in oozie_env.sh • Порт по умолчанию – 11443 <ul style="list-style-type: none"> – Протокол: TCP – Описание: Безопасный порт, используемый для запуска Oozie Server – Доступ пользователям: Да – Параметр конфигурации: OOOZIE_HTTPS_PORT in oozie_env.sh

4.8 Порты NiFi

Список портов, используемых по умолчанию для различных NiFi-сервисов, представлен в таблице.

Таблица 4.8.: Порты NiFi

Сервис	Характеристика
NiFi	<p>Сервер – NiFi сервер</p> <ul style="list-style-type: none"> • Порт по умолчанию – 9090 <ul style="list-style-type: none"> – Протокол: http – Описание: NiFi Server http port – Доступ пользователям: Да • Порт по умолчанию – 9091 <ul style="list-style-type: none"> – Протокол: https – Описание: NiFi Server https port – Доступ пользователям: Да • Порт по умолчанию – 10443 <ul style="list-style-type: none"> – Описание: Certificate Authority – Доступ пользователям: Нет • Порт по умолчанию – 8022 <ul style="list-style-type: none"> – Описание: Удаленный input-порт – Доступ пользователям: Да – Параметр конфигурации: nifi.remote.input.socket.port • Порт по умолчанию – 8021 <ul style="list-style-type: none"> – Описание: Node protocol-порт – Доступ пользователям: Да – Параметр конфигурации: nifi.cluster.node.protocol.port • Порт по умолчанию – 8070 <ul style="list-style-type: none"> – Протокол: http – Описание: NiFi http port – Доступ пользователям: Да • Порт по умолчанию – 9088 <ul style="list-style-type: none"> – Описание: Nifi Protocol Port – Доступ пользователям: Да – Параметр конфигурации: nifi.remote.input.socket.port

4.9 Порты Zookeeper

Список портов, используемых по умолчанию для различных **Zookeeper**-сервисов, представлен в таблице.

Таблица 4.9.: Порты Zookeeper

Сервис	Характеристика
Zookeeper Server	<p data-bbox="418 266 784 296">Сервер – Все Zookeeper Nodes</p> <ul data-bbox="459 296 1450 705" style="list-style-type: none"><li data-bbox="459 296 813 325">• Порт по умолчанию – 2181<ul data-bbox="513 325 1203 451" style="list-style-type: none"><li data-bbox="513 325 727 354">– Протокол: http<li data-bbox="513 354 1203 384">– Описание: Сервис доступа к Zookeeper Server/Quorum<li data-bbox="513 384 870 413">– Доступ пользователям: Да<li data-bbox="513 413 1159 443">– Параметр конфигурации: zookeeper.port/clientPort<li data-bbox="459 451 813 480">• Порт по умолчанию – 2888<ul data-bbox="513 480 1450 577" style="list-style-type: none"><li data-bbox="513 480 1450 510">– Описание: Порт используется Zookeeper для взаимодействия компонентов<li data-bbox="513 510 881 539">– Доступ пользователям: Нет<li data-bbox="513 539 1154 569">– Параметр конфигурации: hbase.zookeeper.peerport<li data-bbox="459 577 813 606">• Порт по умолчанию – 3888<ul data-bbox="513 606 1450 705" style="list-style-type: none"><li data-bbox="513 606 1450 636">– Описание: Порт используется Zookeeper для взаимодействия компонентов<li data-bbox="513 636 881 665">– Доступ пользователям: Нет<li data-bbox="513 665 1174 695">– Параметр конфигурации: hbase.zookeeper.leaderport

Глава 5

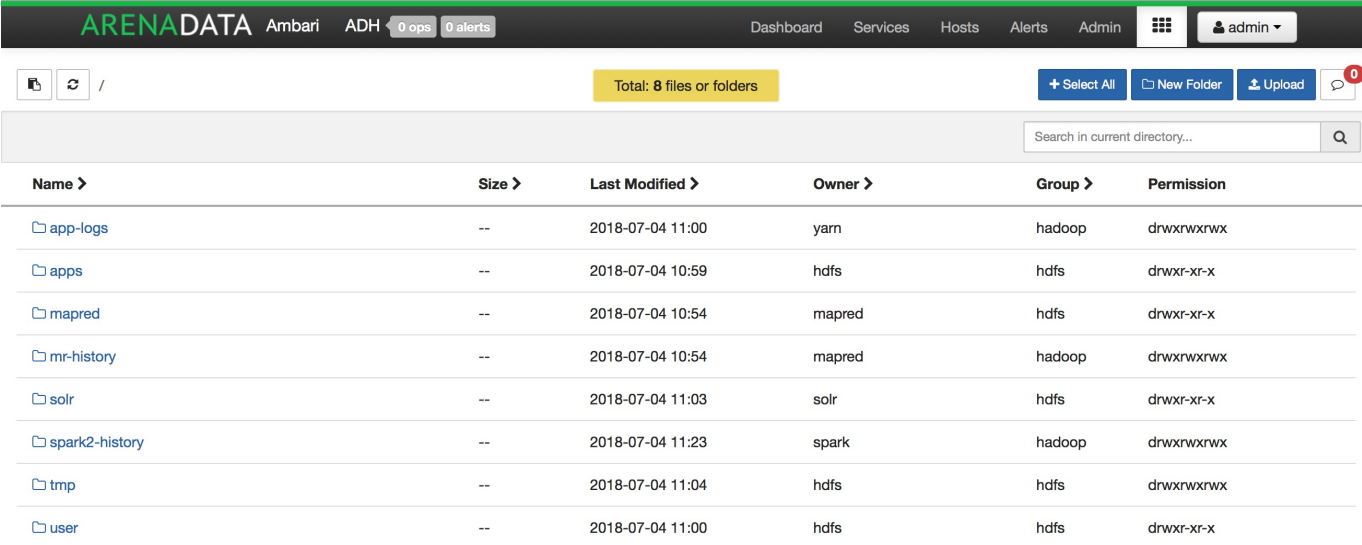
Использование Ambari View

Arenadata Ambari предоставляет возможность использования представления (View) для работы с различными сервисами стека **Hadoop**.

Ambari Views – это способ расширения **Ambari**, который позволяет третьим сторонам подключать новые интерфейсы работы и управления сервисами.

5.1 Files View

Files View позволяет пользователю управлять, просматривать и загружать файлы и папки в **HDFS** (Рис.5.1.).



The screenshot shows the Ambari Files View interface. At the top, there is a navigation bar with the Ambari logo and various menu items like Dashboard, Services, Hosts, Alerts, and Admin. Below the navigation bar, there is a search bar and a table listing files and folders in HDFS. The table has columns for Name, Size, Last Modified, Owner, Group, and Permission. The files listed are app-logs, apps, mapred, mr-history, solr, spark2-history, tmp, and user.

Name >	Size >	Last Modified >	Owner >	Group >	Permission
app-logs	--	2018-07-04 11:00	yarn	hadoop	drwxrwxrwx
apps	--	2018-07-04 10:59	hdfs	hdfs	drwxr-xr-x
mapred	--	2018-07-04 10:54	mapred	hdfs	drwxr-xr-x
mr-history	--	2018-07-04 10:54	mapred	hadoop	drwxrwxrwx
solr	--	2018-07-04 11:03	solr	hdfs	drwxr-xr-x
spark2-history	--	2018-07-04 11:23	spark	hadoop	drwxrwxrwx
tmp	--	2018-07-04 11:04	hdfs	hdfs	drwxrwxrwx
user	--	2018-07-04 11:00	hdfs	hdfs	drwxr-xr-x

Рис.5.1.: Пример экрана Files View

5.1.1 Настройка Kerberos для Files View

Перед настройкой **Kerberos** для **Files View** необходимо сначала сконфигурировать **Kerberos** для **Ambari**, настроив для демона **Ambari Server** принципал **Kerberos** и таблицу ключей.

После настройки **Kerberos** для **Ambari** в разделе “Settings” настраиваемого View необходимо ввести свойства, описанные в таблице:

Параметр	Описание	Значение
WebHDFS Username	Имя пользователя, под которым View будет обращаться к HDFS	<code>\${username}</code>
WebHDFS Authorization	Строка аутентификации для доступа к WebHDFS	<code>auth=KERBEROS;proxyuser=ambari-server</code>

При конфигурации **Kerberos** настройка прокси-пользователя должна быть основным значением принципа **Kerberos** для **Ambari Server**. Например, при настройке сервера **Ambari** для принципа `ambari-server@EXAMPLE.COM` значение – `ambari-server`.

5.2 Capacity Scheduler View

Capacity Scheduler помогает оператору **Hadoop** легко настраивать управление рабочей нагрузкой **YARN** для обеспечения многопользовательской и сильно нагруженной обработки.

В **Capacity Scheduler** представлены ресурсы кластера для создания и управления очередями **YARN** (Рис.5.2.).

The screenshot displays the Ambari Capacity Scheduler View for a 'default' queue. The interface is organized into several panels:

- Queue Overview:** Shows the 'default' queue at 100% capacity, with a green progress bar and a checkmark.
- Scheduler Configuration:** Includes settings for Maximum Applications (10000), Maximum AM Resource (20%), Node Locality Delay (40), Calculator (Default Resource Calculator), Queue Mappings, and Queue Mappings Override (Disabled).
- Capacity Section:** Features a 'Level Total' bar at 100% and two sliders for 'Capacity' and 'Max Capacity', both set to 100%.
- Access Control and Status:** Shows the queue is 'Running'. It lists 'Administer Queue' as 'Anyone', 'Effective Administrators' as '* Anyone', 'Submit Applications' as 'Anyone', and 'Effective Users' as '* Anyone'.
- Resources Section:** Includes 'User Limit Factor' (1), 'Minimum User Limit' (100%), 'Maximum Applications' (Inherited), 'Maximum AM Resource' (Inherited), and an 'Ordering Policy' dropdown.

Рис.5.2.: Пример экрана Capacity View

5.3 Tez View

Tez View предоставляет интерфейс мониторинга задач **Tez**. Используя **Tez View**, можно оптимизировать и ускорить отдельные SQL-запросы или задачи **Pig** с целью получения максимальной производительности в многопользовательской среде **Hadoop** (Рис.5.3).

The screenshot shows the Tez View interface in Ambari. At the top, there's a navigation bar with 'ARENADATA Ambari ADH' and '0 ops 0 alerts'. Below that, a breadcrumb 'Home / All DAGs' and 'Version 0.9.0' are visible. The main content area has a 'Refresh' button and a table of DAGs. The table has columns: Dag Name, Id, Submitter, Status, Progress, Start Time, End Time, Duration, and Application. Two rows are shown, both with 'SUCCEEDED' status and 100% progress. Below the table, there's a 'Load Counters' button and pagination controls.

Dag Name	Id	Submitter	Status	Progress	Start Time	End Time	Duration	Applic
INSERT INTO FOO ...	dag_153069089281...	admin	✓ SUCCEEDED	100%	04 Jul 2018 11:27:32	04 Jul 2018 11:27:39	7s 193ms	applica
OrderedWordCount	dag_153069089281...	ambari-qa	✓ SUCCEEDED	100%	04 Jul 2018 11:01:24	04 Jul 2018 11:01:33	9s 216ms	applica

Рис.5.3.: Пример экрана Tez View

5.4 Hive View

Hive View позволяет пользователю писать и выполнять SQL-запросы в кластере. Показывает историю всех запросов **Hive**, выполняемых на кластере, независимо от того, выполняются ли они из представления **Hive** или из другого источника, такого как **JDBC/ODBC** или **CLI**.

Hive View также дает графическое представление плана выполнения запроса. Это помогает пользователю отлаживать запрос для корректной настройки производительности.

Представление интегрируется с **Tez View**, позволяет пользователю отлаживать любую задачу **Tez**, в том числе отслеживать ход выполнения задачи (будь то **Hive** или **Pig**) во время работы (Рис.5.4).

5.4.1 Настройка HDFS для Hive View

Hive View хранит метаданные пользователя в **HDFS** и по умолчанию используется `/user/${username}`, где `${username}` – это имя пользователя, вошедшего в данный момент в систему и получающего доступ к представлению **Hive**. Для корректной работы **Hive View** необходимо настроить домашний каталог пользователя в **HDFS**.

Important: Поскольку многие пользователи используют по умолчанию администратора Ambari для начала работы с кластером необходимо создать папку `/user/admin` в **HDFS** прежде, чем использовать представление **Hive**

Для того, чтобы создать пользовательские каталоги в **HDFS** для каждого пользователя **Hive View** необходимо:

- Подключиться к хосту в кластере, который включает в себя клиент **HDFS**;

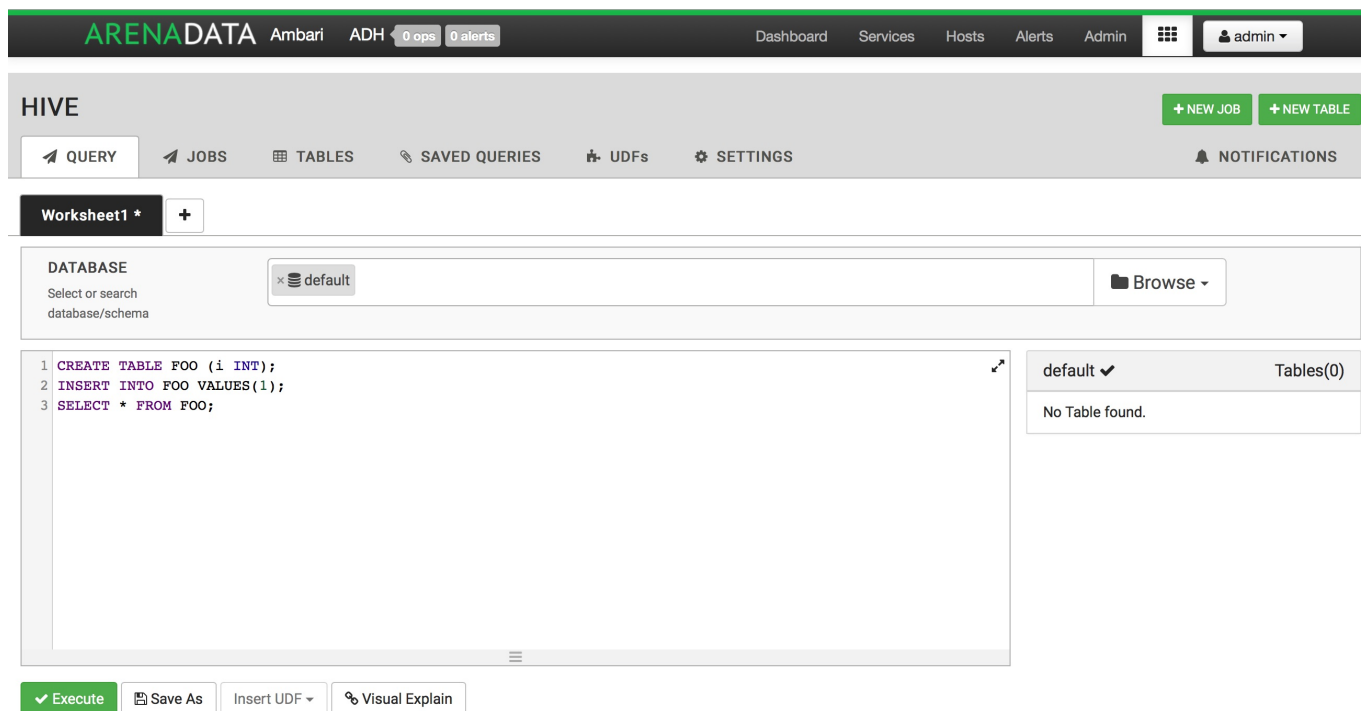


Рис.5.4.: Пример экрана Hive View

- Переключиться на пользователя системной учетной записи hdfs:

```
su - hdfs
```

- Используя клиент HDFS, создать каталог HDFS для пользователя. Например, если имя пользователя *admin*, создать каталог следующим образом:

```
hadoop fs -mkdir /user/admin
```

- Установить права на вновь созданный каталог. Например, если имя пользователя *admin*, сделать данного пользователя владельцем каталога:

```
hadoop fs -chown admin:hadoop /user/admin
```

5.4.2 Настройка SSL для Hive View

Перед настройкой SSL-протокола для **Hive View** необходимо сначала настроить Truststore для **Ambari Server**, импортировав все необходимые ключи в Truststore.

После настройки SSL для **Ambari** в разделе “Settings” настраиваемого View необходимо ввести свойство:

- *Hive Session Parameters* – дополнительные параметры, передаваемые в рамках сессии Hive:

```
sslTrustStore=/path_to_ambari_truststore/ambari_truststore_name.jks;trustStorePassword=*****
```

5.5 Pig View

Pig View аналогичен **Hive View**. Позволяет писать и запускать скрипт **Pig**. Поддерживает сохранение скриптов, а также загрузку и использование существующих **UDF** в скриптах (Рис.5.5.).

The screenshot shows the Ambari Pig View interface. At the top, there is a navigation bar with 'ARENA DATA Ambari' and 'ADH' (0 ops, 0 alerts). The main interface has a left sidebar with 'Scripts', 'UDFs', and 'History' sections. The 'Scripts' section is active, showing a 'Foo' script with options to 'Save', 'Copy', and 'Delete'. The main area displays the Pig script 'Foo' with the following code:

```
1 pv_by_industry = GROUP profile_view by viewee_industry_id
2 pv_avg_by_industry = FOREACH pv_by_industry
3 GENERATE group as viewee_industry_id, AVG(profile_view) AS average_pv;
```

Below the script, there are 'Arguments' and 'Execute' buttons. The 'Arguments' section shows a message: 'This pig script has no arguments defined.' and a '+ Add' button.

Рис.5.5.: Пример экрана Pig View

Important: Помимо приведенных пользовательских представлений существует растущая экосистема Ambari User Views, разрабатываемая сообществом