Arenadata™ Hadoop

Версия - v2.0.0

Руководство администратора по работе с кластером Arenadata Hadoop

2019 Arenadata™

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.

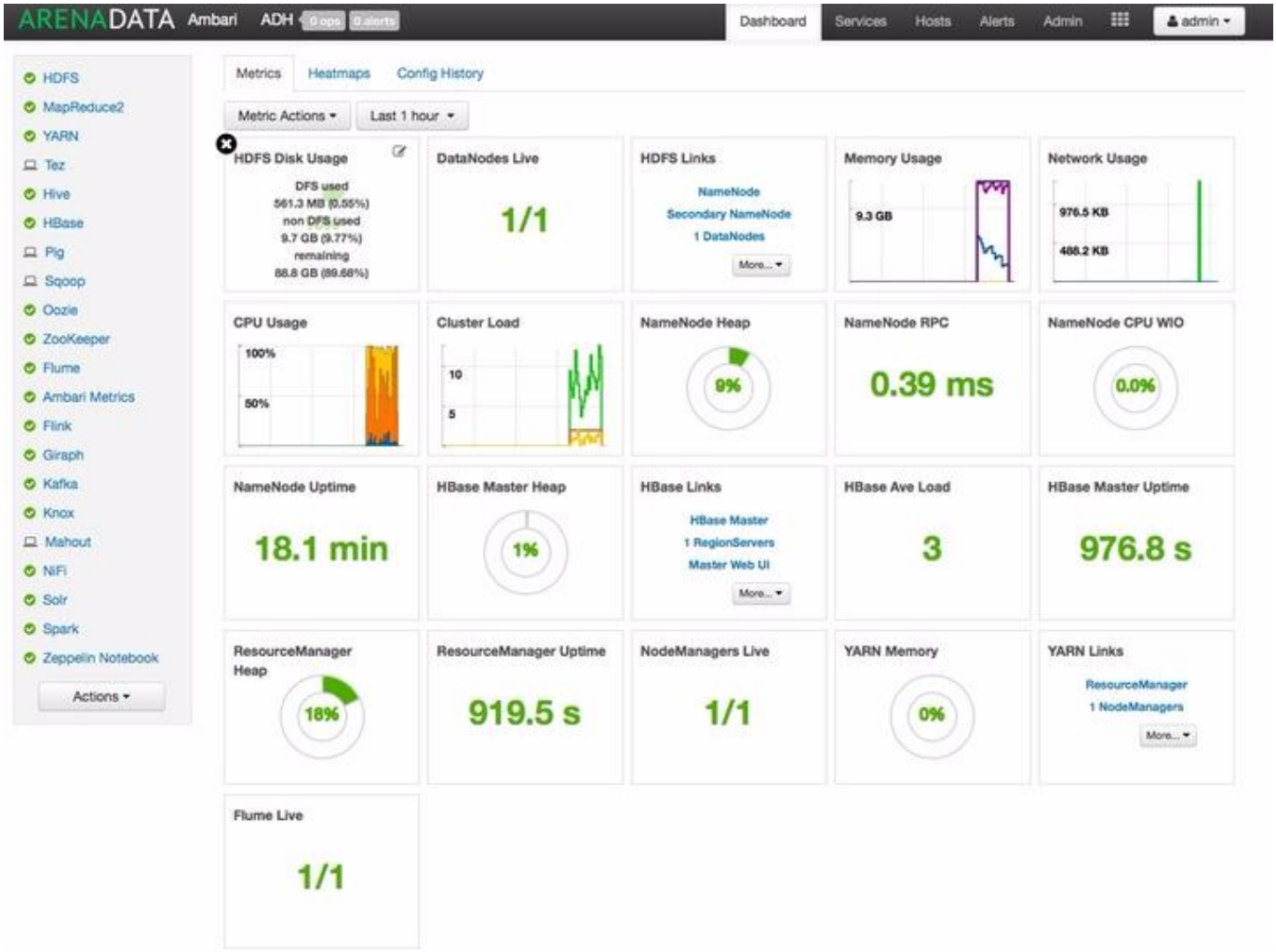


Рис.1.1.: Главная страница Ambari

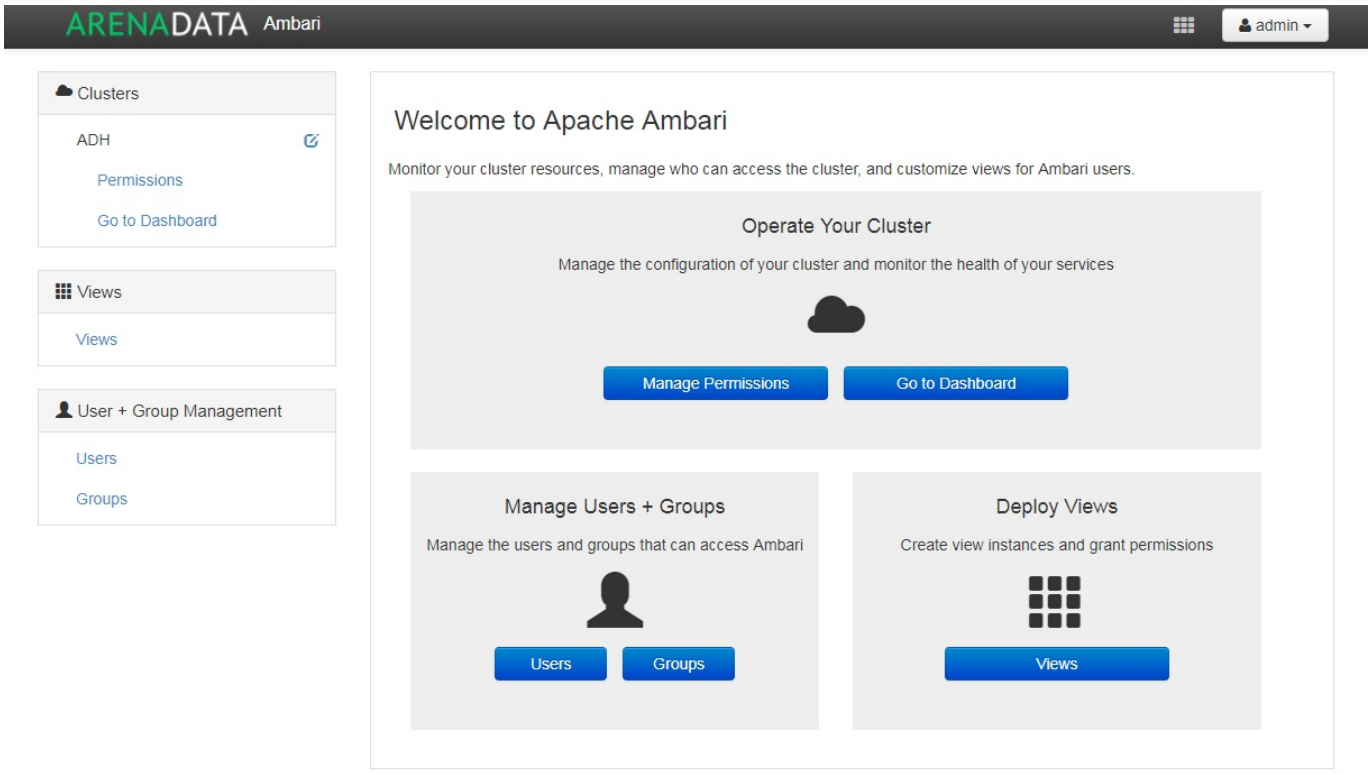


Рис.1.2.: Экран настроек Ambari

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.

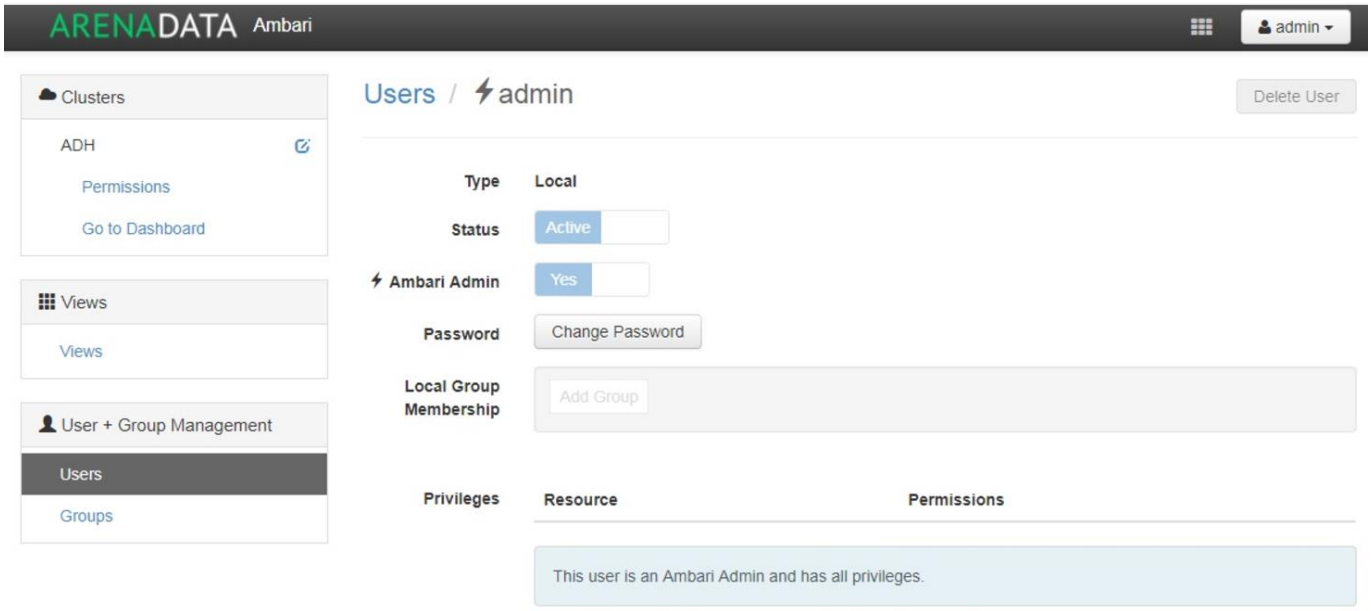


Рис.1.3.: Параметры администратора

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.

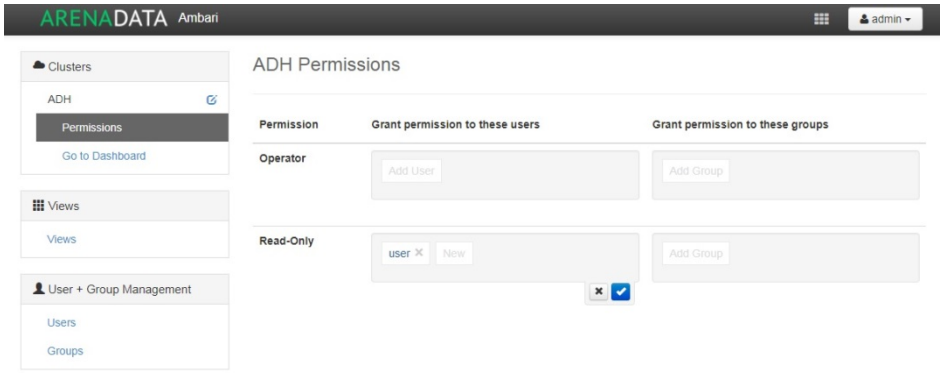


Рис.2.1.: Права доступа к кластеру

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.).

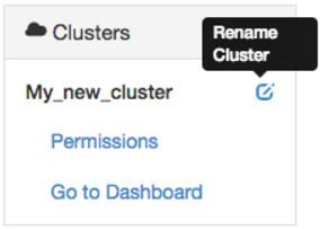


Рис.2.2.: Значок “Rename Cluster”

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.).

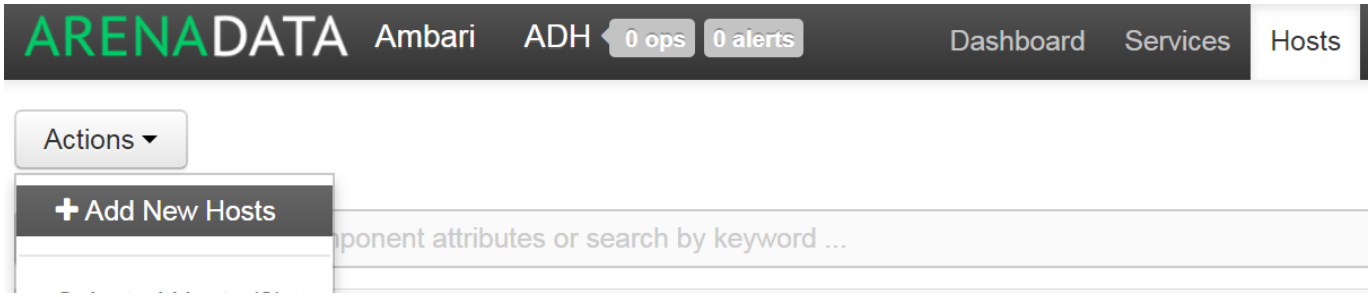


Рис.2.3.: Добавление нового узла

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 добавлемого 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.).

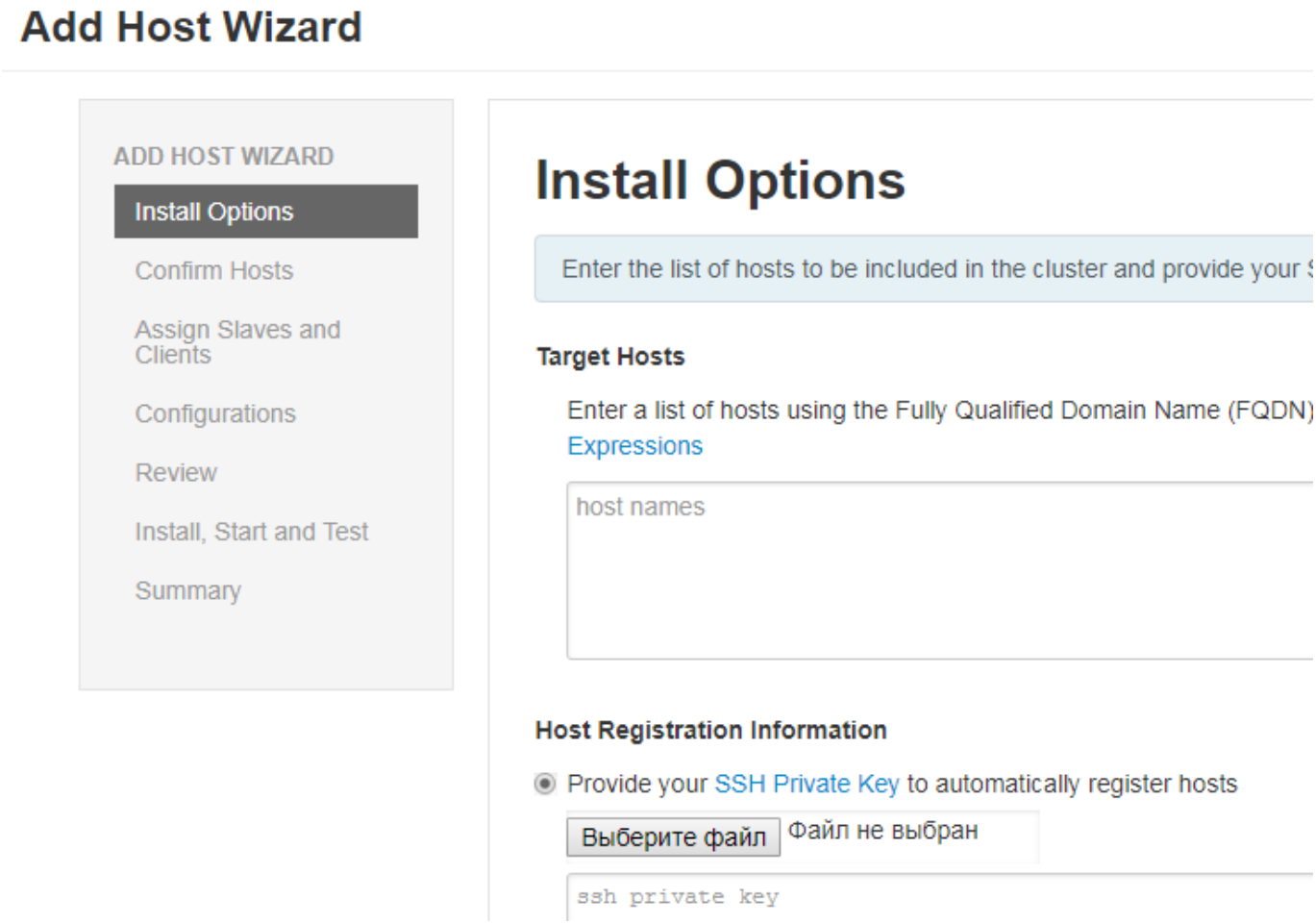


Рис.2.4.: Мастер добавления узла

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.

Таблица3.1.: 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.

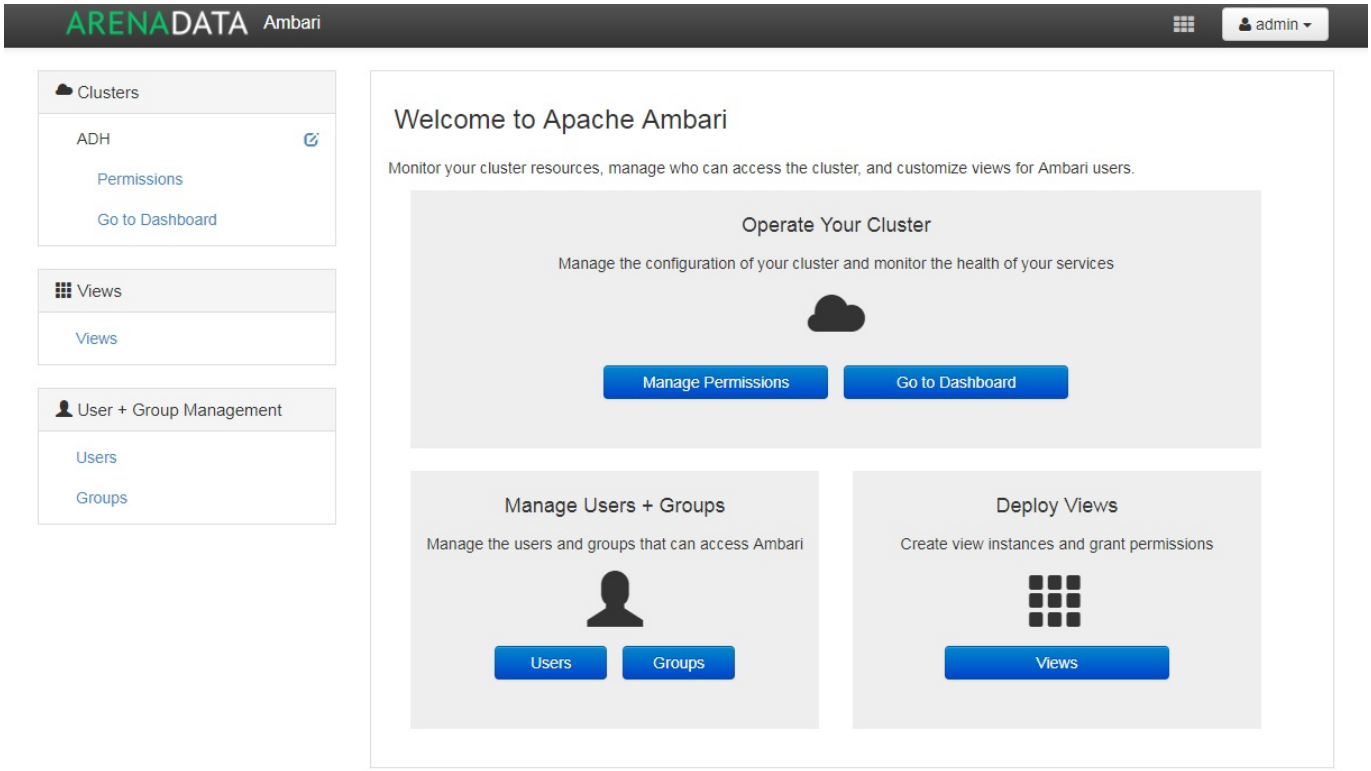


Рис.3.1.: Экран настроек Ambari

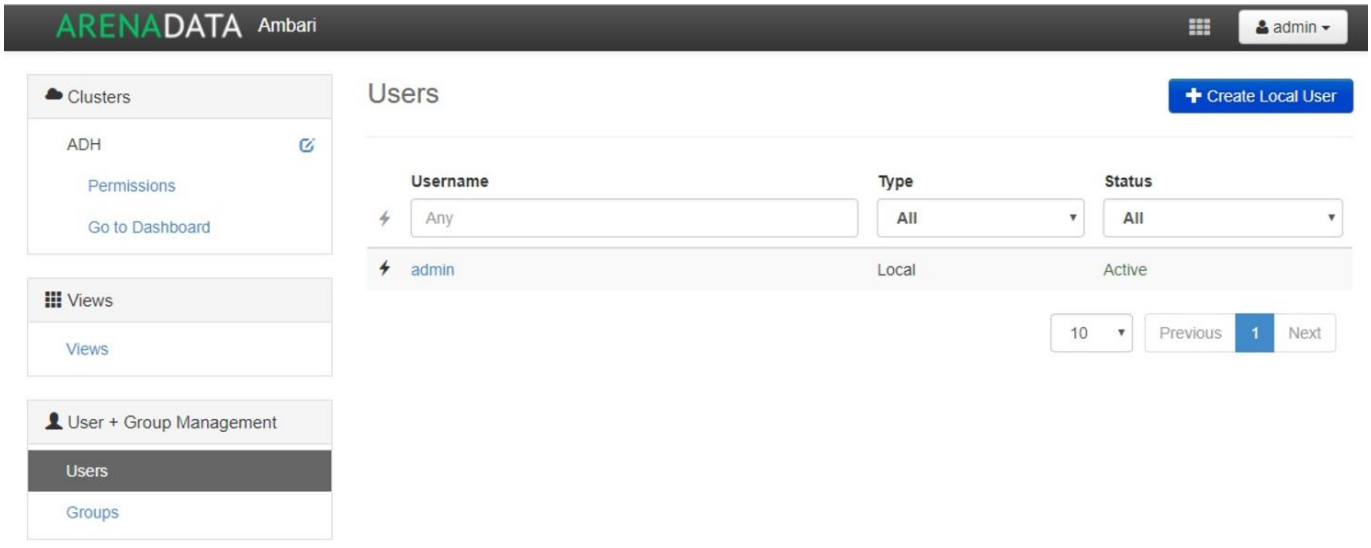


Рис.3.2.: Пользователи Ambari

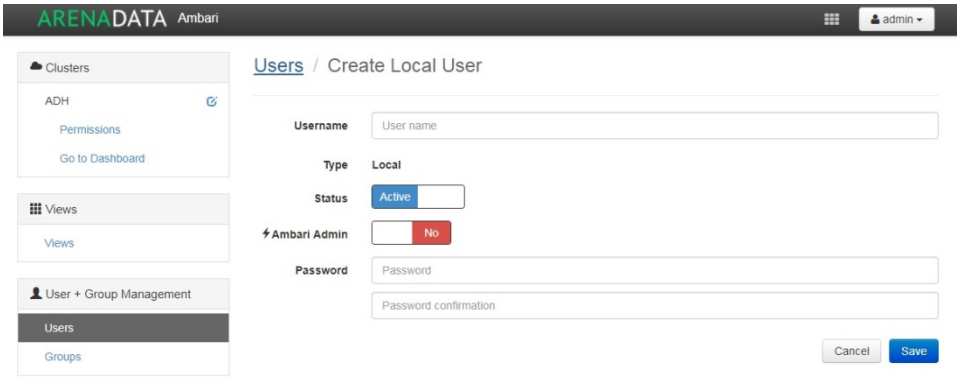


Рис.3.3.: Создание локального пользователя

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.);

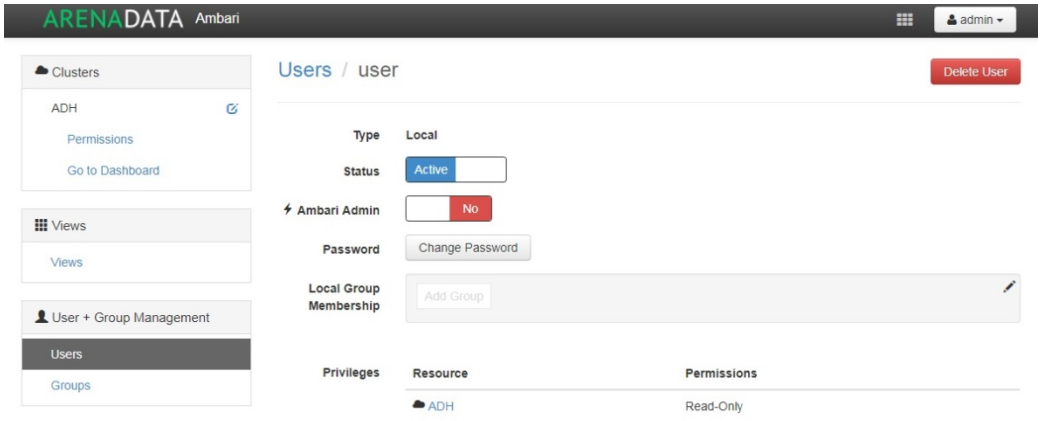


Рис.3.4.: Параметры пользователя

- 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.

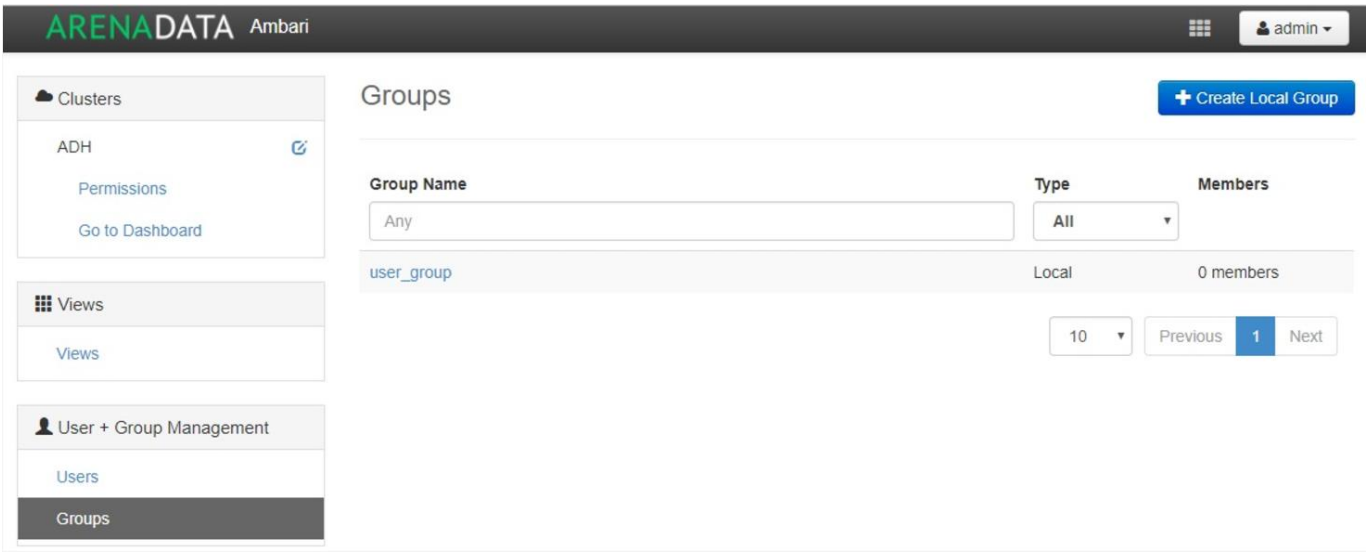


Рис.3.5.: Группы пользователей

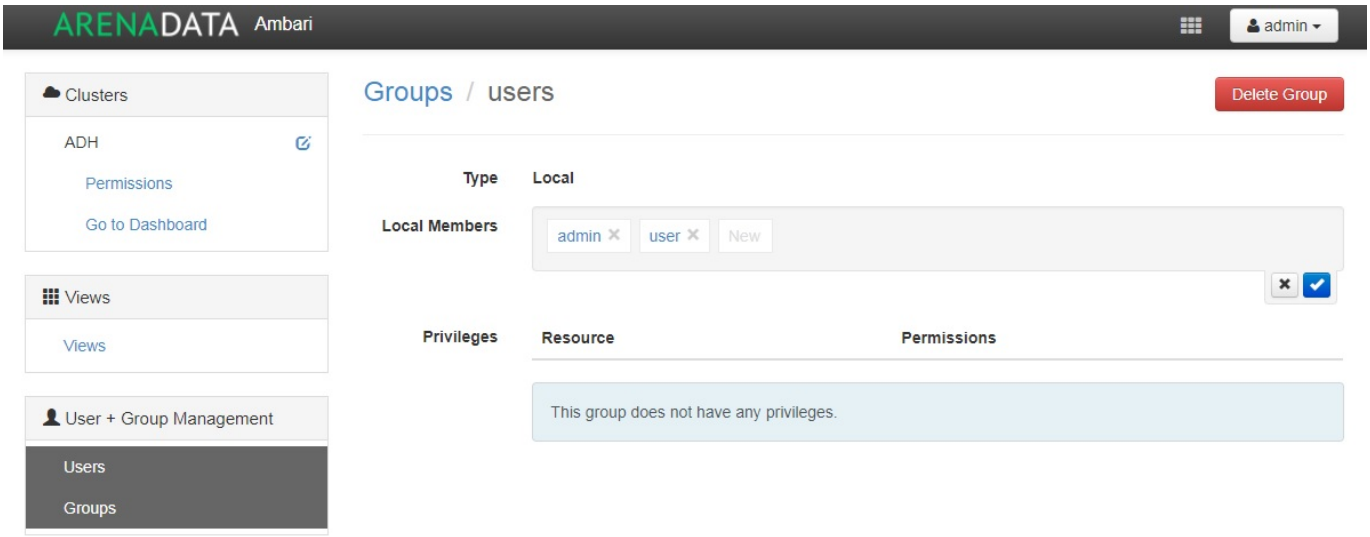


Рис.3.6.: Параметры группы пользователей

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.

Таблица4.1.: Порты HDFS

Сервис	Характеристика
NameNode WebUI	Сервер – Master Nodes • Порт по умолчанию – 50070 – 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 – Протокол: https – Описание: Безопасный http-сервис – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.https.address
NameNode metadata service	Сервер – Master Nodes • Порт по умолчанию – 8020/9000 – 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	Сервер – Все Slave Nodes • Порт по умолчанию – 50075 – 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 – Протокол: https – Описание: Безопасный http-сервис – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: dfs.datanode.https.address • Порт по умолчанию – 50010 – Описание: Передача данных – Параметр конфигурации: dfs.datanode.address • Порт по умолчанию – 50020 – Протокол: IPC – Описание: Операции с метаданными – Доступ пользователям: Нет – Параметр конфигурации: dfs.datanode.ipc.address
Secondary NameNode	Сервер – Secondary NameNodes • Порт по умолчанию – 50090 – Протокол: http – Описание: Чекпоинт для метаданных NameNode – Доступ пользователям: Нет – Параметр конфигурации: dfs.secondary.http.address
HDFS HFTP	Сервер – Master NameNodes • Порт по умолчанию – 50470 – Протокол: https – Описание: Порт для доступа к HFTP файловой системы – Параметр конфигурации: dfs.https.port
JournalNode Web UI	Сервер – Master NameNodes • Порт по умолчанию – 8480 – 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 – 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.

Таблица4.2.: Порты MapReduce

Сервис	Характеристика
JobTracker WebUI	Сервер – Master Nodes • Порт по умолчанию – 50030 – Протокол: http – Описание: Web-интерфейс для работы JobTracker – Доступ пользователям: Да – Параметр конфигурации: mapred.job.tracker.http.address
JobTracker	Сервер – Master Nodes • Порт по умолчанию – 8021 – Протокол: IPC – 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 • Порт по умолчанию – 50060 – 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	• Порт по умолчанию – 51111 – Протокол: http – Описание: Web-интерфейс для истории заданий – Доступ пользователям: Да – Параметр конфигурации: mapreduce.history.server.http.address
MapReduce Shuffle Port	• Порт по умолчанию – 13562 – 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.

Таблица4.3.: Порты YARN

Сервис	Характеристика
ResourceManager WebUI	Сервер – Master Nodes • Порт по умолчанию – 8088 – Протокол: http – Описание: Web-интерфейс для Resource Manager – Доступ пользователям: Да – Параметр конфигурации: yarn.resourcemanager.webapp.address
ResourceManager	Сервер – Master Nodes (ResourceManager Node) • Порт по умолчанию – 8050 – 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 – 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 – Протокол: http – Описание: Прокси для Resource Manager – Доступ пользователям: Да – Параметр конфигурации: yarn.web-proxy.address • Порт по умолчанию – 8141 – Протокол: http – Описание: Адрес планировщика – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.resourcemanager.admin.address
Scheduler	Сервер – Master Nodes • Порт по умолчанию – 8030 – Описание: Адрес планировщика – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.resourcemanager.scheduler.address

Сервис	Характеристика
NodeManager	Сервер – Master Nodes • Порт по умолчанию – 45454 – Протокол: http – Описание: Адрес NodeManager – Параметр конфигурации: yarn.nodemanager.address Сервер – Slave Nodes • Порт по умолчанию – 8040 – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.localizer.address • Порт по умолчанию – 8042 – Протокол: http – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.webapp.address • Порт по умолчанию – 8044 – Протокол: https – Описание: NodeManager – Параметр конфигурации: yarn.nodemanager.webapp.https.address
Timeline Server	Сервер – Master Nodes • Порт по умолчанию – 10200 – Протокол: http – Описание: Адрес Timeline Server – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.timeline-service.address • Порт по умолчанию – 8188 – Протокол: http – Описание: Адрес Timeline Server Webapp – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: yarn.timeline-service.webapp.address • Порт по умолчанию – 8190 – 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.

Таблица4.4.: Порты Hive

Сервис	Характеристика
Hive Server2	Сервер – Hive Server машина • Порт по умолчанию – 10000 – 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 – 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	Сервер – Master Nodes • Порт по умолчанию – 8021 – Протокол: IPC – Описание: Для публикации заданий – Доступ пользователям: Да (все клиенты, которым требуется запуск MR, Hive, Pig. Задачи, использующие HCatalog)
Hive Web UI	Сервер – Hive Server машина • Порт по умолчанию – 9999 – Протокол: thrift – Описание: WebUI для Hive – Доступ пользователям: Да – Параметр конфигурации: hive.hwi.listen.port • Порт по умолчанию – 9933 – 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.

Таблица4.5.: Порты WebHCat

Сервис	Характеристика
WebHCat Server	Сервер – WebHCat Server машина • Порт по умолчанию – 50111 – 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.

Таблица4.6.: Порты HBase

Сервис	Характеристика
HMaster	Сервер – Master Nodes (HBase Master Node и back-up HBase Master node) • Порт по умолчанию – 60000 – Доступ пользователям: Да – Параметр конфигурации: hbase.master.port
HMaster Info Web UI	Сервер – Master Nodes (HBase Master Node и back-up HBase Master node) • Порт по умолчанию – 60010 – Протокол: http – Описание: Порт для HBase Master UI – Доступ пользователям: Да – Параметр конфигурации: hbase.master.info.port
Region Server	Сервер – Все Slave Nodes • Порт по умолчанию – 60020 – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: hbase.regionserver.port • Порт по умолчанию – 60030 – Протокол: http – Доступ пользователям: Да (администратор, разработчик, поддержка) – Параметр конфигурации: hbase.regionserver.info.port
HBase Thrift Server	Сервер – Все Thrift Server • Порт по умолчанию – 9090 – Описание: Порт, используемый HBase Thrift-сервером – Доступ пользователям: Да
HBase Thrift Server Web UI	Сервер – Все Thrift Server • Порт по умолчанию – 9090 – 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.

Таблица4.7.: Порты Oozie

Сервис	Характеристика
Oozie	Сервер – Oozie Server • Порт по умолчанию – 11000 – Протокол: TCP – Описание: Порт, используемый для запуска Oozie Server – Доступ пользователям: Да – Параметр конфигурации: OOZIE_HTTP_PORT in oozie_env.sh • Порт по умолчанию – 11001 – Протокол: TCP – Описание: Порт, используемый админ-консолью для запуска Oozie Server – Доступ пользователям: Нет – Параметр конфигурации: OOZIE_ADMIN_PORT in oozie_env.sh • Порт по умолчанию – 11443 – Протокол: TCP – Описание: Безопасный порт, используемый для запуска Oozie Server – Доступ пользователям: Да – Параметр конфигурации: OOZIE_HTTPS_PORT in oozie_env.sh

4.8 Порты NiFi

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

Таблица4.8.: Порты NiFi

Сервис	Характеристика
NiFi	Сервер – NiFi сервер • Порт по умолчанию – 9090 – Протокол: http – Описание: NiFi Server http port – Доступ пользователям: Да • Порт по умолчанию – 9091 – Протокол: https – Описание: NiFi Server https port – Доступ пользователям: Да • Порт по умолчанию – 10443 – Описание: Certificate Authority – Доступ пользователям: Нет • Порт по умолчанию – 8022 – Описание: Удаленный input-порт – Доступ пользователям: Да – Параметр конфигурации: nifi.remote.input.socket.port • Порт по умолчанию – 8021 – Описание: Node protocol-порт – Доступ пользователям: Да – Параметр конфигурации: nifi.cluster.node.protocol.port • Порт по умолчанию – 8070 – Протокол: http – Описание: NiFi http port – Доступ пользователям: Да • Порт по умолчанию – 9088 – 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.

Таблица4.9.: Порты Zookeeper

Сервис	Характеристика
Zookeeper Server	Сервер – Все Zookeeper Nodes • Порт по умолчанию – 2181 – Протокол: http – Описание: Сервис доступа к Zookeeper Server/Quorum – Доступ пользователям: Да – Параметр конфигурации: zookeeper.port/clientPort • Порт по умолчанию – 2888 – Описание: Порт используется Zookeeper для взаимодействия компонентов – Доступ пользователям: Нет – Параметр конфигурации: hbase.zookeeper.peerport • Порт по умолчанию – 3888 – Описание: Порт используется Zookeeper для взаимодействия компонентов – Доступ пользователям: Нет – 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.).

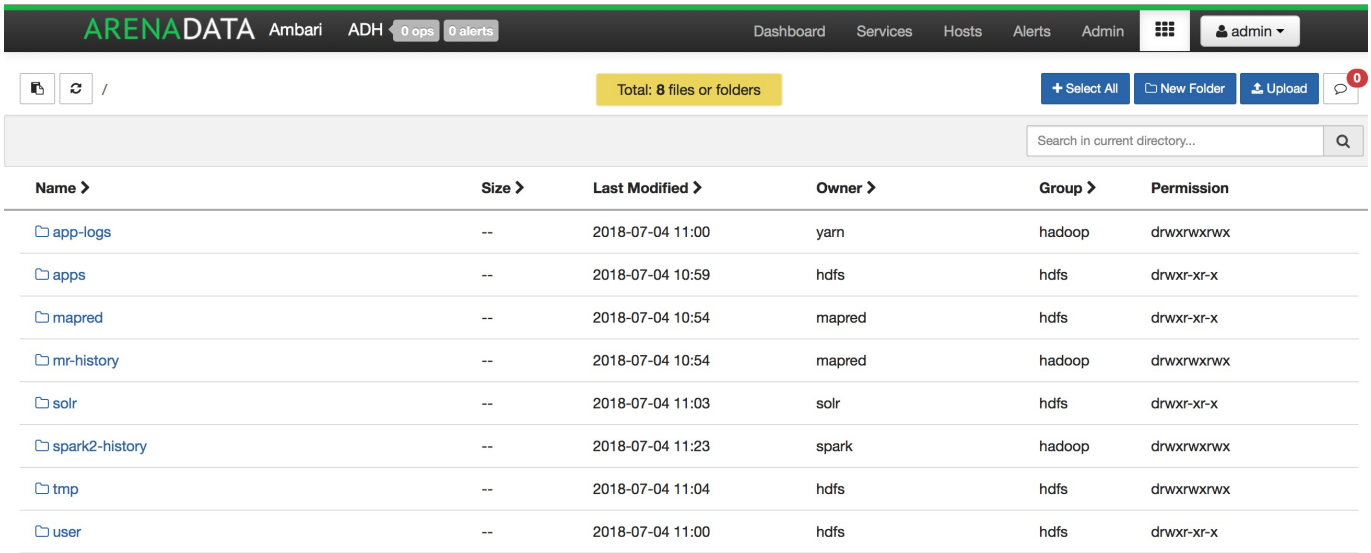


Рис.5.1.: Пример экрана Files View

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	`${username}`
WebHDFS Authorization	Строка аутентификации для доступа к WebHDFS	`auth=KERBEROS;proxyuser=ambari-server`

При конфигурации **Kerberos** настройка прокси-пользователя должна быть основным значением принципала **Kerberos** для **Ambari Server**. Например, при настройке сервера **Ambari** для приницпала `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.).

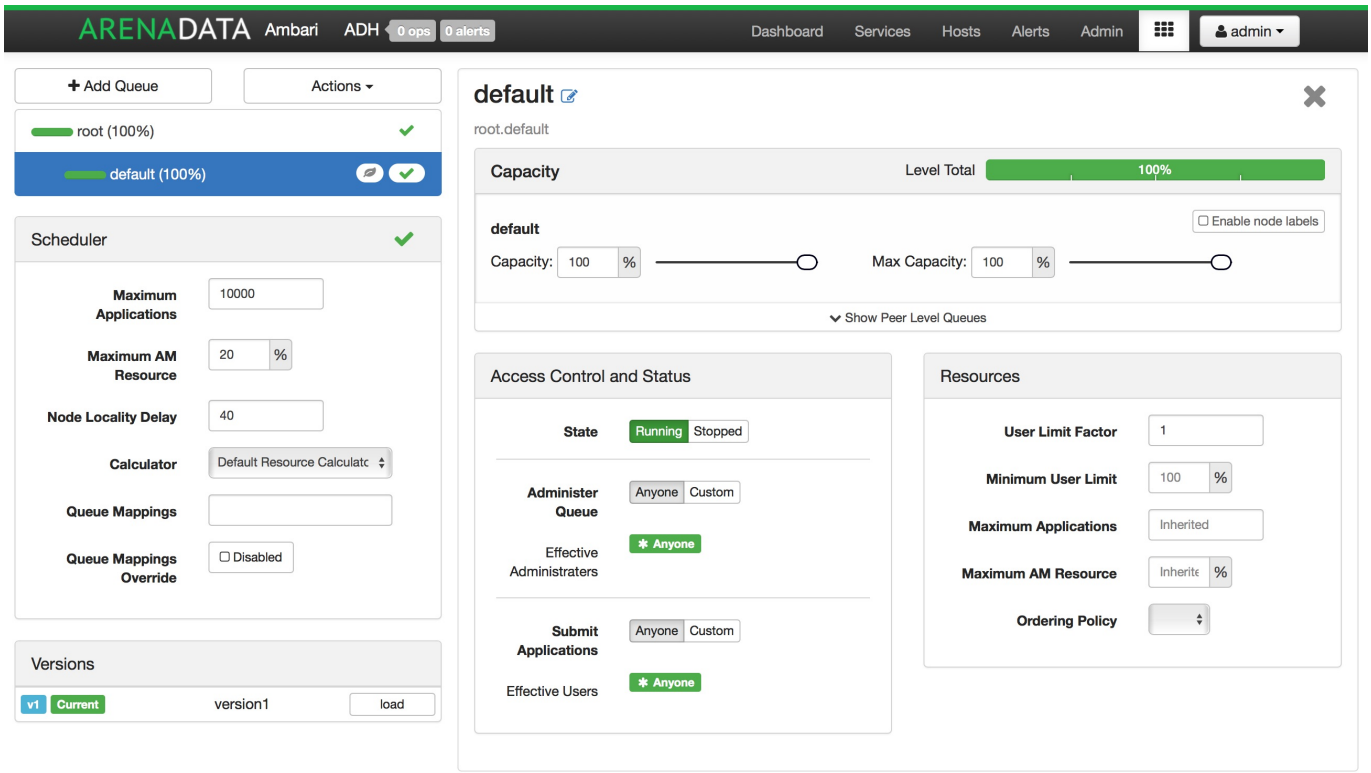


Рис.5.2.: Пример экрана Capacity View

5.3 Tez View

Tez View предоставляет интерфейс мониторинга здадач **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.).

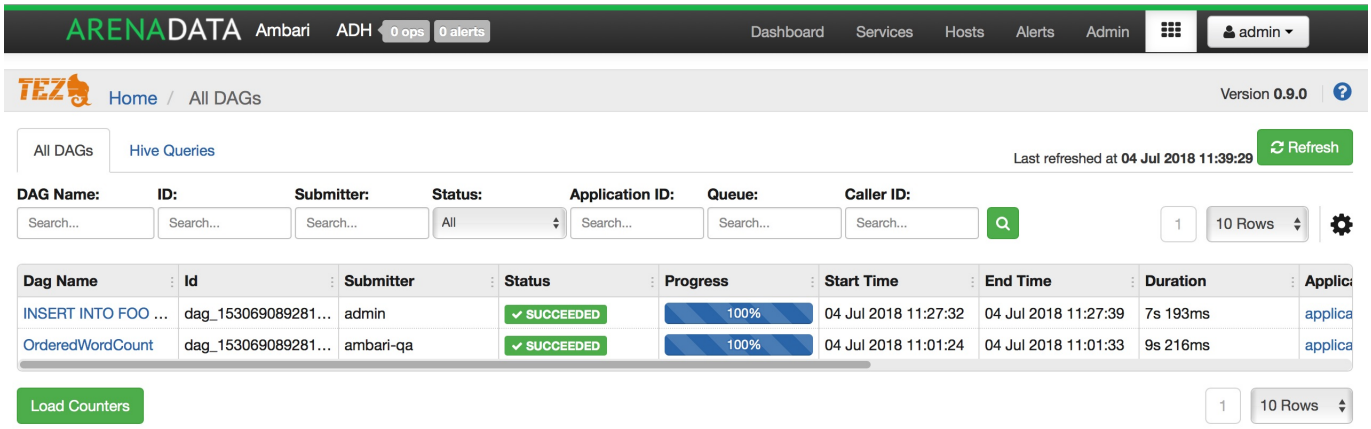


Рис.5.3.: Пример экрана Tez View

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;

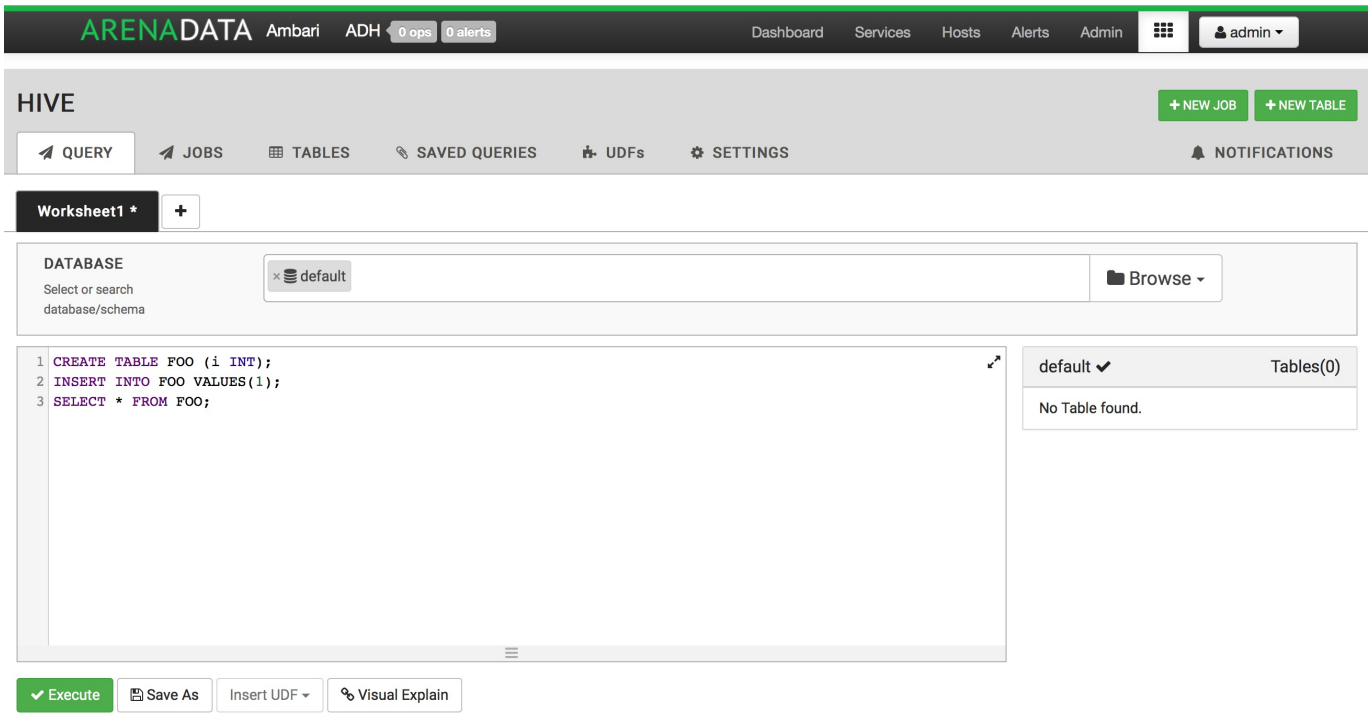


Рис.5.4.: Пример экрана Hive View

- 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.).

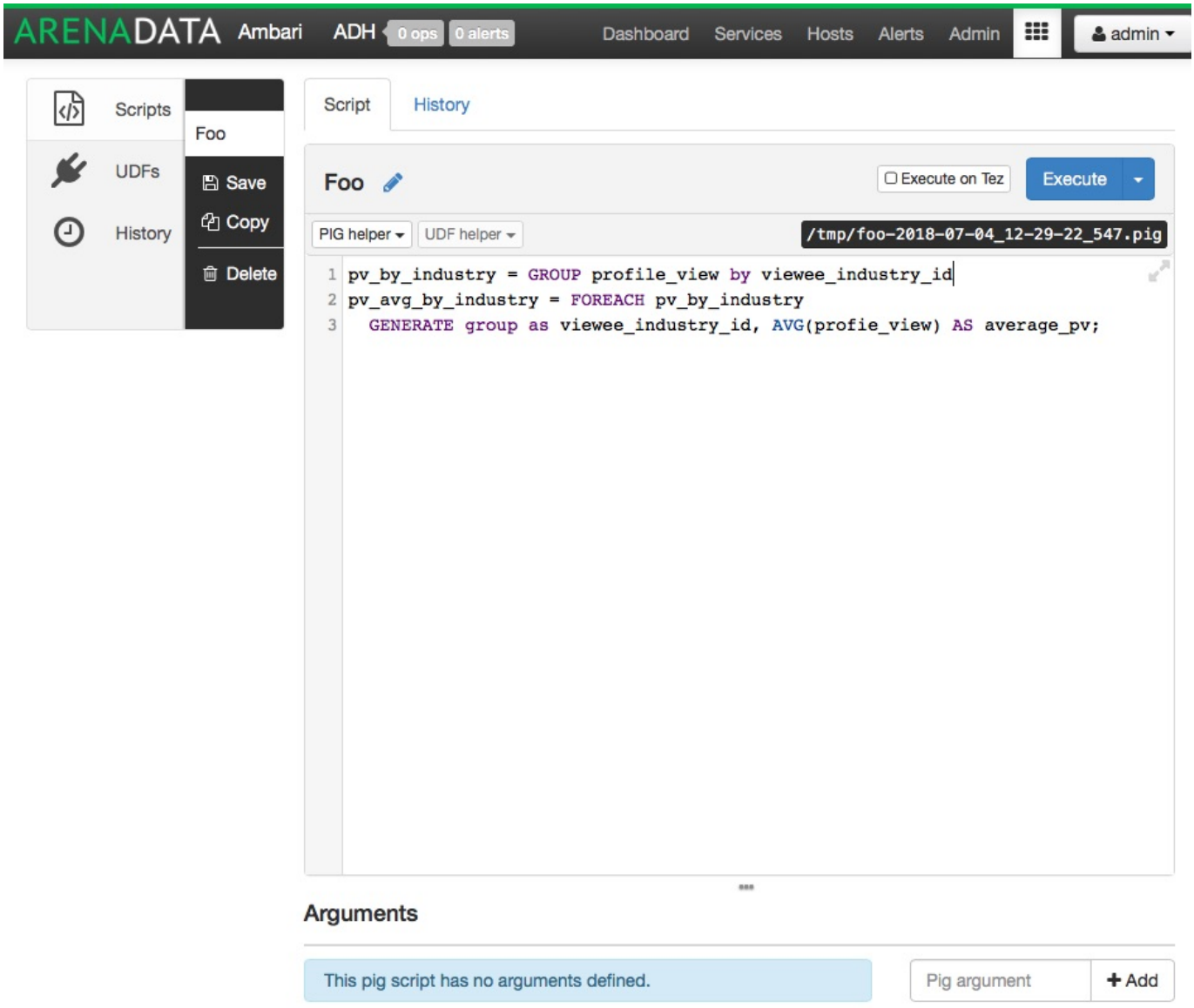


Рис.5.5.: Пример экрана Pig View

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