Arenadata™ Hadoop

Версия - v1.6.1

Использование Ambari View

2019 Arenadata™

Оглавление

1 Files View 3

2 Capacity Scheduler View 4

3 Tez View 5

4 Hive View 6

5 Pig View 7

6 Настройка HDFS для Hive View 9

7 Настройка Kerberos для Files View 10

8 Настройка SSL для Hive View 11

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.

Инструкция может быть полезна администраторам, программистам, разработчикам и сотрудникам подразделений информационных технологий, осуществляющих внедрение сервера.

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

Глава 1

Files View

Files View позволяет пользователю управлять, просматривать и загружать файлы и папки в **HDFS**.

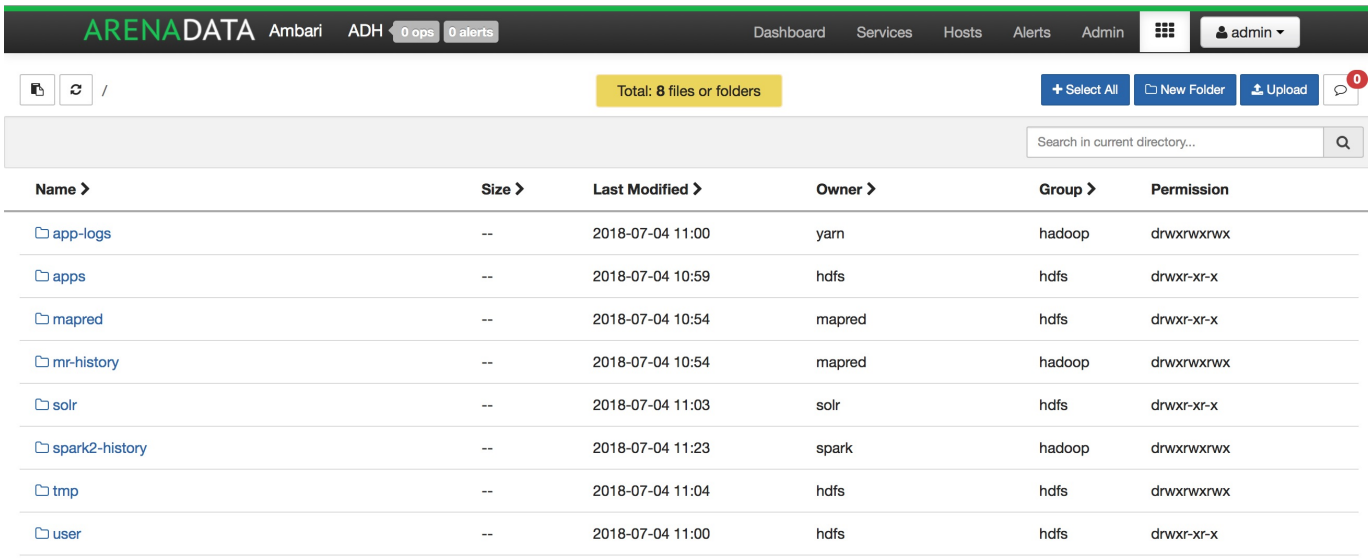


Рис.1.1.: Пример экрана Files View

Глава 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**.

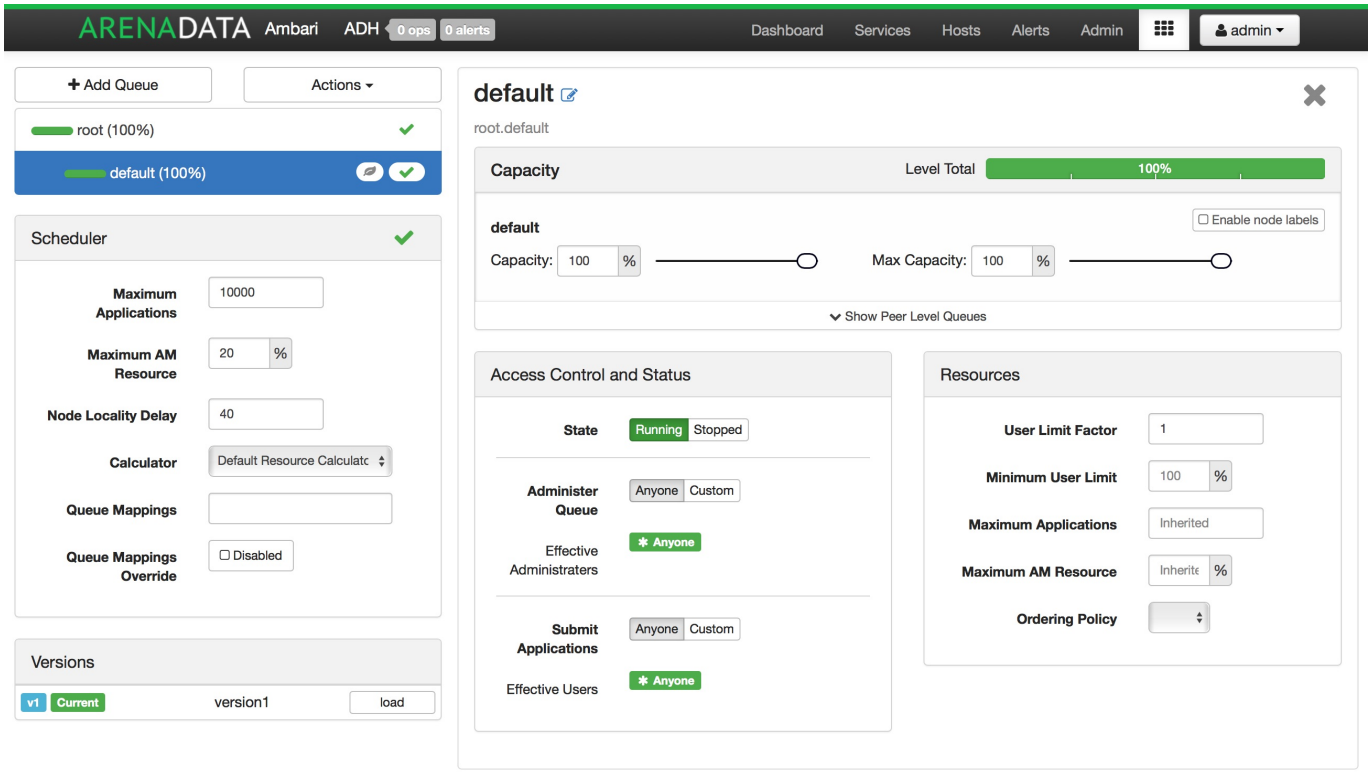


Рис.2.1.: Пример экрана Capacity View

Глава 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**.

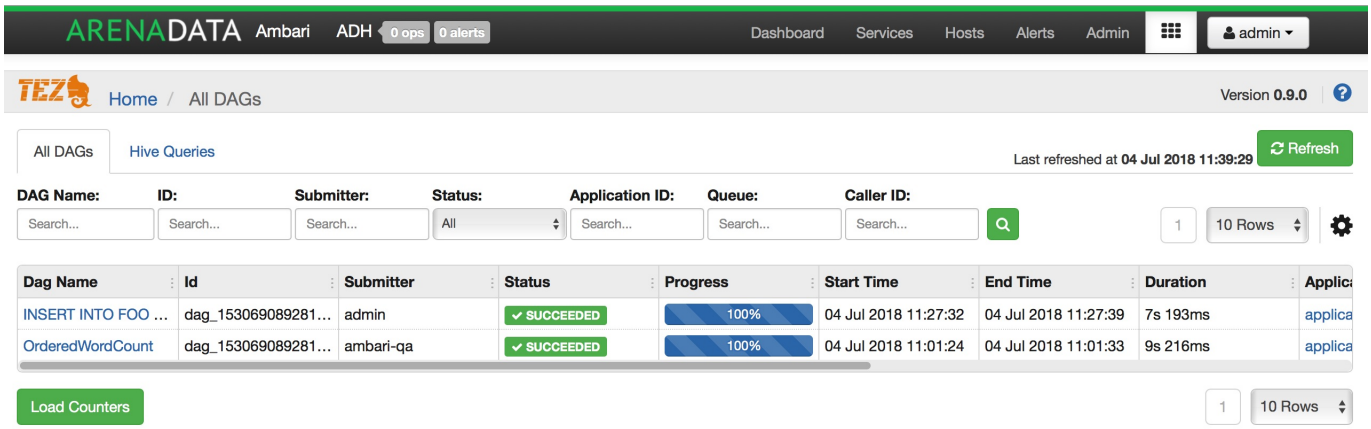


Рис.3.1.: Пример экрана Tez View

Глава 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.

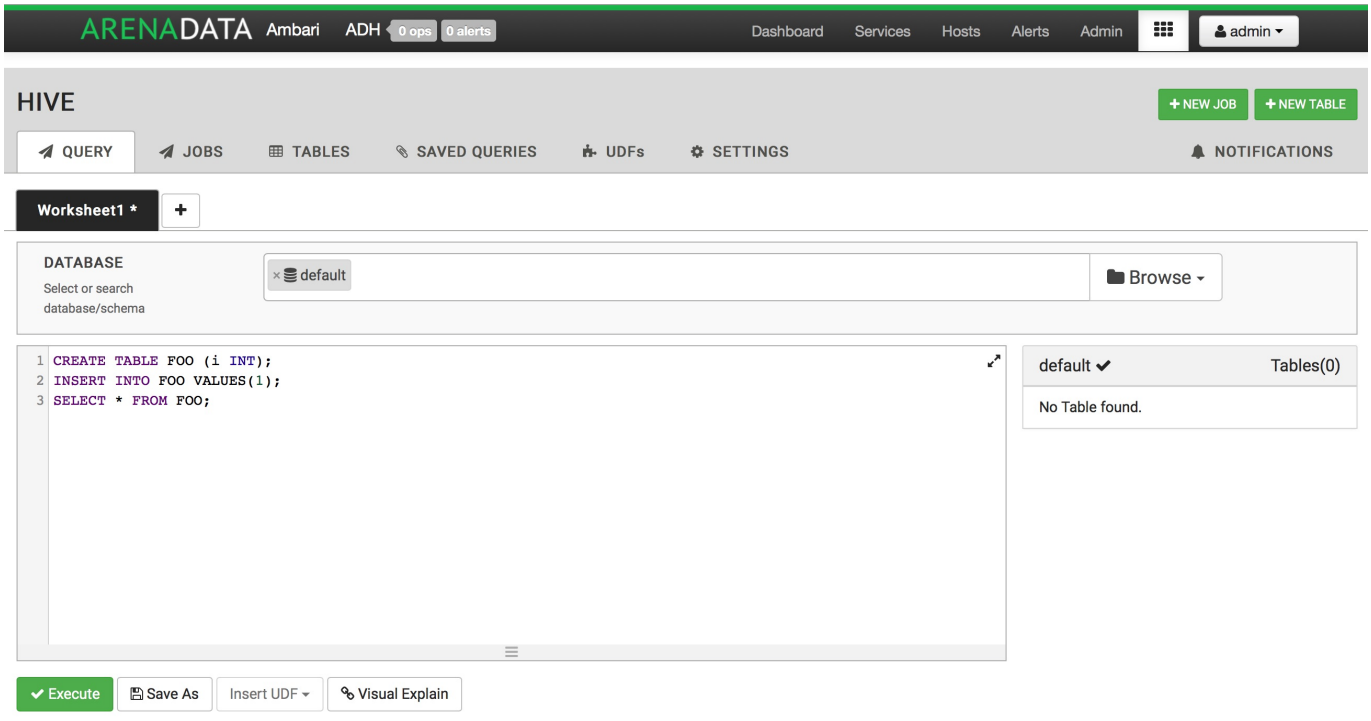


Рис.4.1.: Пример экрана Hive View

Глава 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.

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

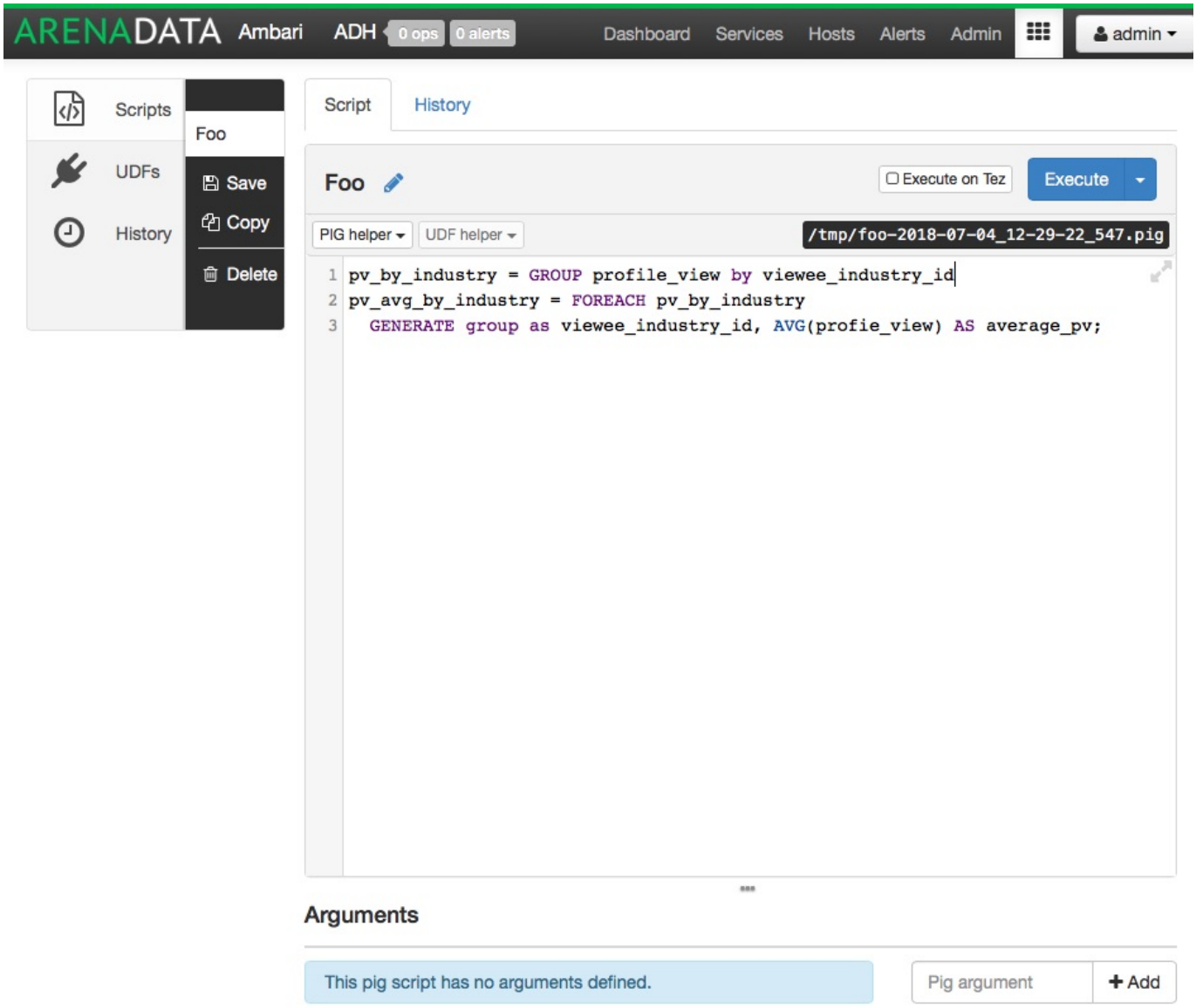


Рис.5.1.: Пример экрана Pig View

Глава 6

Настройка HDFS для Hive View

Hive View хранит метаданные пользователя в HDFS, 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 этого 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.

Чтобы создать пользовательские каталоги в HDFS, для каждого пользователя Hive View:

- подключитесь к хосту в кластере, который включает в себя клиент HDFS;
- переключитесь на пользователя системной учетной записи hdfs:

```
su - hdfs
```

- используя клиент HDFS, создайте каталог HDFS для пользователя. Например, если имя Вашего пользователя admin, то создайте каталог следующим образом:

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

- установите права на вновь созданный каталог. Например, если имя Вашего пользователя admin, то сделаете этого пользователя владельцем каталога:

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

Глава 7

Настройка Kerberos для Files View

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

После настройки Kerberos для Ambari в разделе Settings настраиваемого View введите следующие свойства:

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 настройке Kerberos настройка прокси-пользователя должна быть основным значением принципала Kerberos для Ambari Server. Например, если вы настроили сервер Ambari для приницпала `ambari-server@EXAMPLE.COM`, то это значение будет `ambari-server`.

Глава 8

Настройка SSL для Hive View

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

После настройки SSL для Ambari в разделе Settings настраиваемого View введите следующие свойства:

Параметр	Описание	Значение
Hive Session Parameters	Дополнительные параметры передаваемыее в рамках сессии Hive	`sslTrustStore=/path_to_ambari_truststore/`