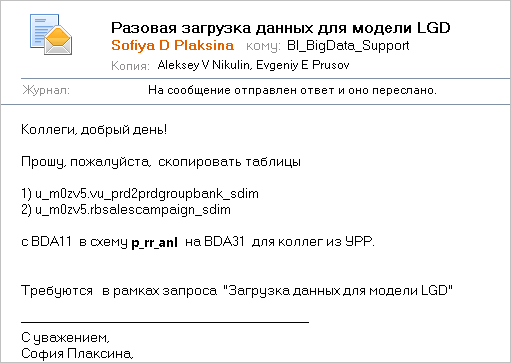
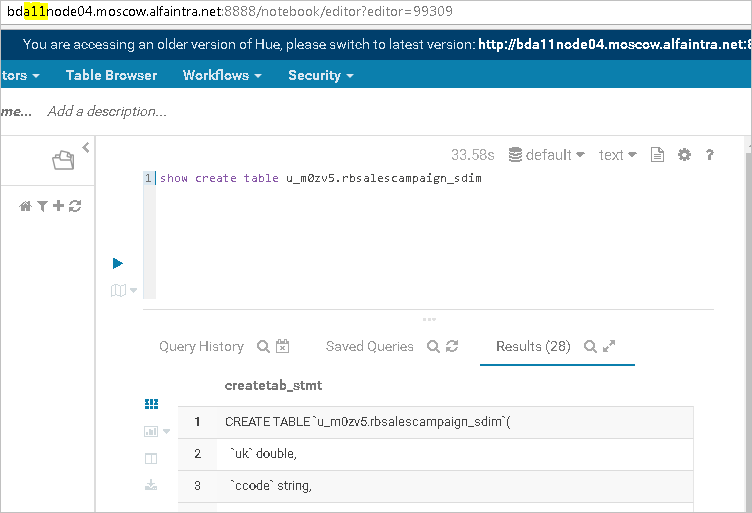
Копирование таблицы с одного кластера на другой

Сначала создаем на том кластере, куда надо скопировать, такую же таблицу.

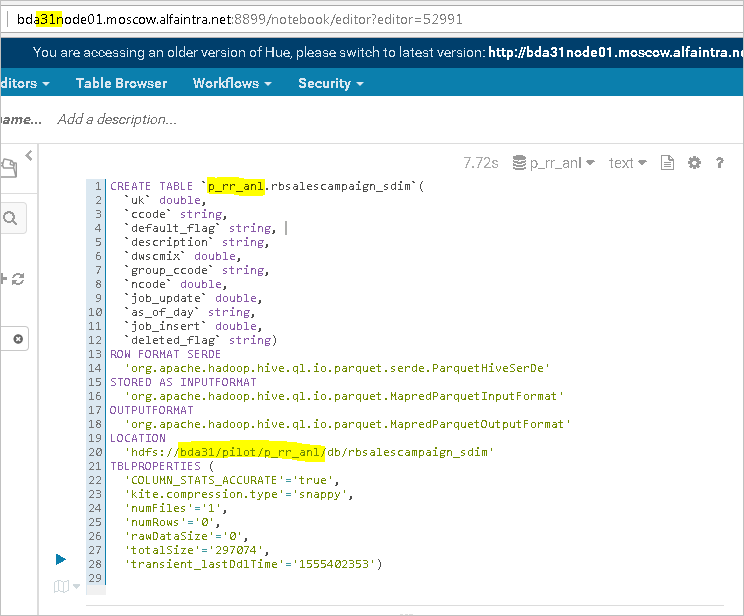
Например, задача такая:



Смотрим как таблица создана на bda**1**1:

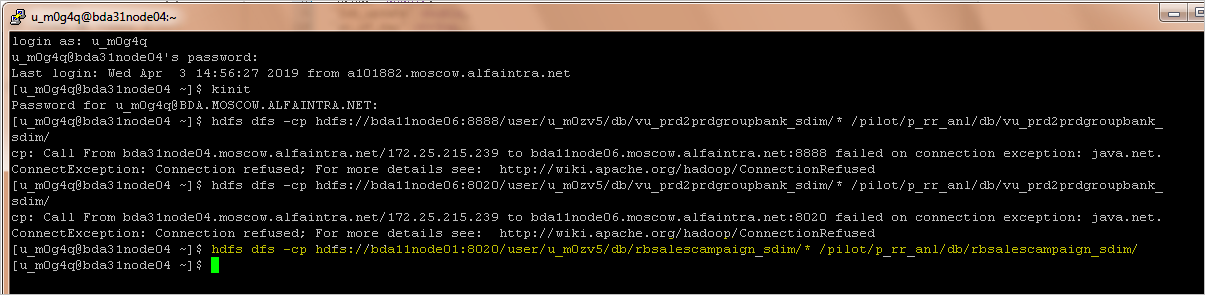


и создаем такую же на bda**3**1, меняя схему и путь в Location:



Дальше копируем файла hdfs такой командой:

**hdfs dfs -cp hdfs://bda11node01:8020/user/u\_m0zv5/db/rbsalescampaign\_sdim/\* /pilot/p\_rr\_anl/db/rbsalescampaign\_sdim/**



При этом если заходим через терминал на bda**3**1, то указываем полный путь до источника на bda**1**1:

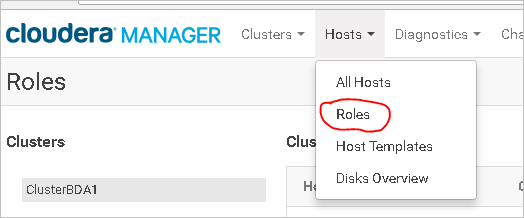
**hdfs://bda11node01:8020**

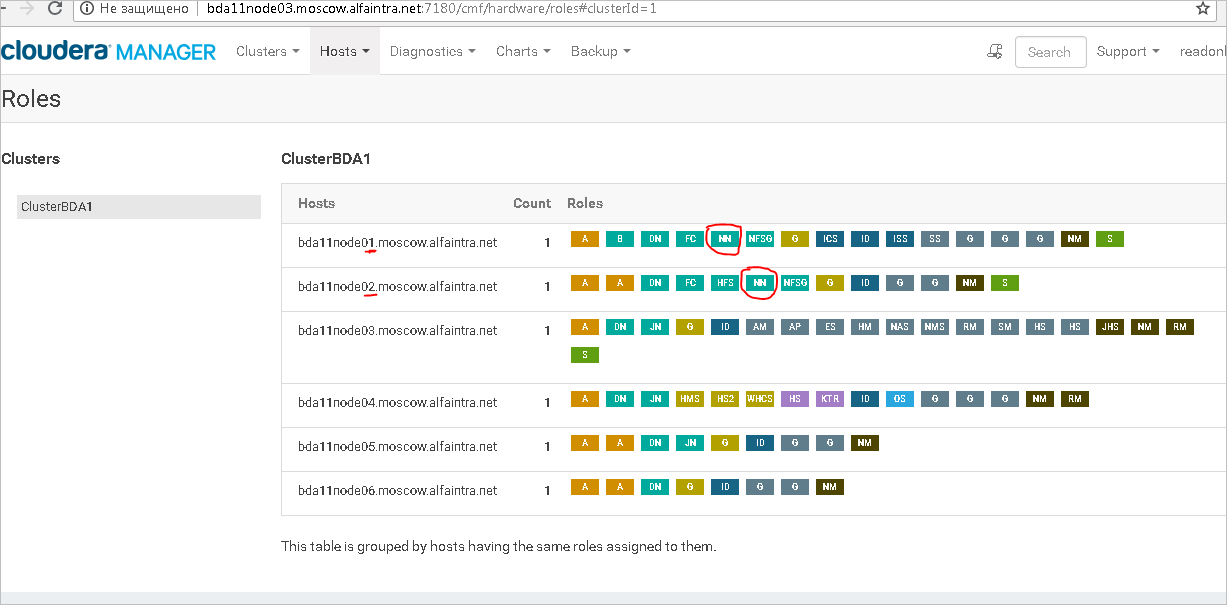
Можно наоборот, зайти на bda**1**1 и указать полный путь таргета :

hdfs dfs -cp /user/u\_m0zv5/db/vu\_prd2prdgroupbank\_sdim/\* **hdfs://bda31node02:8020**/pilot/p\_rr\_anl/db/vu\_prd2prdgroupbank\_sdim/

Важно понять, какая нода на данный момент является NameNode.

Для этого заходим в Cloudera MANEGER - > Hosts -> Roles и смотрим у каких нод(их две) зеленый значок NN

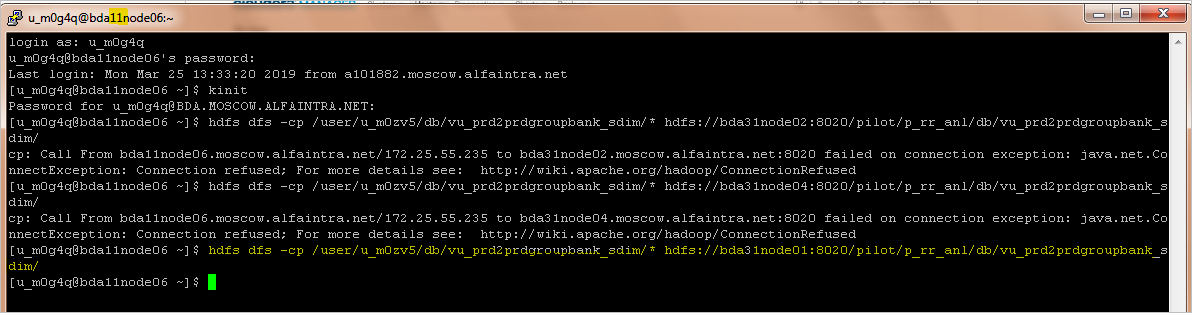




На графике видно что одна из них **bda11node01** ее и указываем как путь к источнику.

Порт на всех кластерах – **8020**

Пример выполнения с bda**1**1:



Копирование из одной области в другую на ОДНОМ кластере

На примере задачи по письму:

*Коллеги, добрый день!*

*Прошу срочно скопировать по запросу* [*http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=279707977*](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=279707977) *от Стаховского Геннадия*

*данные таблицы d\_dmfr.moo\_allassets\_stat, находятся на bda31в директории - /dev/d\_dmfr/db/moo\_allassets\_stat*

*в пилотную схему p\_ufpa на BDA31.*

*Данные таблицы нужно для предоставлению отчета высшему руководству.*

*------------------------------------------------------------------------  
С уважением,  
София Плаксина,*

Также создаем новую таблицу в требуемой области и выполняем две команды:

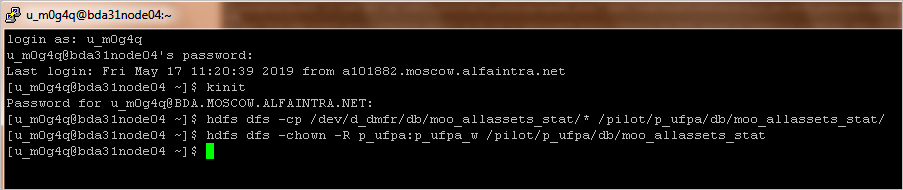
-для копирования данных:

**hdfs dfs -cp /dev/d\_dmfr/db/moo\_allassets\_stat/\* /pilot/p\_ufpa/db/moo\_allassets\_stat/**

- и меняем владельца созданной таблицы на саму область, т.е. подставляем p\_ufpa:

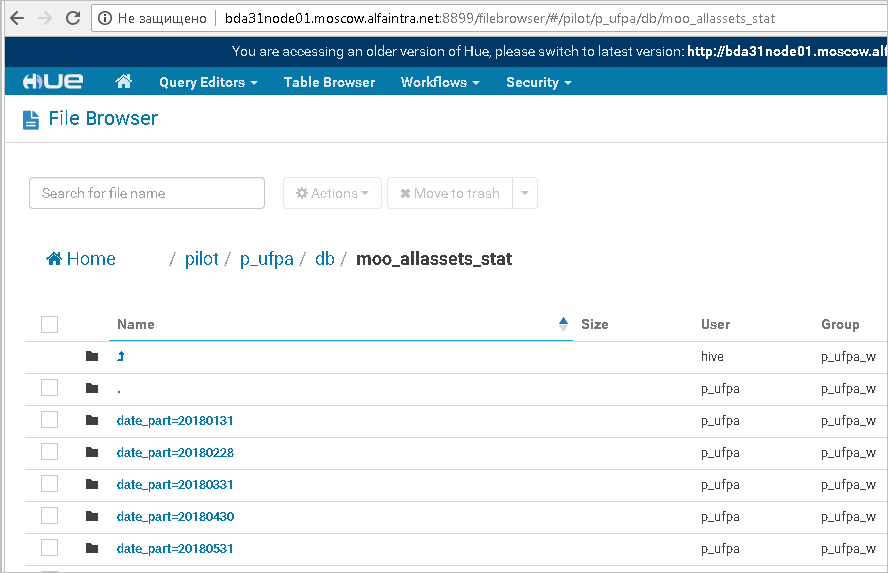
**hdfs dfs -chown -R p\_ufpa:p\_ufpa\_w /pilot/p\_ufpa/db/moo\_allassets\_stat**

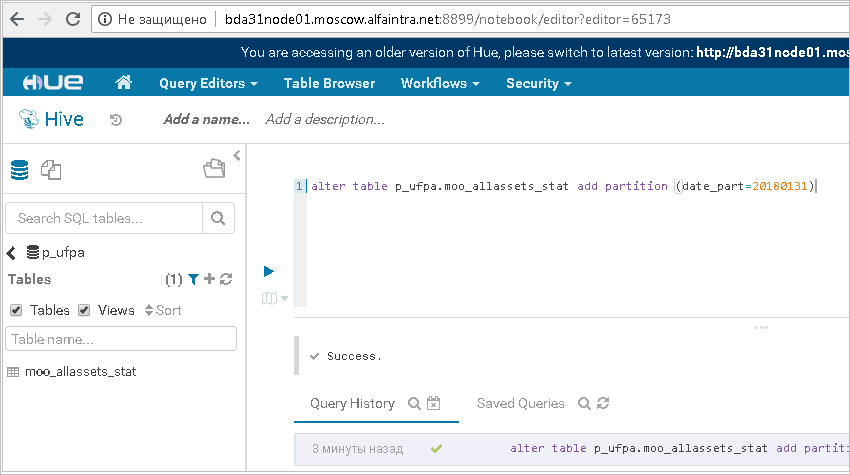
(если этого не сделать, то владельцем таблицы буду я, т.к. создавал ее под своей учеткой)



Т.к. данная таблица **партиционированная**, то для каждой партиции**!** надо выполнить:

**alter table p\_ufpa.moo\_allassets\_stat add partition (date\_part=20180131)** чтобы добавить эти данные таблице, иначе не будут видны (будут просто лежать в каталоге таблицы, но не будут с ней связаны)





**select \* from p\_ufpa.moo\_allassets\_stat where date\_part = 20180131;**

