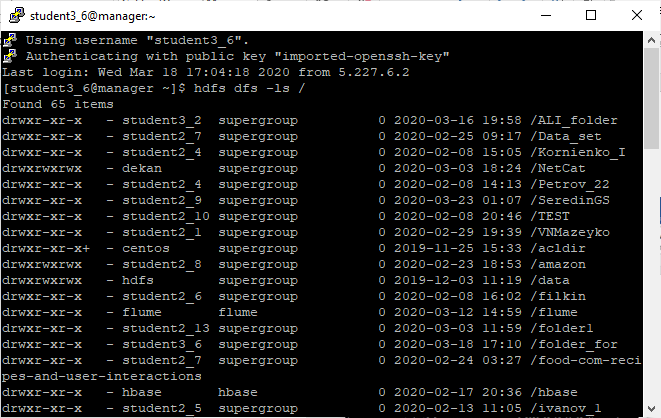
Практическое задание к лекции по HDFS (Лекция №1)

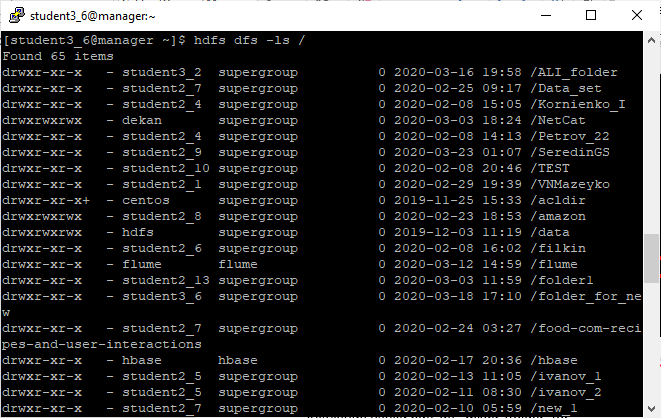
1. Опробовать консольные утилиты для работы с кластером

\* Создать/скопировать/удалить папку

Создание папки hdfs dfs -mkdir /folder\_for



Копирование папки hdfs dfs -mv /folder\_for /folder\_for\_new



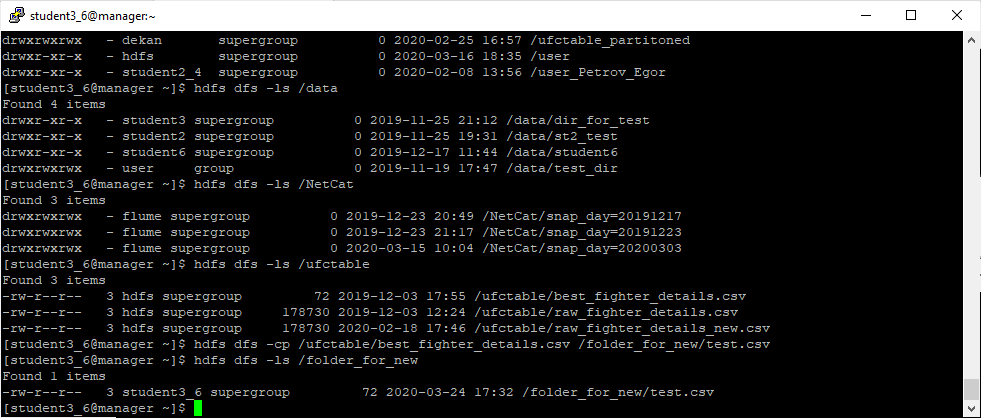
Удаление папки hdfs dfs -rm /folder\_for\_new

\* Положить в HDFS любой файл

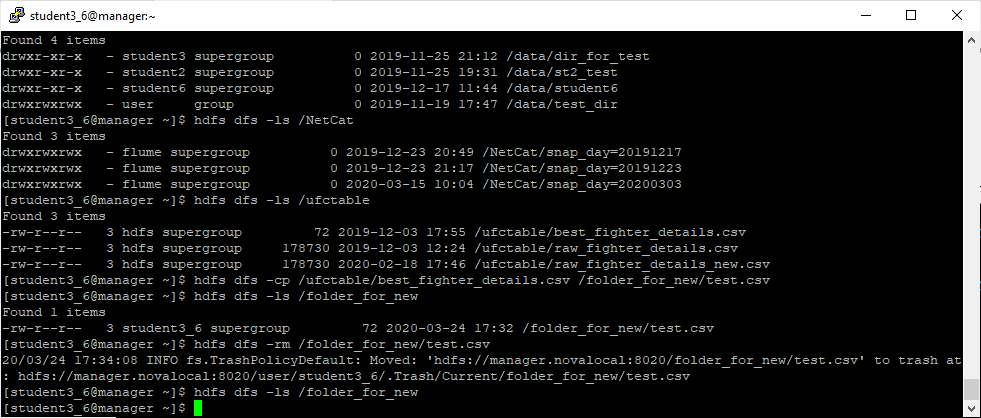
Не совсем понимаю как из Windows перекинуть в HDFS. В Linux понимаю с помощью copyFromLocal.

\* Скопировать/удалить этот файл

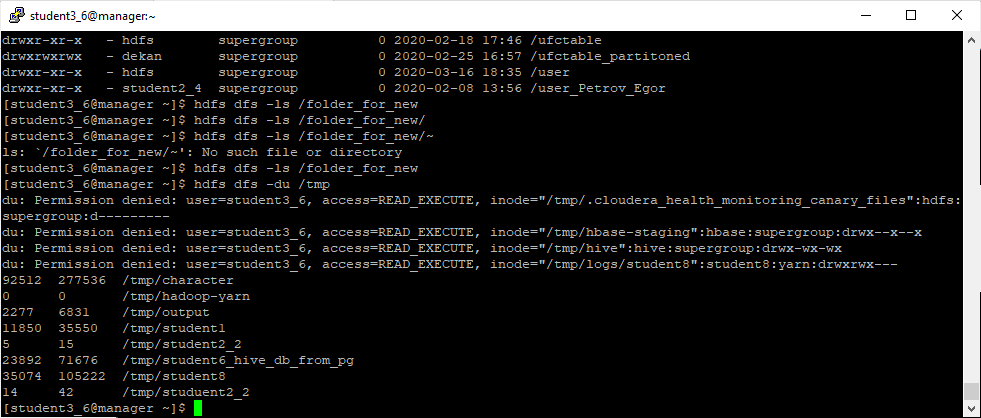
- Скопировать файл hdfs dfs -cp /ufstable/best\_fighter\_details.csv /folder\_for\_new/test.csv



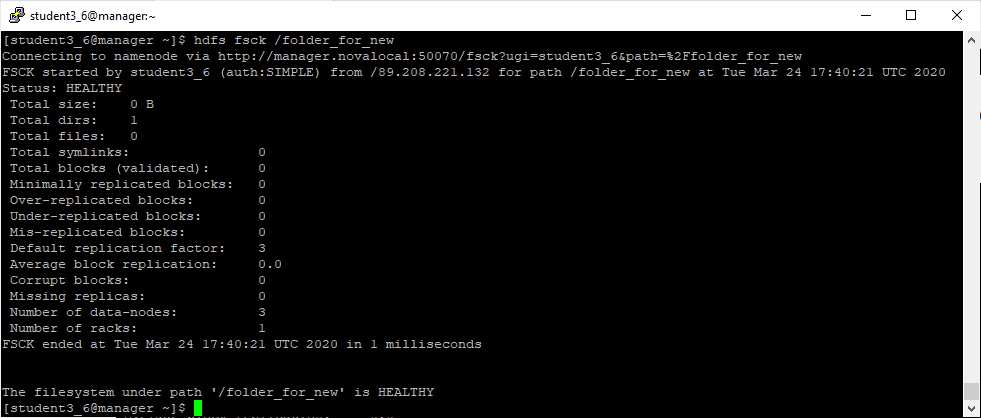
- Удалить файл hdfs dfs -rm /folder\_for\_new/test.csv



\* Просмотреть размер любой папки hdfs dfs -du /tmp



\* Посмотреть как файл хранится на файловой системе (см. команду fsck) hdfs fsck /folder\_for\_new

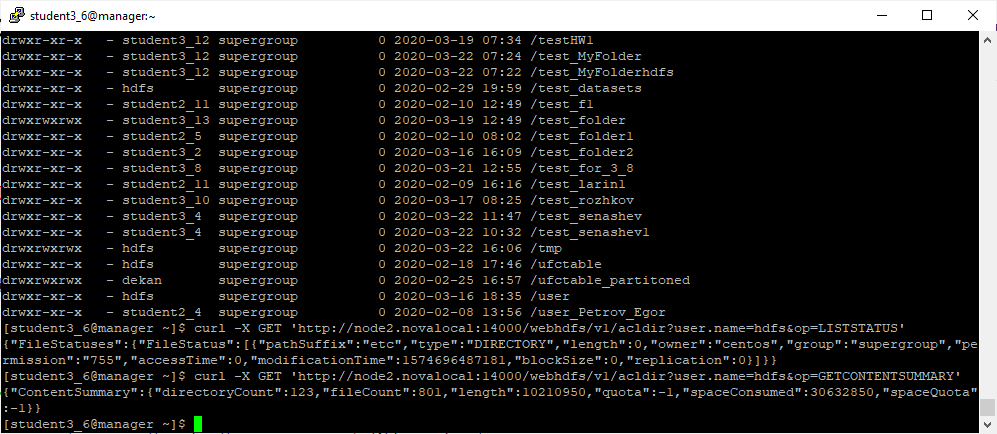


\* Установить нестандартный фактор репликации (см. команду setrep)

hdfs dfs –setrep –R 2 /folder\_for\_new

3. Опробовать rest-доступ для работы с кластером

\* Используя утилиту CURL



\* Используя python3

from hdfs3 import HDFileSystem

hdfs = HDFileSystem(host='89.208.221.132', port=8020, user='student3\_6')

hdfs.ls('/tmp')

5. [Для любителей администрирования] Опробовать NFS доступ. Предварительно связаться со мной чтобы я открыл нужные порты.

6. [Для любителей программирования] Достучаться до файловой системы используя python и библиотеку libhdfs3