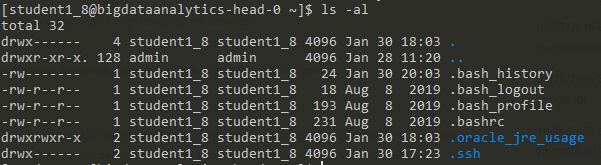
***Задания:***

***1. Опробовать консольные утилиты для работы с кластером***

Вопрос не по теме – обнаружила скрытые файлы в своей папке на кластере. Откуда они там?) Это следы прошлых курсов?



***Создать/скопировать/удалить папку***

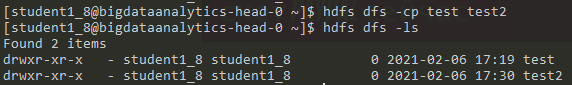
*Создала директорию*

**[student1\_8@bigdataanalytics-head-0 ~]$ hdfs dfs -mkdir test**

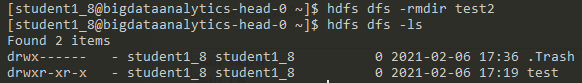
*Убедилась, что она появилась*



*Скопировала папку, убедилась, что она появилась*

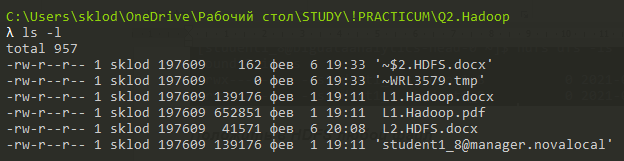


*Удалила папку и убедилась, что ее нет. Увидела, что появилась «корзина»*



***Положить в HDFS любой файл***

*Перешла в папку на локалке и выбрала самый большой файл*

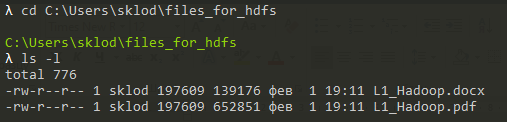


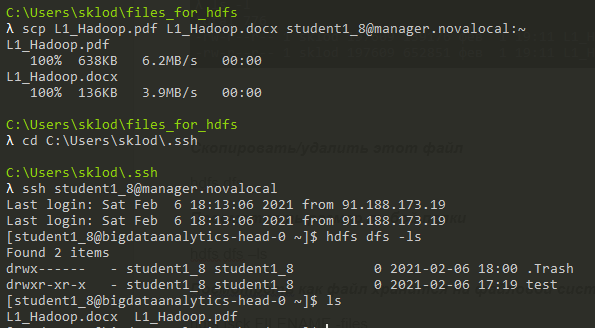
*И попыталась скопировать его на кластер, но меня ждало разочарование*



*После нескольких десятков переходов по папкам и вариаций написания скрипта мой КПД стремился к нулю(((( НО гугл творит чудеса и я таки нашла где-то комментарий, что такая ошибка может вылетать, если в пути имени есть директории с названием на русском языке! Велик и могуч русский язык, блин…*

*Переложила файлы в другое место и чудо!*

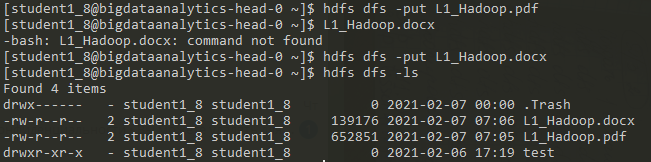




*Не понимаю только… файлы в итоге лежат не там? Через hdfs dfs –ls их не видно((((*

*….*

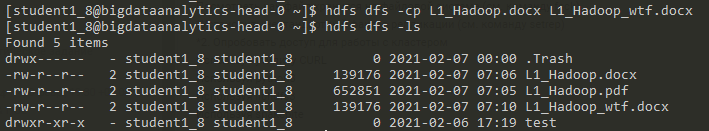
*В общем, я потратила много времени, чтоб осознать, что файлы попали на сервер, но не попали в Hadoop. Нагуглить внятный алгоритм передвижения файлов я так и не смогла, пришлось действовать интуитивно методом проб и ошибок*



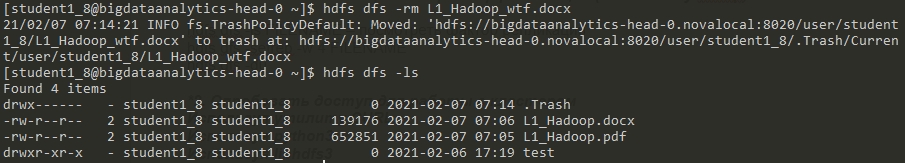
*Кажется, я более менее разобралась)*

***Скопировать/удалить этот файл***

*Скопировала файл и убедилась, что он появился*

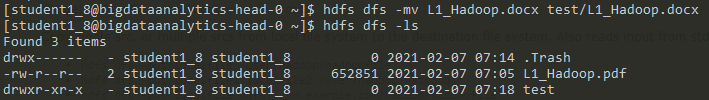


*А потом отправила свое творение в Trash*



***Просмотреть размер любой папки***

*Поскольку у меня только одна папка и она пустая, я решила положить туда один из файлов*

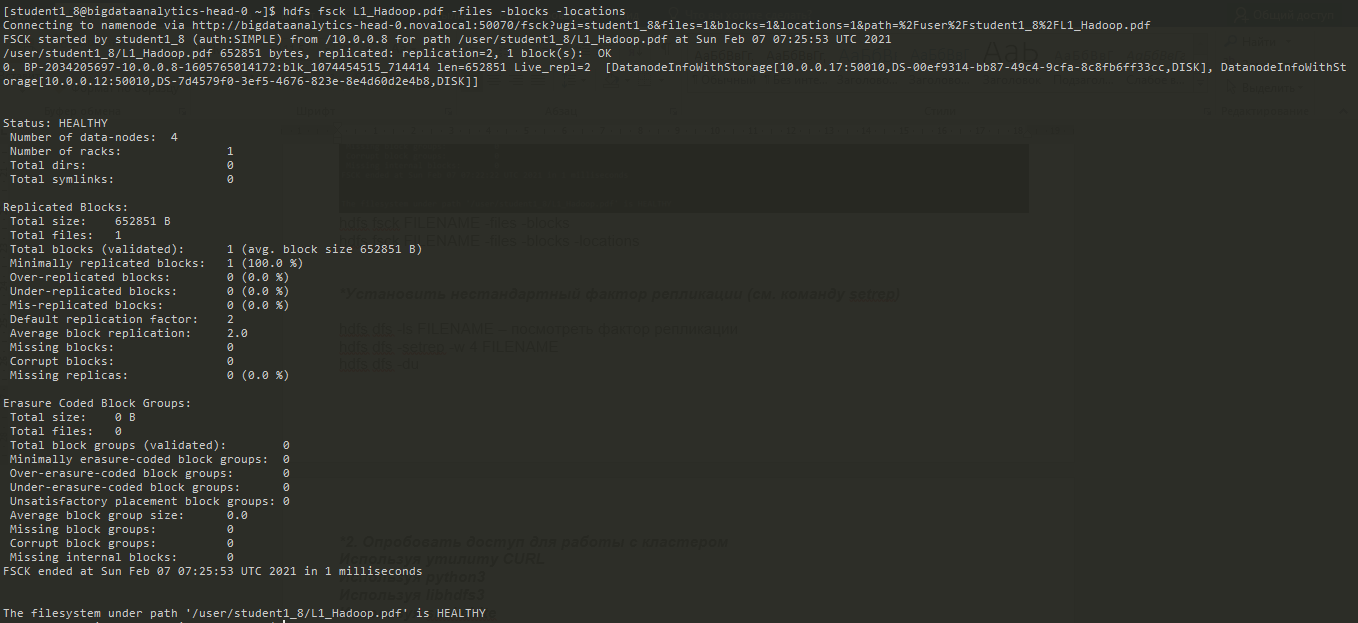


*А потом проверила размер папки и убедилась, что файл таки там*



***Посмотреть как файл хранится на файловой системе (см. команду fsck)***

*Проверяю все операторы* -files -blocks -locations



[student1\_8@bigdataanalytics-head-0 ~]$ hdfs fsck L1\_Hadoop.pdf -files -blocks –locations  
Connecting to namenode via <http://bigdataanalytics-head-0.novalocal:50070/fsck?ugi=student1_8&files=1&blocks=1&locations=1&path=%2Fuser%2Fstudent1_8%2FL1_Hadoop.pdf>  
FSCK started by student1\_8 (auth:SIMPLE) from /10.0.0.8 for path /user/student1\_8/L1\_Hadoop.pdf at Sun Feb 07 07:25:53 UTC 2021  
/user/student1\_8/L1\_Hadoop.pdf 652851 bytes, replicated: replication=2, 1 block(s): OK  
0. BP-2034205697-10.0.0.8-1605765014172:blk\_1074454515\_714414 len=652851 Live\_repl=2 [DatanodeInfoWithStorage[10.0.0.17:50010,DS-00ef9314-bb87-49c4-9cfa-8c8fb6ff33cc,DISK], DatanodeInfoWithStorage[10.0.0.12:50010,DS-7d4579f0-3ef5-4676-823e-8e4d60d2e4b8,DISK]]

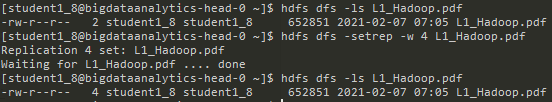
Status: HEALTHY  
Number of data-nodes: 4  
Number of racks: 1  
Total dirs: 0  
Total symlinks: 0  
Replicated Blocks:  
Total size: 652851 B  
Total files: 1  
Total blocks (validated): 1 (avg. block size 652851 B)  
Minimally replicated blocks: 1 (100.0 %)  
Over-replicated blocks: 0 (0.0 %)  
Under-replicated blocks: 0 (0.0 %)  
Mis-replicated blocks: 0 (0.0 %)  
Default replication factor: 2  
Average block replication: 2.0  
Missing blocks: 0  
Corrupt blocks: 0  
Missing replicas: 0 (0.0 %)  
Erasure Coded Block Groups:  
Total size: 0 B  
Total files: 0  
Total block groups (validated): 0  
Minimally erasure-coded block groups: 0  
Over-erasure-coded block groups: 0  
Under-erasure-coded block groups: 0  
Unsatisfactory placement block groups: 0  
Average block group size: 0.0  
Missing block groups: 0  
Corrupt block groups: 0  
Missing internal blocks: 0  
SCK ended at Sun Feb 07 07:25:53 UTC 2021 in 1 milliseconds

The filesystem under path '/user/student1\_8/L1\_Hadoop.pdf' is HEALTHY

*При попытке найти блок получаю только* **Permission denied**

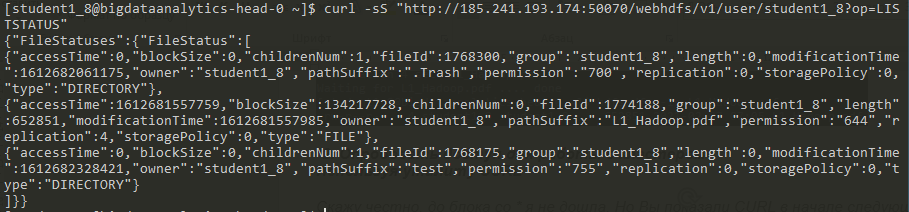
***\*Установить нестандартный фактор репликации (см. команду setrep)***

*Фактор репликации поменяла с 2 на 4*



***\*2. Опробовать доступ для работы с кластером  
Используя утилиту CURL***

*Скажу честно, до блока со \* я не дошла. Но Вы показали CURL в начале следующего урока, и я просто проверила, как это работает*



*Не представляю, сколько часов гугл мне понадобилось бы, чтоб найти эту длинную строчку, которую я у вас списала*

***Используя python3  
Используя libhdfs3  
Используя snakebite***