1. Опробовать консольные утилиты для работы с кластером

* Создать/скопировать/удалить папку

`hdfs dfs -mkdir /student6\_1 `

`hdfs dfs -cp /student6\_1 /student6\_1 \_copy`

`hdfs dfs -rm -R /student6\_1 \_copy`

* Положить в HDFS любой файл

`hdfs dfs -copyFromLocal /home/student6\_1 /.bashrc /student6\_1 /`

`hdfs dfs -ls /student6\_1 `

Результат:

```

Found 1 items

-rw-r--r-- 3 student6\_1 supergroup 231 2020-07-19 23:55 /student6\_1 /.bashrc

```

* Скопировать/удалить этот файл

`hdfs dfs -get /student6\_1/.bashrc /tmp`

`hdfs dfs -copyToLocal /student6\_1/.bashrc /tmp`.

Порядок аргументов такой же. Сперва откуда, потом куда.

Проверяем командой `ls -la /tmp`.

Сразу удалим этот файл с рабочей станции: `rm /tmp/.bashrc`.

Скопируем файл: `hdfs dfs -cp /student6\_1/.bashrc /student6\_1/.bashrc\_copy`. Удалим дубликат: `hdfs dfs -rm /student6\_1/.bashrc\_copy`

* Просмотреть размер любой папки

`hdfs dfs -du /student6\_1/`.

Результат:

```

231 693 /student6\_1/.bashrc

```

Этот файл файл 231 байт

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

`hdfs fsck /student6\_1/`

Результат:

```

Total size: 231 B

Total dirs: 1

Total files: 1

Total symlinks: 0

Total blocks (validated): 1 (avg. block size 231 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: 3

Average block replication: 3.0

Corrupt blocks: 0

Missing replicas: 0 (0.0 %)

Number of data-nodes: 3

Number of racks: 1

```

* Установить нестандартный фактор репликации (см. команду setrep)

```

[student6\_1@manager /]$ hdfs dfs -setrep 2 /student6\_1/

Replication 2 set: /student6\_1/.bashrc

```

Сделаем копию файла ещё раз:

`hdfs dfs -cp /student6\_1/.bashrc /student6\_1/.bashrc\_copy`

Проверим содержимое папки:

```

[student6\_1@manager /]$ hdfs dfs -du -h /student6\_1/

231 462 /student6\_1/.bashrc

231 693 /student6\_1/.bashrc\_copy

```

Нестандартный фактор репликации применился к существующим файлам. Но к новым файлам в этой папке применяется дефолтный фактор.

1. Опробовать rest-доступ для работы с кластером

* Используя утилиту CURL

`curl -X GET 'http://node2.novalocal:14000/webhdfs/v1/acldir?user.name=student6\_1&op=LISTSTATUS'`.

Результат:

```json

{

"FileStatuses": {

"FileStatus": [

{

"pathSuffix": "etc",

"type": "DIRECTORY",

"length": 0,

"owner": "centos",

"group": "supergroup",

"permission": "755",

"accessTime": 0,

"modificationTime": 1574696487181,

"blockSize": 0,

"replication": 0

}

]

}

}

```

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

```

from hdfs3 import HDFileSystem

hdfs=HDFileSystem(host='manager.novalocal',port=8020)

hdfs.ls('/user/instructor')

```