

См. флаги “-mkdir” и “-touchz”

1. Создайте папку в корневой HDFS-папке

```
hdfs dfs -mkdir /test
```

2. Создайте в созданной папке новую вложенную папку.

```
hdfs dfs -mkdir /test/test_child
```

3. Что такое Trash в распределенной FS? Как сделать так, чтобы файлы удалялись сразу, минуя “Trash”?

Trash – папка с удаленными файлами. Удаление сразу, команда `-skipTrash`

4. Создайте пустой файл в подпапке из пункта 2.

```
hdfs dfs -touchz /test/test_child/test_file
```

5. Удалите созданный файл.

```
hdfs dfs -rm -skipTrash /test/test_child/test_file
```

6. Удалите созданные папки.

```
hdfs dfs -rm -r -skipTrash /test
```

См. флаги “-put”, “-cat”, “-tail”, “-cp”

1. Скопируйте любой в новую папку на HDFS

```
hdfs dfs -mkdir /new_folder; hdfs dfs -put test.txt /new_folder/
```

2. Выведите содержимое HDFS-файла на экран.

```
hdfs dfs -cat /new_folder/test.txt
```

3. Выведите содержимое нескольких последних строчек HDFS-файла на экран.

```
hdfs dfs -tail /new_folder/test.txt
```

4. Выведите содержимое нескольких первых

строчек HDFS-файла на экран.

```
hdfs dfs -head /new_folder/test.txt
```

5. Переместите копию файла в HDFS на новую локацию.

```
hdfs dfs -cp /new_folder/test.txt /NEW_FOLDER
```

Задачи:

1. Изменить replication factor для файла. Как долго занимает время на увеличение / уменьшение числа реплик для файла?

```
hdfs dfs -setrep -w 2 /new_folder/test.txt
```

Изменение replication Factor происходит быстро ~ 5 сек

2. Найдите информацию по файлу, блокам и их расположениям с помощью “hdfs fsck”

```
hdfs fsck /new_folder/test.txt -files -blocks  
-locations
```

3. Получите информацию по любому блоку из п.2 с помощью “hdfs fsck -blockId”.

Обратите внимание на Generation Stamp (GS number).

```
hdfs fsck -blockId blk_1073741831
```