Скачаем Hadoop:

Установим:

tar -xvf hadoop-2.8.5.tar.gz

Установим пути в файл ~/.bashrc:

```
export JAVA_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64
export HADOOP_HOME=$HOME/hadoop-2.8.5
export HADOOP_HDFS_HOME=$HADOOP_HOME
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
```


Посмотрим версию:

Получим список файлов на удаленном сервере:

Найдем файл part-r-00000:

Выполним команду lsr:

cat:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000

```
Let       6
Lewis     4
License 8
License.          2
Like      1
Literary          13
Little  2
Lizard  2
Lizard) 1
Lizard, 2
Lizard's          1
Lobster 5
Lobster;          1
London  1
Long    2
Longitude         2
Lory    3
Lory,   3
Lory.   1
Luckily 1
M,      1
M?"     1
MARMALADE",       1
ME,"    2
MERCHANTABILITY 1
ME"     1
MILLENNIUM        1
Ma!"    1
Mabel   1
Mabel!  1
```

Самое часто встречающееся слово нельзя быстро найти, так как различных слов в тексте очень много

awk:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000 | awk '{print $1":"$2}'


```
fact.:1
fact.":1
fading:1
failure.:1
faint:1
faintly:1
fair:1
fairly,":1
fairy-tales,:1
fall:6
fall,:1
fallen:4
falling:2
familiarly:1
family:1
fan:8
fan!:1
fan,:1
fancied:2
fancy:3
fancy,:2
fancying:1
fancy—Who's:1
fanned:1
fanning:1
far:9
far!":1
far,:1
far,":2
farm-yard—while:1
```

Найдем все the в тексте:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000 | awk '{print $1":"$2}' | grep the

Численная сортировка:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000 | awk '{print $1":"$2}' | grep the | sort -t: -nk2

Найдем слово Alice:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000 |awk '{print $1":"$2}' | grep Alice | sort -t: -nk2

Получить корректный ответ мешают знаки препинания.

Уберем знаки препинания:

hdfs dfs -cat hdfs://192.168.121.147:9000/lab4/test1out/part-r-00000 | awk '

{gsub(/[[:punct:]]/,""); print $1":"$2}' | grep Alice | sort -t: -nk2

Создадим на диске файл и запишем туда состав группы:

Переместим файл в папку exam:

hdfs dfs -put 50916kekwait hdfs://192.168.121.147:9000/exam/

hdfs dfs -ls 50916kekwait hdfs://192.168.121.147:9000/exam/

Переместим файл в папку lab4:

```
tor/hadoop-2.8.5/share/hadoop/common/lib/hadoop-auth-2.8.5.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUti
l
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
administrator@clickhousestudent:~/50916kekw$ hdfs dfs -ls 50916kekwait hdfs://192.168.121.147:9000/exam/
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/administra
tor/hadoop-2.8.5/share/hadoop/common/lib/hadoop-auth-2.8.5.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUti
l
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
-rw-rw-r--   1 administrator administrator        291 2020-12-16 07:52 50916kekwait
Found 1 items
-rw-r--r--   3 administrator supergroup           291 2020-12-16 07:53 hdfs://192.168.121.147:9000/exam/50916kekwait
administrator@clickhousestudent:~/50916kekw$ nano 50916kekwait
administrator@clickhousestudent:~/50916kekw$ hdfs dfs -put 50916kekwait hdfs://192.168.121.147:9000/lab4
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/administra
tor/hadoop-2.8.5/share/hadoop/common/lib/hadoop-auth-2.8.5.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUti
l
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
put: Permission denied: user=administrator, access=WRITE, inode="/lab4":hadoop:supergroup:drwxr-xr-x
administrator@clickhousestudent:~/50916kekw$ hdfs dfs -put 50916kekwait hdfs://192.168.121.147:9000/lab4
```

put: Permission denied: user=administrator, access=WRITE, inode="/lab4":hadoop:supergroup:drwxr-xr-x