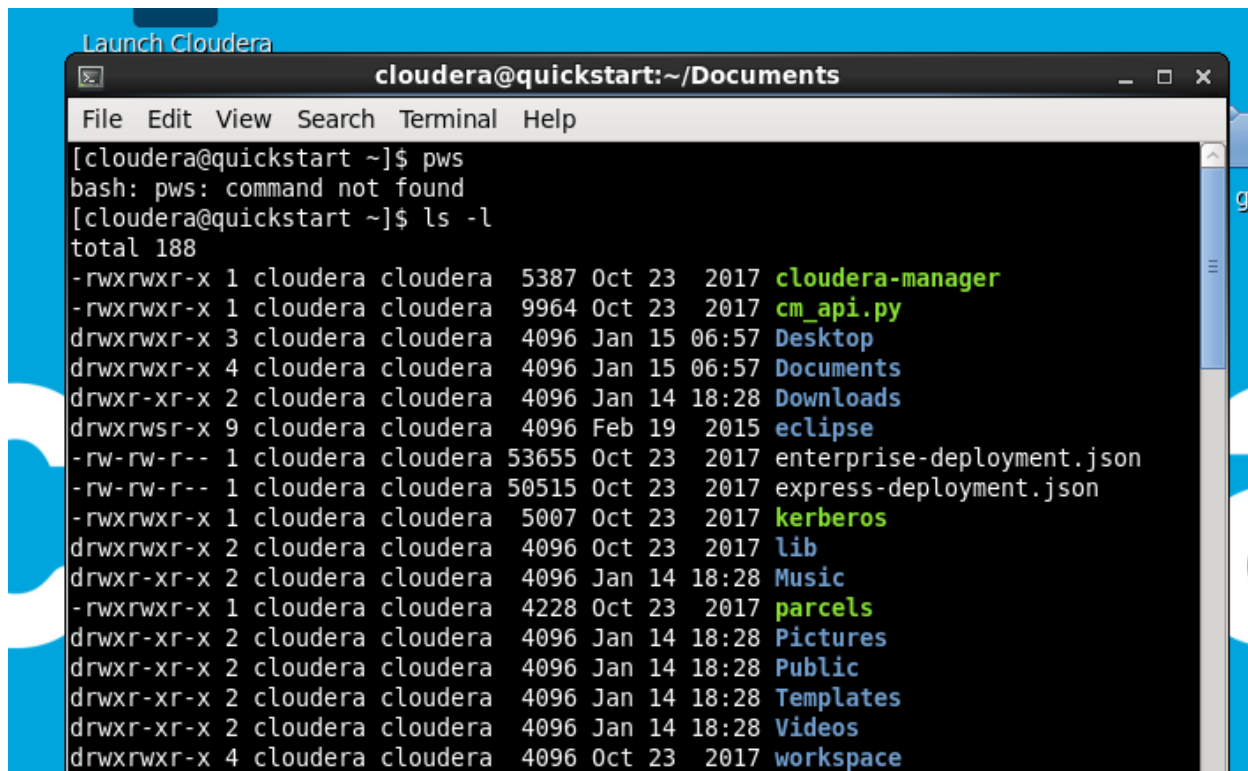


После запуска терминала проверим где мы находимся и посмотрим содержимое иеректории:

pwd

ls -l



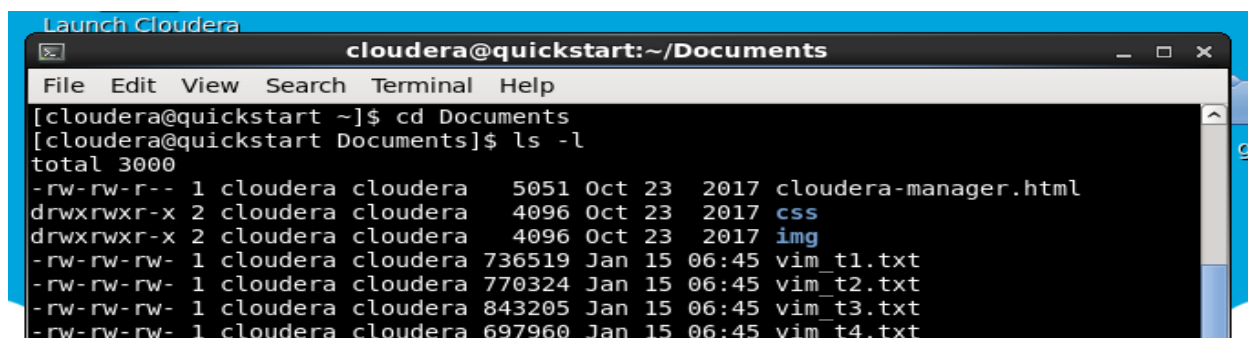
```
Launch Cloudera
cloudera@quickstart:~/Documents
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ pwd
bash: pwd: command not found
[cloudera@quickstart ~]$ ls -l
total 188
-rwxrwxr-x 1 cloudera cloudera 5387 Oct 23 2017 cloudera-manager
-rwxrwxr-x 1 cloudera cloudera 9964 Oct 23 2017 cm_api.py
drwxrwxr-x 3 cloudera cloudera 4096 Jan 15 06:57 Desktop
drwxrwxr-x 4 cloudera cloudera 4096 Jan 15 06:57 Documents
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Downloads
drwxrwsr-x 9 cloudera cloudera 4096 Feb 19 2015 eclipse
-rw-rw-r-- 1 cloudera cloudera 53655 Oct 23 2017 enterprise-deployment.json
-rw-rw-r-- 1 cloudera cloudera 50515 Oct 23 2017 express-deployment.json
-rwxrwxr-x 1 cloudera cloudera 5007 Oct 23 2017 kerberos
drwxrwxr-x 2 cloudera cloudera 4096 Oct 23 2017 lib
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Music
-rwxrwxr-x 1 cloudera cloudera 4228 Oct 23 2017 parcels
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Pictures
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Public
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Templates
drwxr-xr-x 2 cloudera cloudera 4096 Jan 14 18:28 Videos
drwxrwxr-x 4 cloudera cloudera 4096 Oct 23 2017 workspace
```

Мы в основной домашней директории.

Файлы томов Толстова находятся в папке Documents. Покажем их в терминале:

cd Documents

ls -l



```
Launch Cloudera
cloudera@quickstart:~/Documents
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ cd Documents
[cloudera@quickstart Documents]$ ls -l
total 3000
-rw-rw-r-- 1 cloudera cloudera 5051 Oct 23 2017 cloudera-manager.html
drwxrwxr-x 2 cloudera cloudera 4096 Oct 23 2017 css
drwxrwxr-x 2 cloudera cloudera 4096 Oct 23 2017 img
-rw-rw-rw- 1 cloudera cloudera 736519 Jan 15 06:45 vim_t1.txt
-rw-rw-rw- 1 cloudera cloudera 770324 Jan 15 06:45 vim_t2.txt
-rw-rw-rw- 1 cloudera cloudera 843205 Jan 15 06:45 vim_t3.txt
-rw-rw-rw- 1 cloudera cloudera 697960 Jan 15 06:45 vim_t4.txt
```

Все 4 тома на месте.

Посмотрим содержимое папки `hdfs`. Папка должна быть пуста. После этого скопируем в нее наши 4 файла и проверим то, что они скопированы:

```
hdfs dfs -ls
```

```
hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t1.txt
```

```
hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t2.txt
```

```
hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t3.txt
```

```
hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t4.txt
```

```
hdfs dfs -ls
```

```
[cloudera@quickstart Documents]$ hdfs dfs -ls
[cloudera@quickstart Documents]$ hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t1.txt
^[[[cloudera@quickstart Documents]$ hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t2.txt
^[[A^[[B[cloudera@quickstart Documents]$ hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t3.txt
[cloudera@quickstart Documents]$ hdfs dfs -copyFromLocal /home/cloudera/Documents/vim_t4.txt
[cloudera@quickstart Documents]$ hdfs dfs -ls
Found 4 items
-rw-r--r-- 1 cloudera cloudera 736519 2023-01-15 07:02 vim_t1.txt
-rw-r--r-- 1 cloudera cloudera 770324 2023-01-15 07:02 vim_t2.txt
-rw-r--r-- 1 cloudera cloudera 843205 2023-01-15 07:03 vim_t3.txt
-rw-r--r-- 1 cloudera cloudera 697960 2023-01-15 07:03 vim_t4.txt
```

Home / user / cloudera

	Name	Size	User	Group	Permissions	Date
	↑		hdfs	supergroup	drwxr-xr-x	October 23, 2017 09:17 AM
	.		cloudera	cloudera	drwxr-xr-x	January 15, 2023 07:03 AM
	vim_t1.txt	719.3 KB	cloudera	cloudera	-rw-r--r--	January 15, 2023 07:02 AM
	vim_t2.txt	752.3 KB	cloudera	cloudera	-rw-r--r--	January 15, 2023 07:02 AM
	vim_t3.txt	823.4 KB	cloudera	cloudera	-rw-r--r--	January 15, 2023 07:03 AM
	vim_t4.txt	681.6 KB	cloudera	cloudera	-rw-r--r--	January 15, 2023 07:03 AM

Show 45 of 4 items Page 1 of 1

Копирование выполнено верно.

Объединим наши 4 файла в один и назовем общий файл merge.txt:

```
hdfs dfs -getmerge /user/cloudera /home/cloudera/Downloads/merge.txt
```

```
ls -l
```

```
^[[[cloudera@quickstart Documents]$ hdfs dfs -getmerge /user/cloudera /home/clo
ra/Documents/merge.txt
[cloudera@quickstart Documents]$ ls -l
total 5980
-rw-rw-r-- 1 cloudera cloudera 5051 Oct 23 2017 cloudera-manager.html
drwxrwxr-x 2 cloudera cloudera 4096 Oct 23 2017 css
drwxrwxr-x 2 cloudera cloudera 4096 Oct 23 2017 img
-rw-r--r-- 1 cloudera cloudera 3048008 Jan 15 07:11 merge.txt
-rw-rw-rw- 1 cloudera cloudera 736519 Jan 15 06:45 vim_t1.txt
-rw-rw-rw- 1 cloudera cloudera 770324 Jan 15 06:45 vim_t2.txt
-rw-rw-rw- 1 cloudera cloudera 843205 Jan 15 06:45 vim_t3.txt
-rw-rw-rw- 1 cloudera cloudera 697960 Jan 15 06:45 vim_t4.txt
```

Файлы романа объединены.

Скопируем merge.txt в hdfs и изменим права доступа к файлу:

```
hdfs dfs -copyFromLocal /home/cloudera/Documents/merge.txt
```

```
hdfs dfs -chmod 755 merge.txt
```

```
hdfs dfs -ls
```

```
[cloudera@quickstart Documents]$ hdfs dfs -copyFromLocal /home/cloudera/Documen
ts/merge.txt
[cloudera@quickstart Documents]$ hdfs dfs -chmod 755 merge.txt
[cloudera@quickstart Documents]$ hdfs dfs -ls
Found 5 items
-rwxr-xr-x 1 cloudera cloudera 3048008 2023-01-15 07:15 merge.txt
-rw-r--r-- 1 cloudera cloudera 736519 2023-01-15 07:02 vim_t1.txt
-rw-r--r-- 1 cloudera cloudera 770324 2023-01-15 07:02 vim_t2.txt
-rw-r--r-- 1 cloudera cloudera 843205 2023-01-15 07:03 vim_t3.txt
-rw-r--r-- 1 cloudera cloudera 697960 2023-01-15 07:03 vim_t4.txt
```

Права доступа к файлу успешно изменены.

Выведем размер файла merge.txt:

```
hdfs dfs -du -h /user/cloudera/merge.txt
```

```
[cloudera@quickstart Documents]$ hdfs dfs -du -h /user/cloudera/merge.txt
2.9 M  2.9 M  /user/cloudera/merge.txt
```

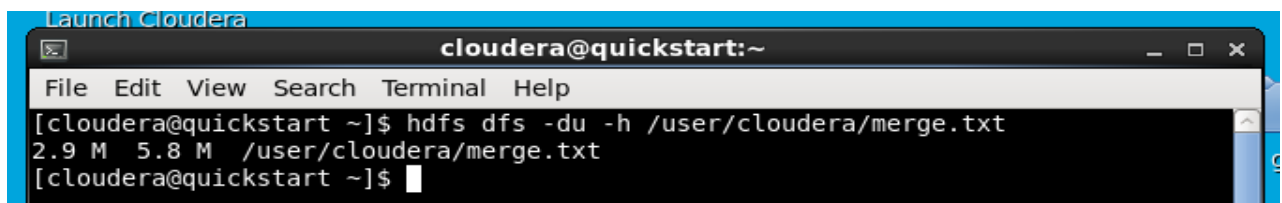
Поменяем количество репликации и снова проверим размер файла:

```
hdfs dfs -setrep -w 2 /user/cloudera/merge.txt
```

```
hdfs dfs -du -h /user/cloudera/merge.txt
```

```
[cloudera@quickstart Documents]$ hdfs dfs -setrep -w 2 /user/cloudera/merge.txt
Replication 2 set: /user/cloudera/merge.txt
Waiting for /user/cloudera/merge.txt .....
.....
```

Р.с.: т.к. есть какая-то проблема со скоростью репликации, то я остановил процесс, но размер файла все же изменился.



```
Launch Cloudera
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -du -h /user/cloudera/merge.txt
2.9 M  5.8 M  /user/cloudera/merge.txt
[cloudera@quickstart ~]$
```

Подсчитаем количество строк в файле:

```
hadoop fs -cat /user/cloudera/merge.txt | wc -l
```

```
[cloudera@quickstart Documents]$ hadoop fs -cat /user/cloudera/merge.txt | wc -l
10272
```