

1. После разворачивания всех контейнеров рекурсивно скопируем все скачанные файлы книг с локального компьютера в контейнер datanode-1, попутно создав в нём папку data_books:

Команда:

```
docker cp books/ docker-hadoop-hive-parquet_datanode_1:data_books/
```

Ответ:

```
Successfully copied 2.2MB to docker-hadoop-hive-parquet_datanode_1:data_books/
```

2. Подключаемся к datanode-1:

Команда:

```
docker exec -it docker-hadoop-hive-parquet_datanode_1 bash
```

Ответ:

```
root@9e77eda57491:/# l
```

3. Объединим все книги в один файл:

Команда:

```
cat data_books/* > data_books/war_and_peace.txt
```

4. Отправим полученный файл в файловую систему HDFS:

Команда:

```
hadoop fs -put data_books/war_and_peace.txt /user/hadoop/war_and_peace.txt
```

5. Запросим содержимое личной папки на HDFS:

Команда:

```
hadoop fs -ls /user/hadoop/
```

Ответ:

```
Found 1 items
-rw-r--r--  3 root hadoop    3048008 2023-08-07 09:11 /user/hadoop/war_and_peace.txt
```

6. Поменяем права на файл: полные для владельца файла, а для остальных права на чтение и исполнение:

Команда:

```
hadoop fs -chmod 755 /user/hadoop/war_and_peace.txt
```

7. Проверим заново права доступа:

Команда:

```
hadoop fs -ls /user/hadoop/
```

Ответ:

```
Found 1 items
-rwxr-xr-x  3 root hadoop    3048008 2023-08-07 09:11 /user/hadoop/war_and_peace.txt
```

8. Узнаем размер конкретного файла в мб. :

Команда:

```
hadoop fs -du -h /user/hadoop/war_and_peace.txt
```

Ответ:

```
2.9 M /user/hadoop/war_and_peace.txt
```

9. Но команда выше покажет просто размер файла, а какой фактор репликации или сколько памяти занимает файл с учётом репликации мы не узнаем, поэтому воспользуемся другой командой, и увидим что текущий фактор репликации равен 3, помимо размера файла в мб:

Команда:

```
hadoop fs -ls -h /user/hadoop/
```

Ответ:

```
Found 1 items
-rwxr-xr-x  3 root hadoop    2.9 M 2023-08-07 09:11 /user/hadoop/war_and_peace.txt
```

9. Устанавливаем фактор репликации равный 2-м:

Команда:

```
hadoop fs -setrep 2 /user/hadoop/war_and_peace.txt
```

Ответ:

```
Replication 2 set: /user/hadoop/war_and_peace.txt
```

10. Проверим какой теперь фактор репликации:

Команда:

```
hadoop fs -ls -h /user/hadoop/
```

Ответ:

```
Found 1 items
-rwxr-xr-x  2 root hadoop    2.9 M 2023-08-07 09:11 /user/hadoop/war_and_peace.txt
```

11. Наконец, подсчитаем количество строк в файле:

Команда:

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

Ответ:

```
10272
```