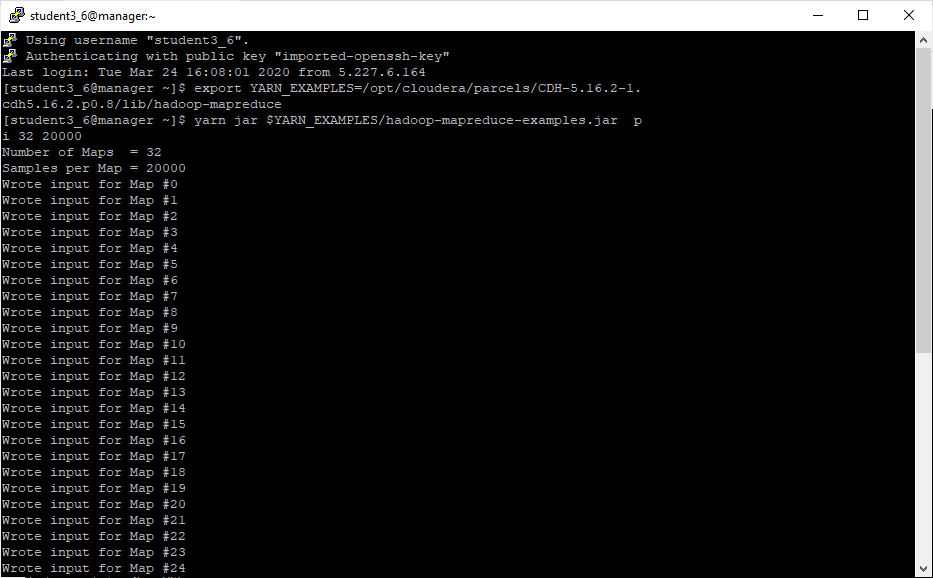
Практическое задание к лекции по HDFS (Лекция №2)

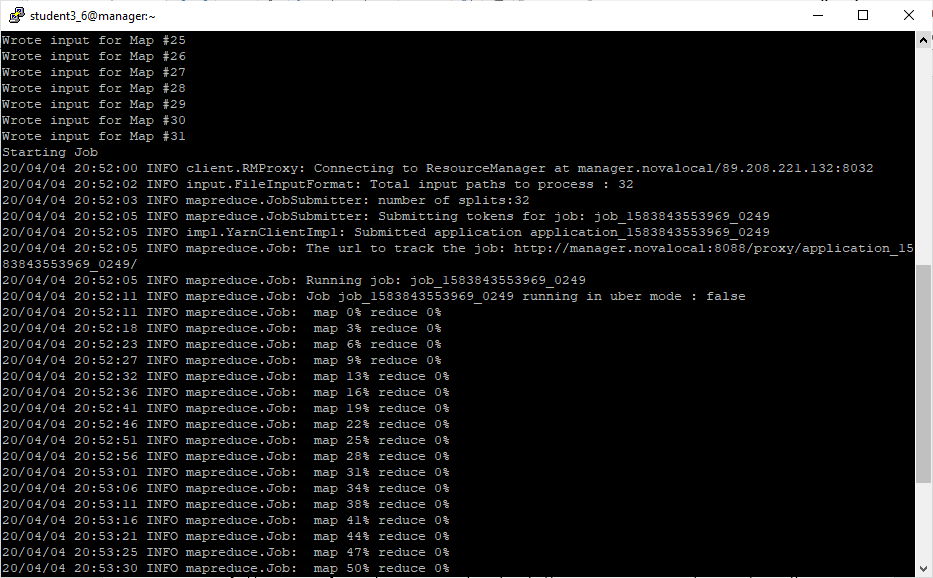
0. [Исследовательское задание]

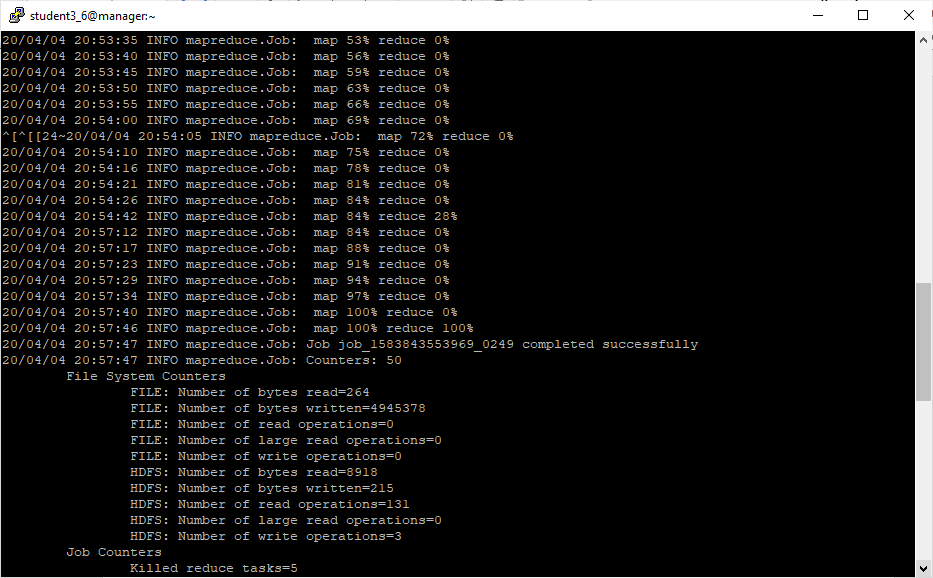
Сколько узлов можно потерять в кластере из 10 узлов без потери данных? Из 100 узлов?

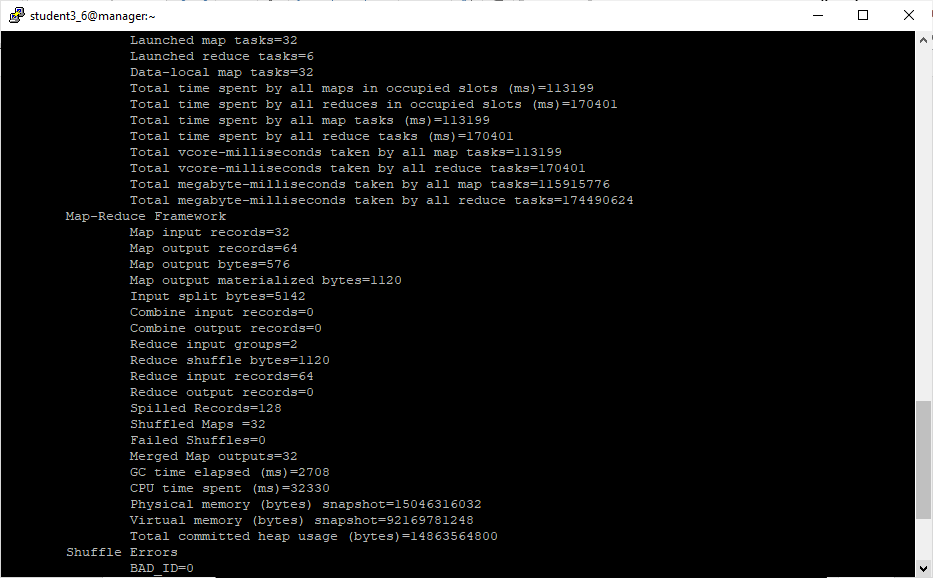
1. Опробовать запуски map-reduce задач для кластера используя hadoop-mapreduce-examples.jar.

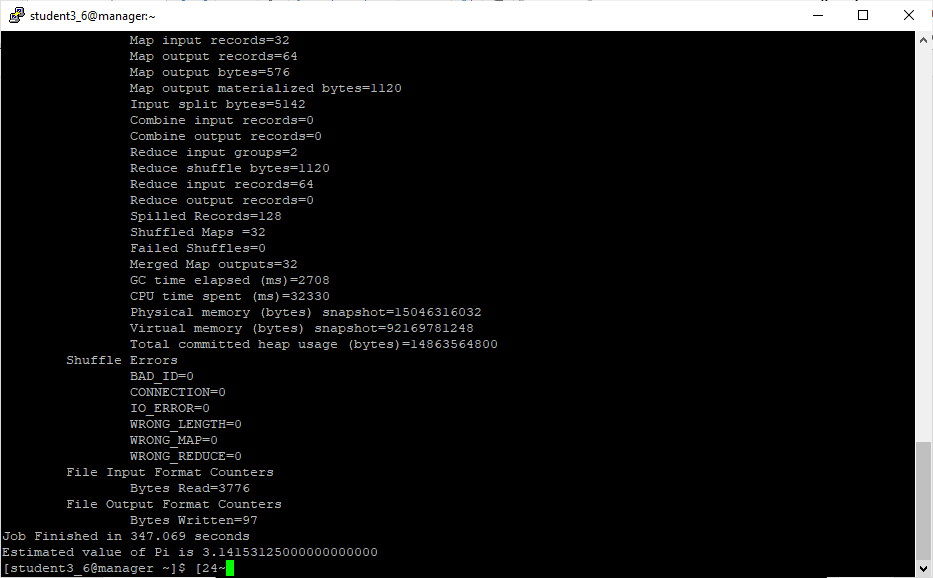
Запускаем задачу:





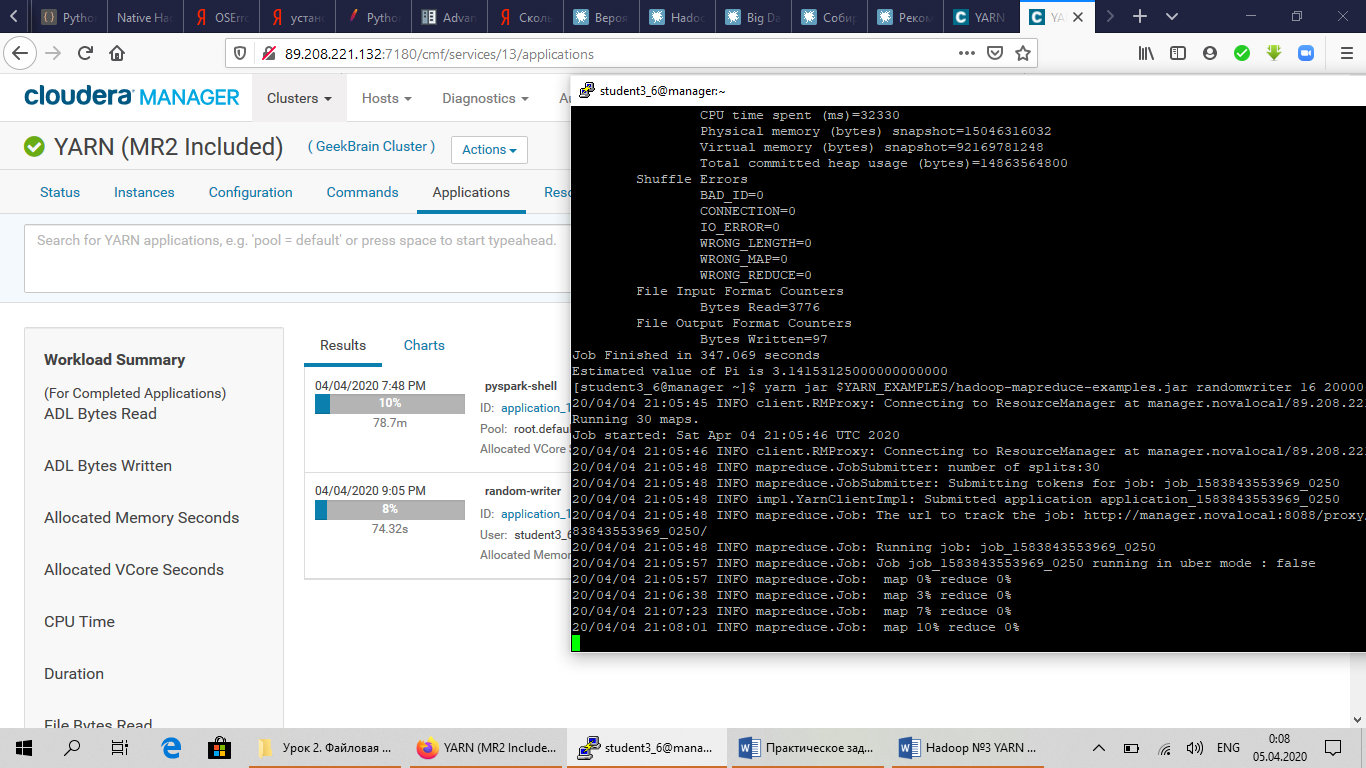


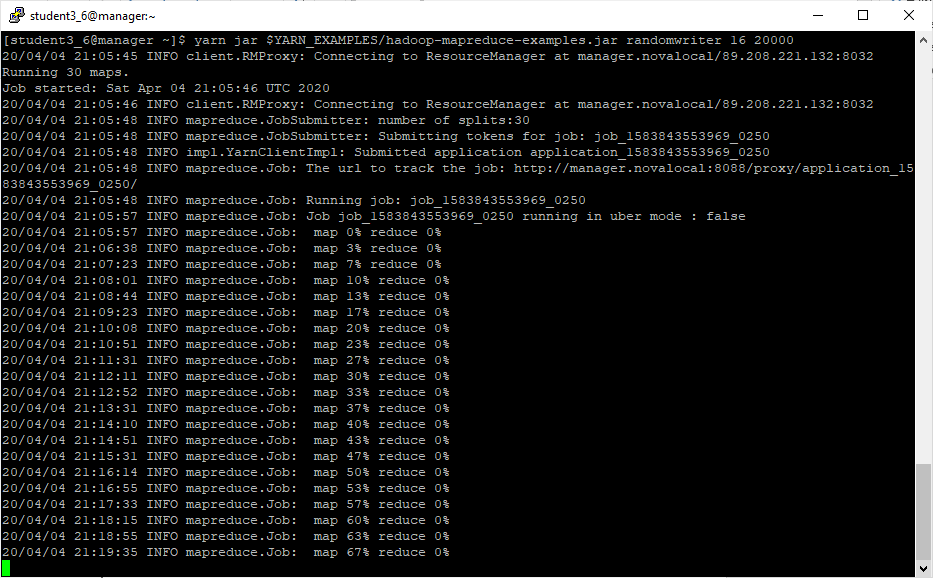


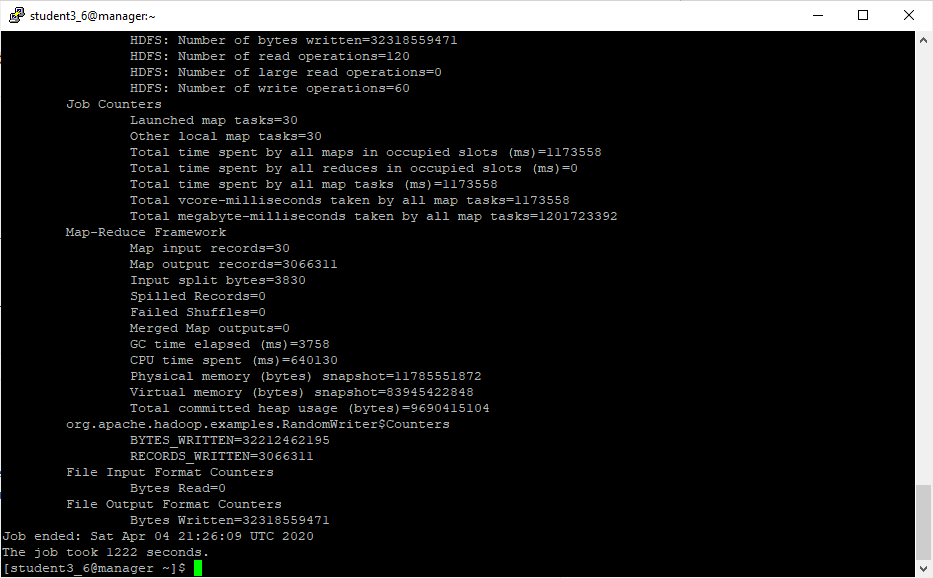


2. Выполнить три любых задачи включенных в этот JAR.

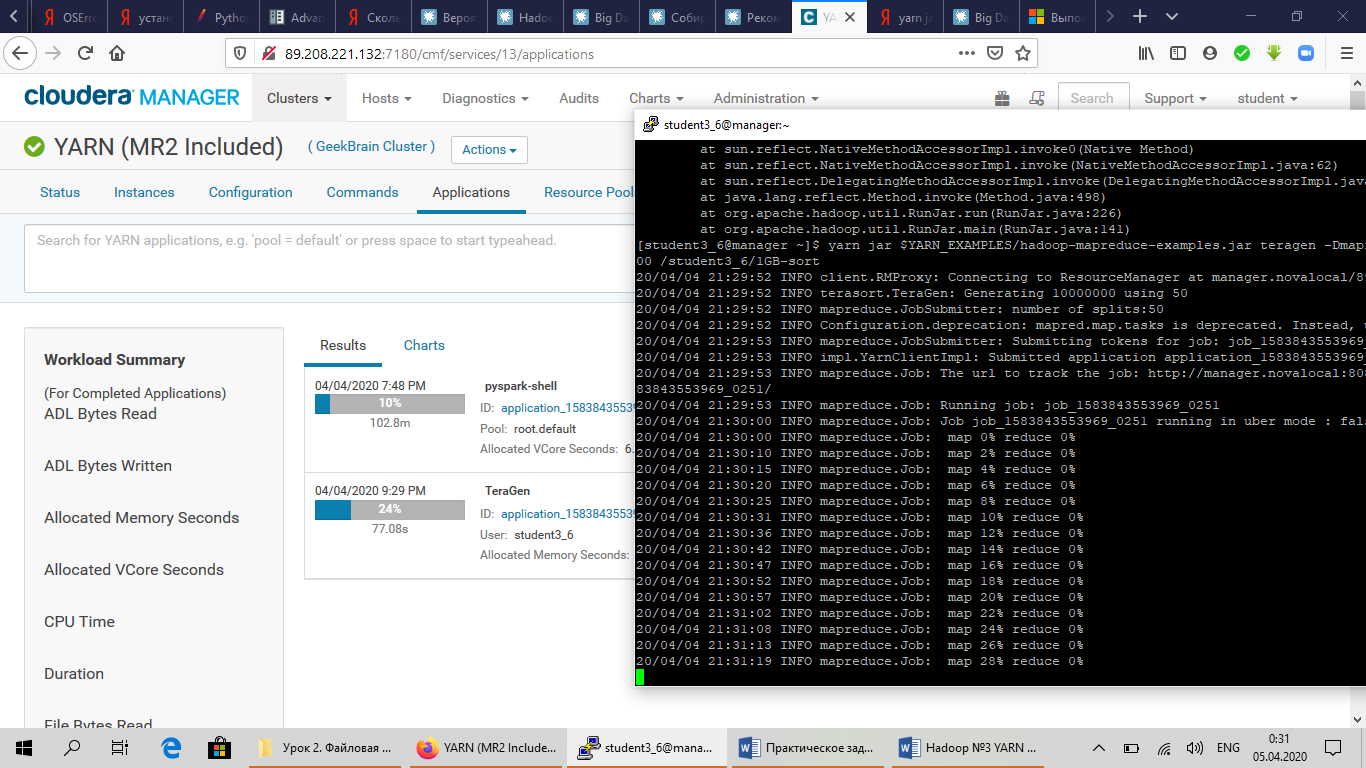
2.1. randomwriter

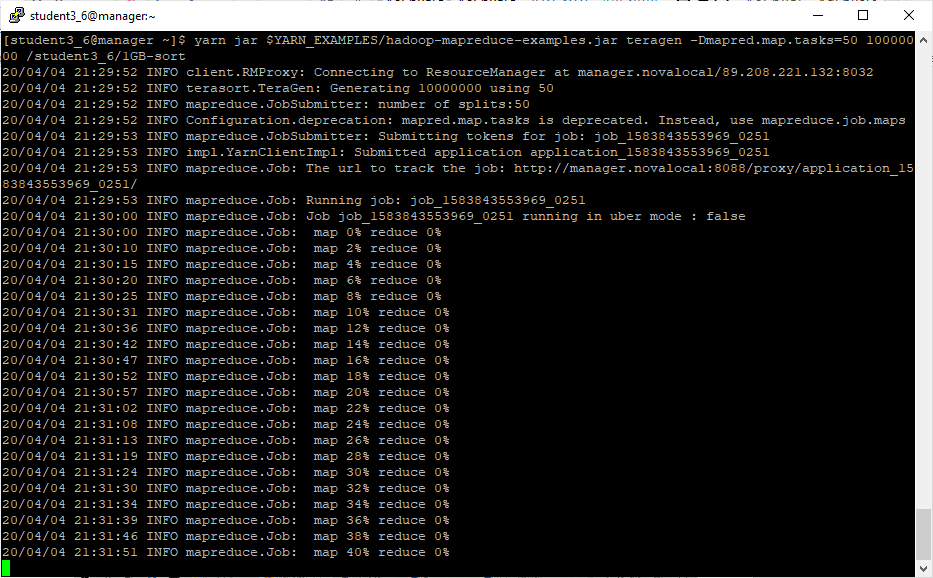


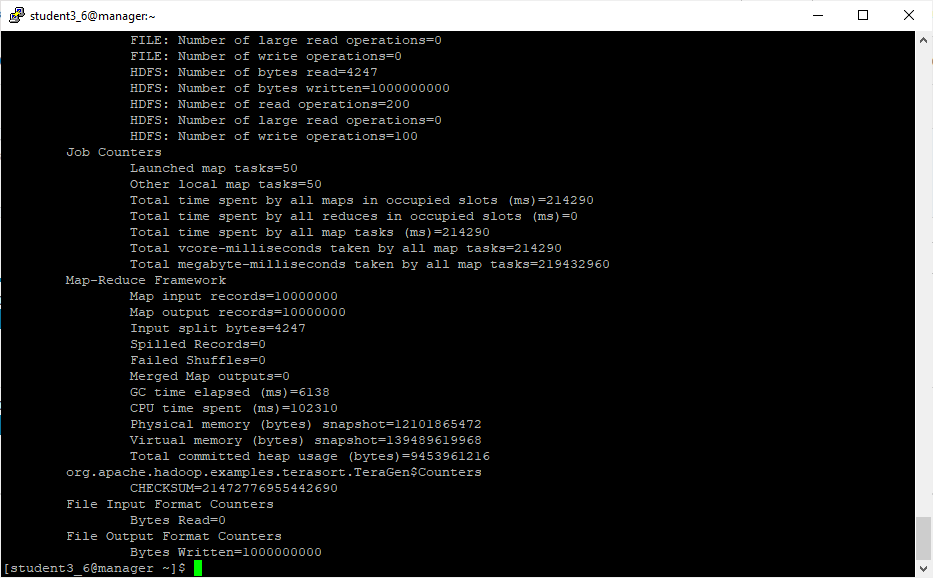




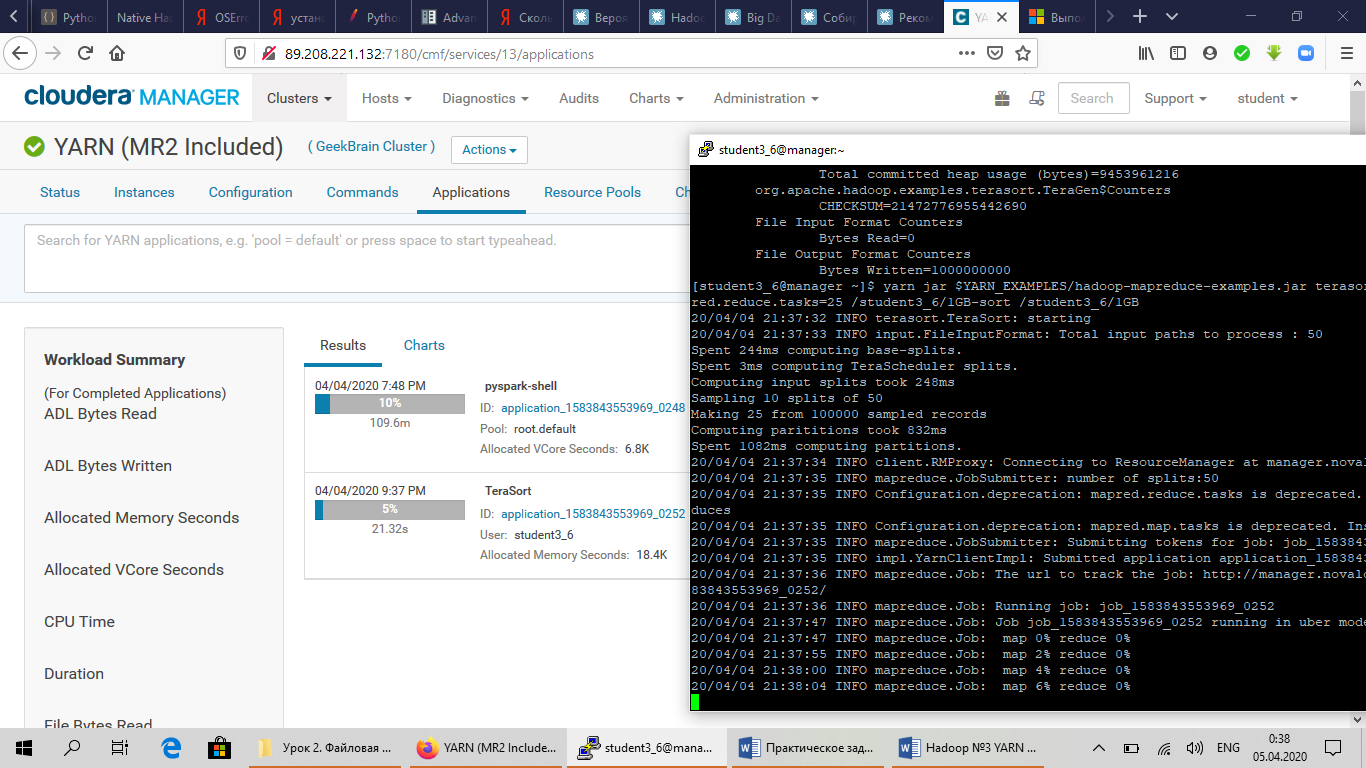
2.2. teragen

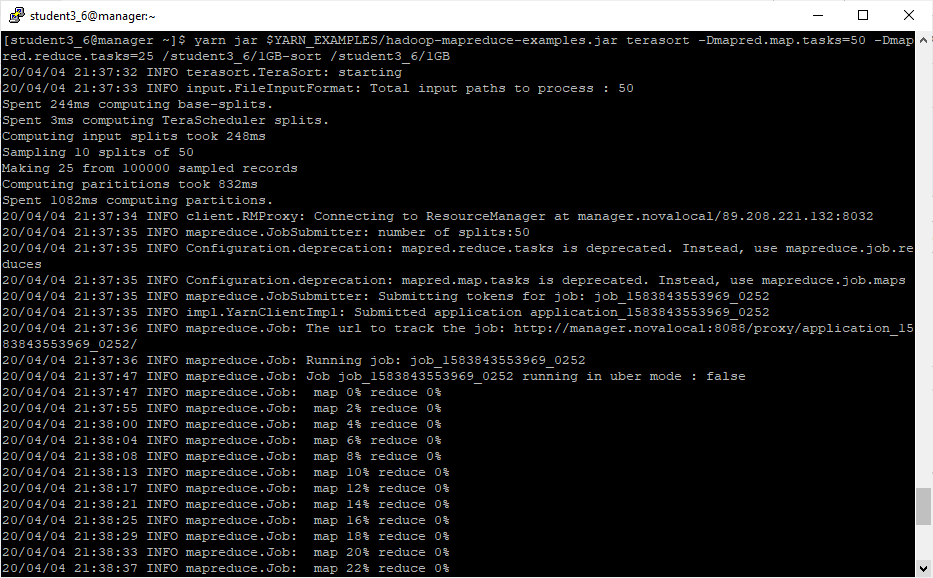




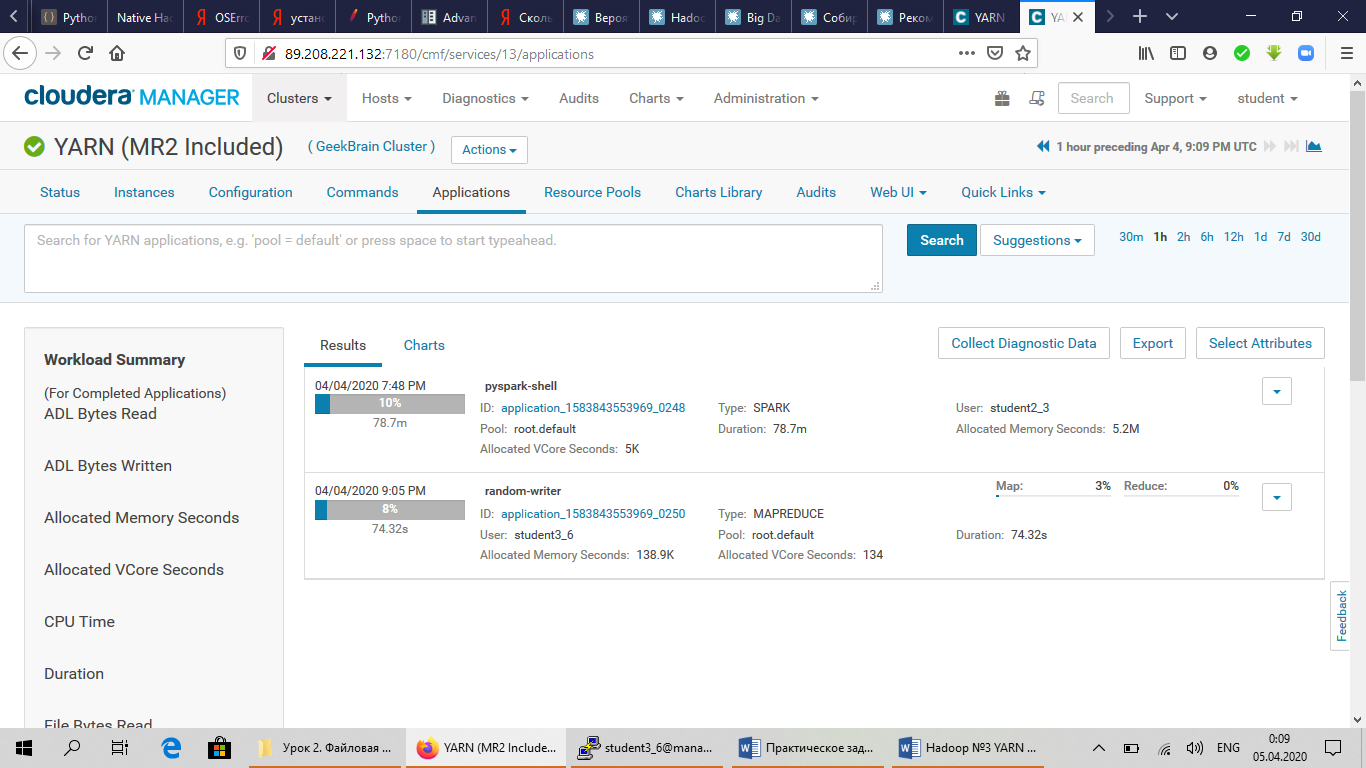


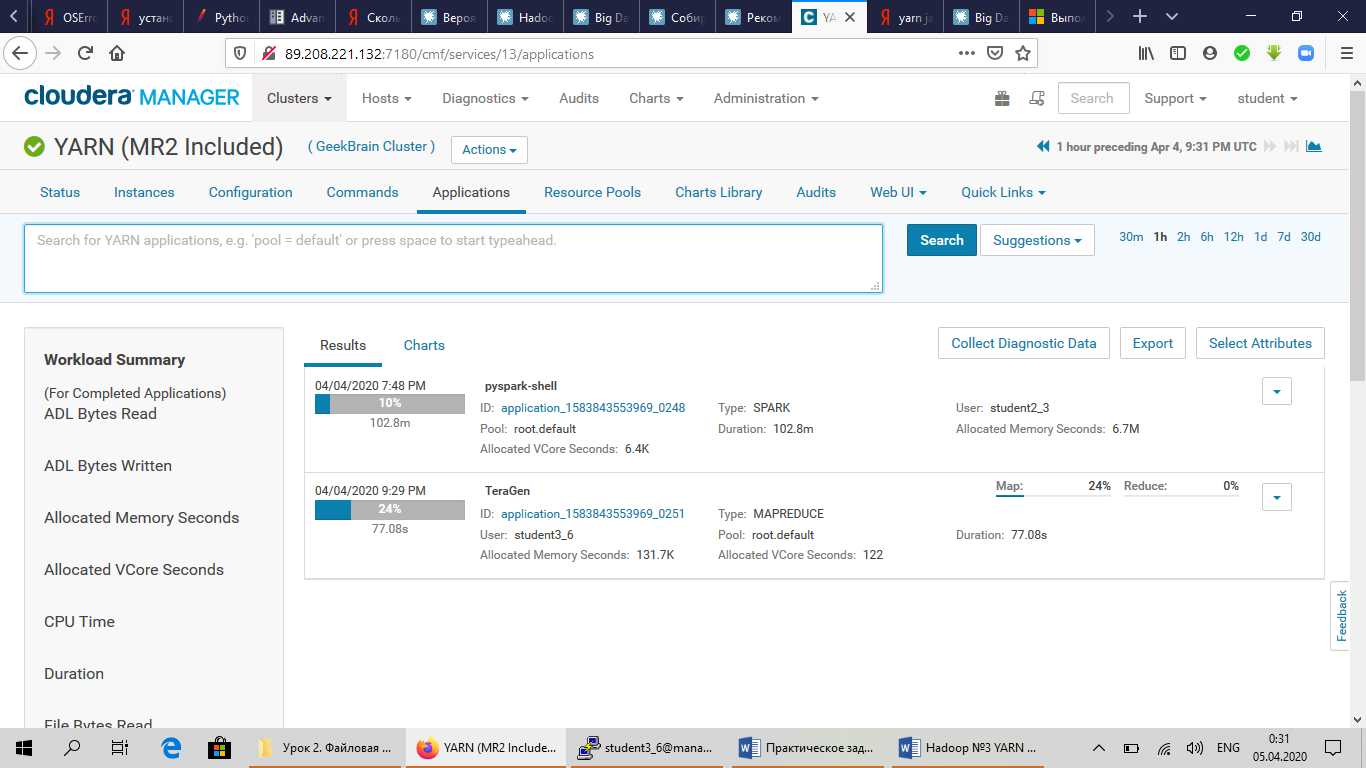
2.3. terasort

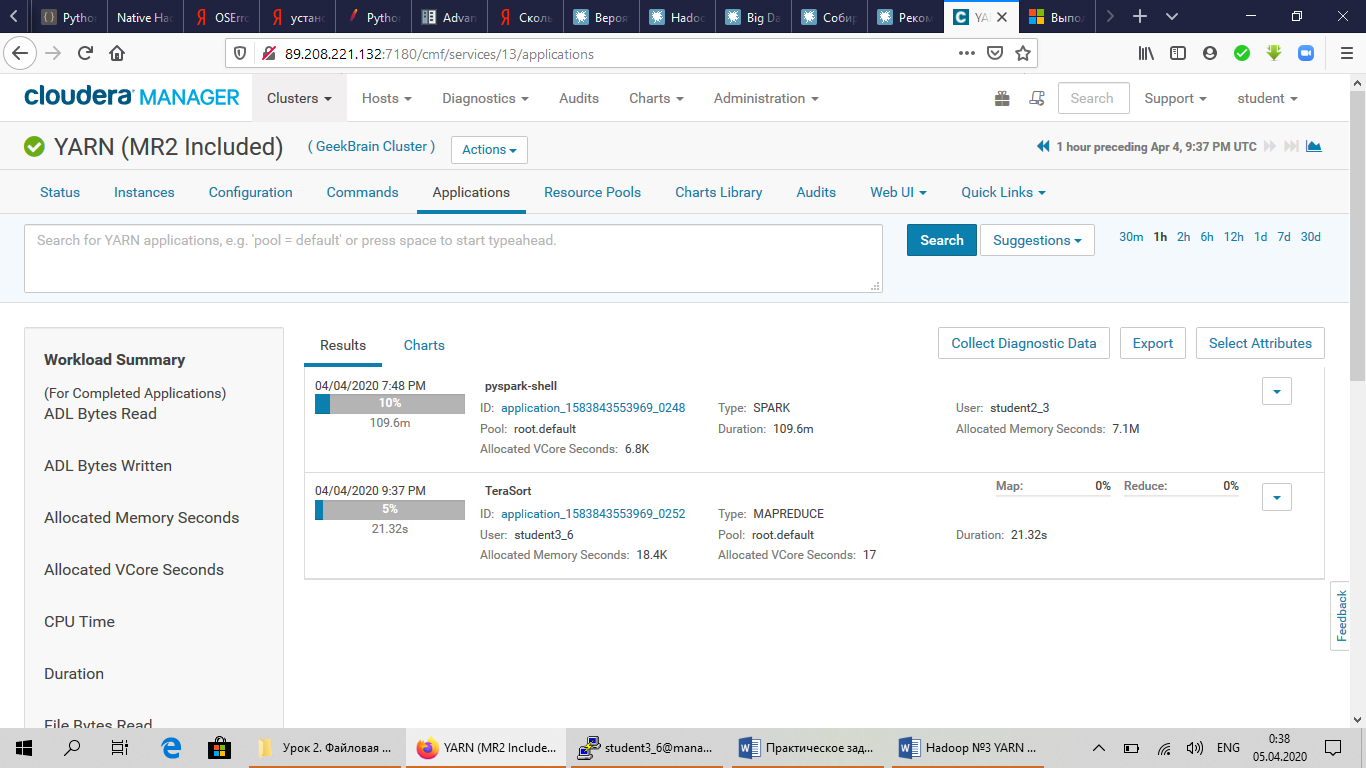




3. Найти свои задачи в интерфейсе Cloudera Manager







4. Опробовать навигацию по интерфейсу YARN

Навигацию опробовал.

5. Сделать документ со скриншотами того, чтовы видели.

Документ создан.

6. [Факультативное, для тех кто знает JAVA] Собрать программу для MR на Java и запустить ее. Wordcount будет вполне достаточен.

7. [Задание на 5++] Повторить вот этот пример https://www.michael-noll.com/tutorials/writing-an-hadoop-mapreduce-program-in-python/