МИНИСТЕРСТВО НАУКИ и ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Курганский государственный университет»

(ФГБОУ ВО «КГУ»)

Кафедра «Безопасность информационных и автоматизированных систем»

ОТЧЕТ

По лабораторной работе №5

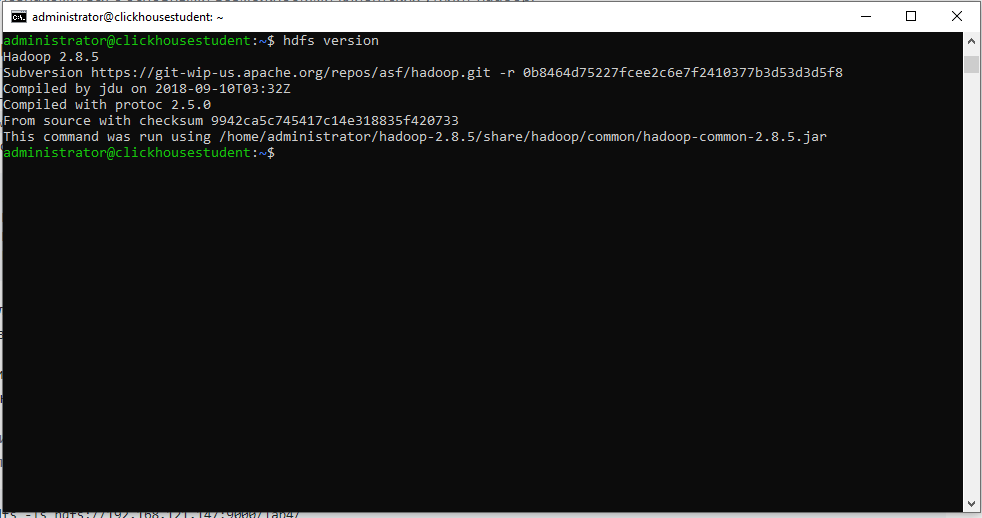
Работу выполнил

студент группы ИТ-40917

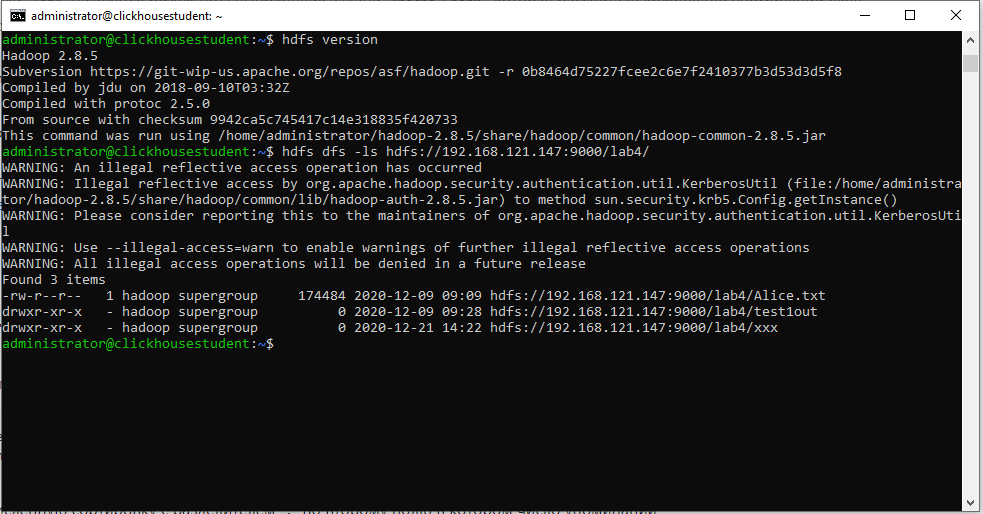
Скутин Роман

Курган 2020

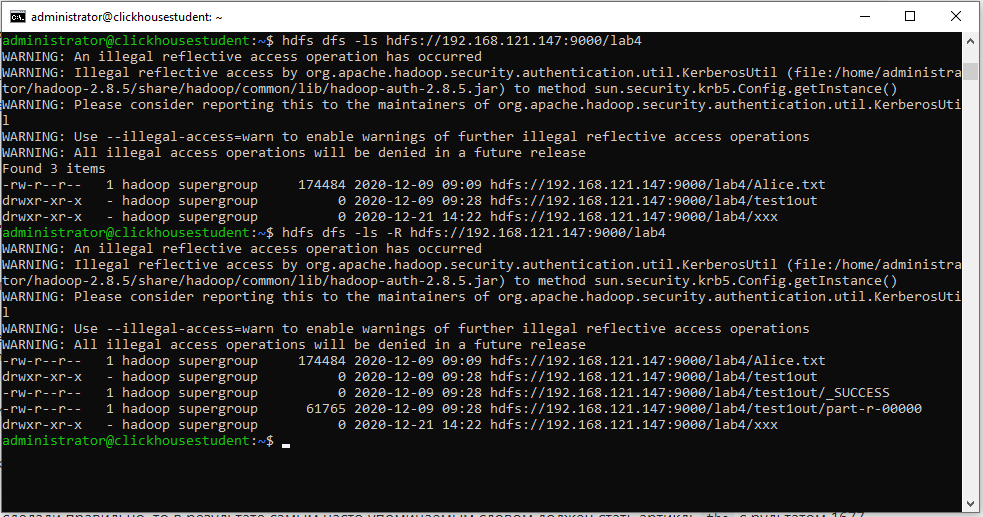
Посмотрим версию hdfs



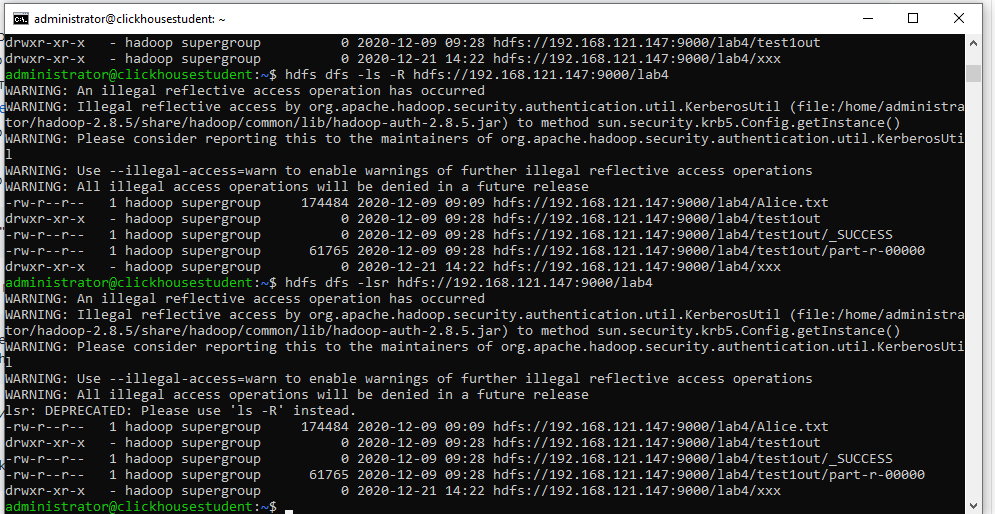
Посмотрим список файлов в папке lab4 на удаленном сервере



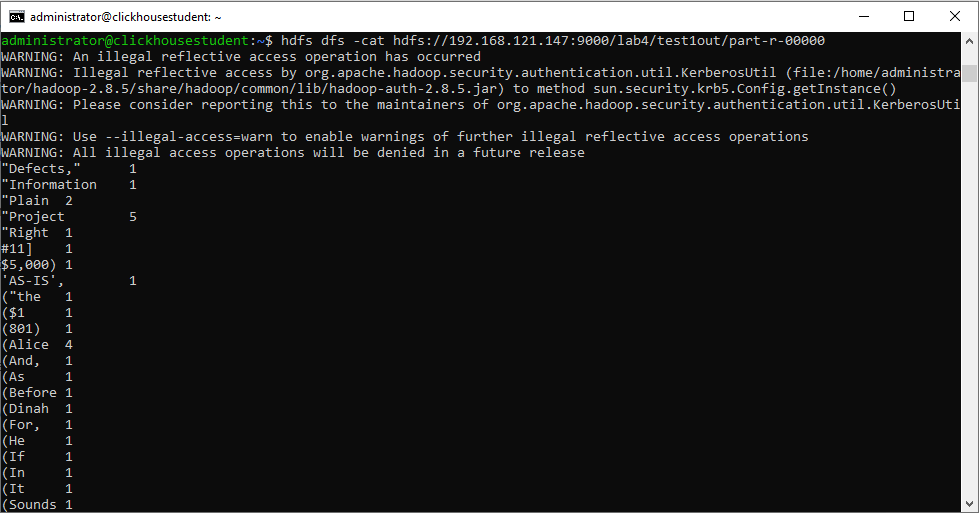
Найдём файл файл part-r-00000 с помощью команды –ls и ключа -R



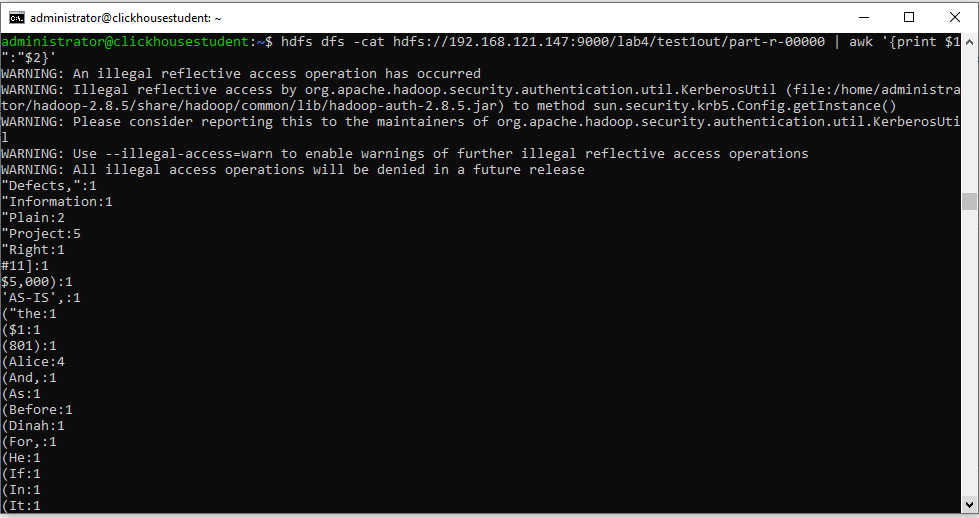
Аналогичный результат даёт команда -lsr



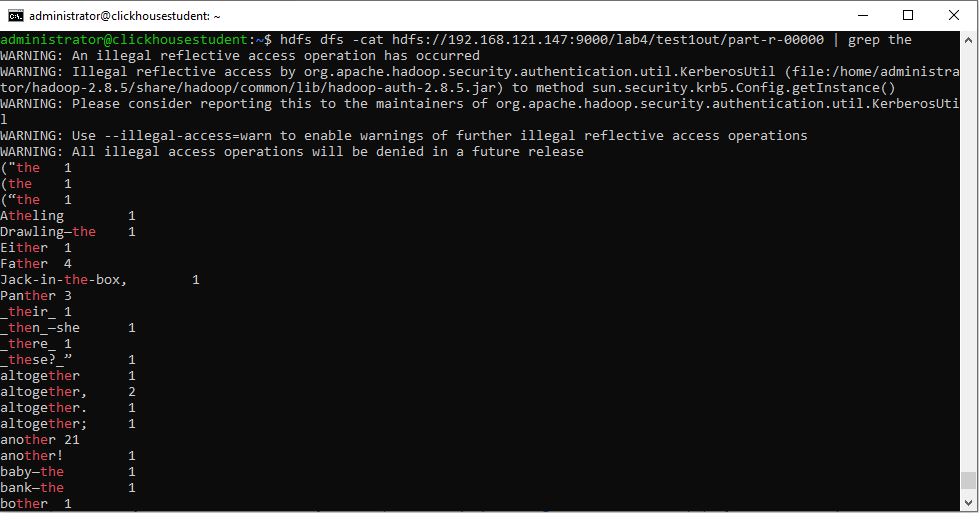
Прочитаем файл part-r-00000 с помощью команды cat



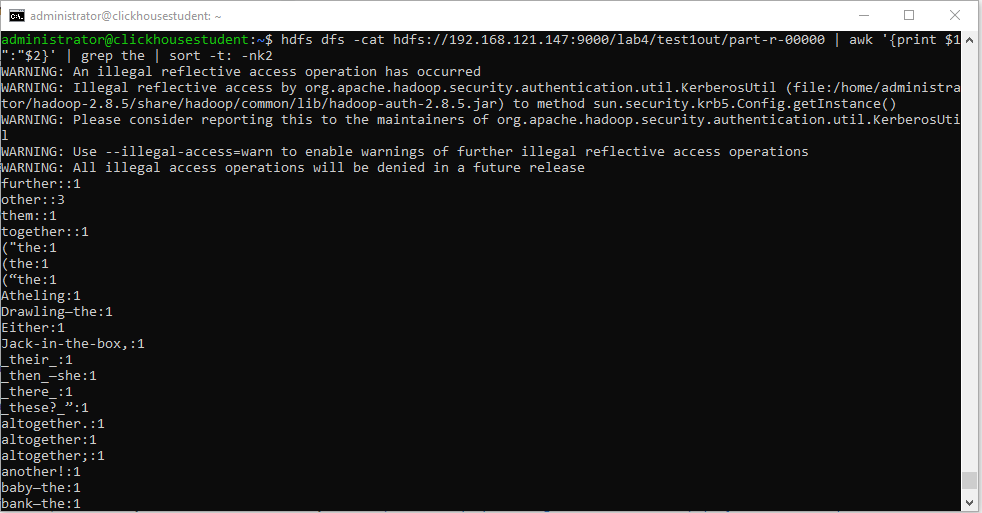
С помощью интерпретатора awk разобьём содержимое файла part-r-00000 на колонки с разделителем



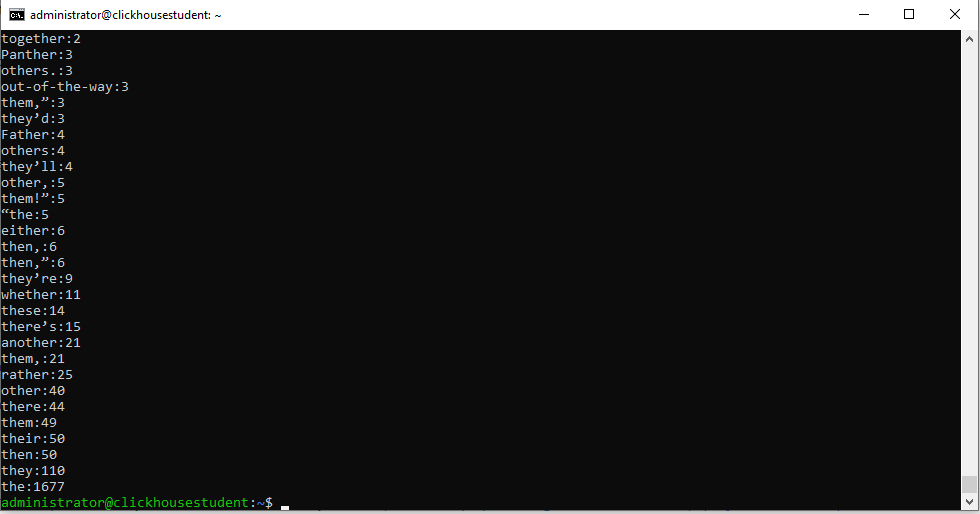
Найдем все “the” в тексте



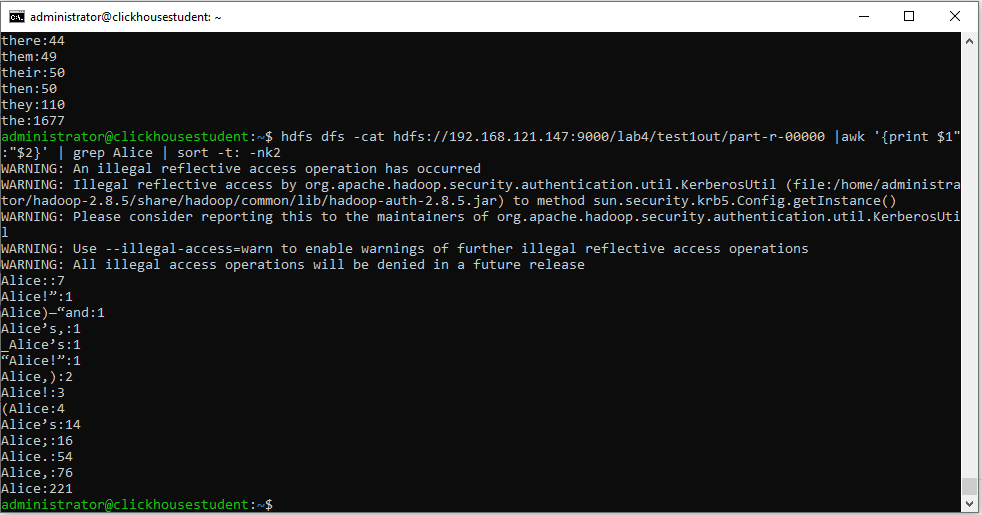
Численно отсортируем



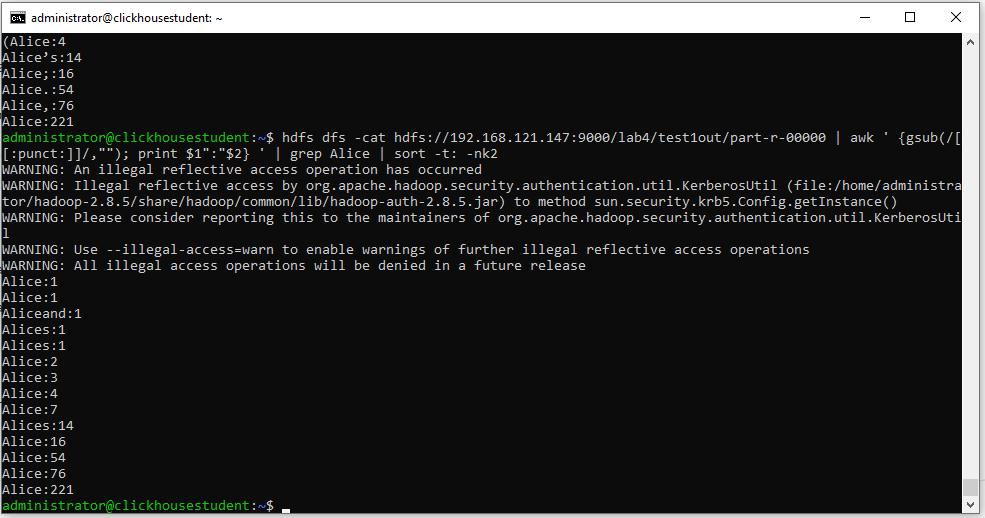
Самым упоминаемым стало слово the с числом вхождений 1677



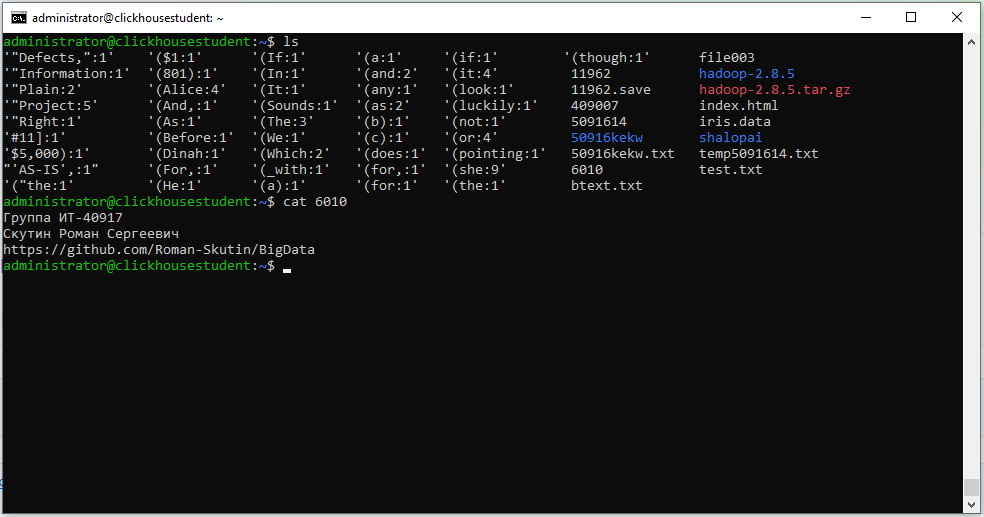
Найдём слово Alice



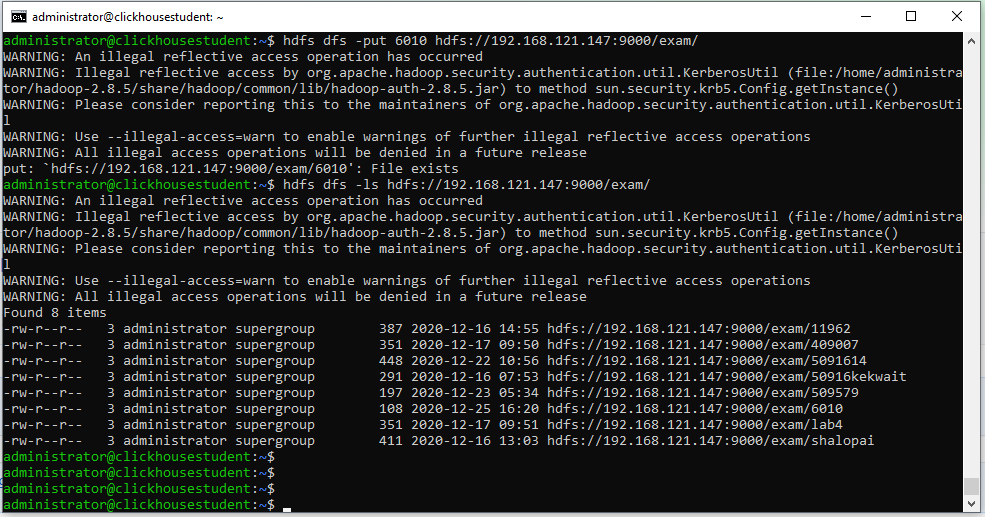
Чтобы получить корректный результат уберём знаки препинания



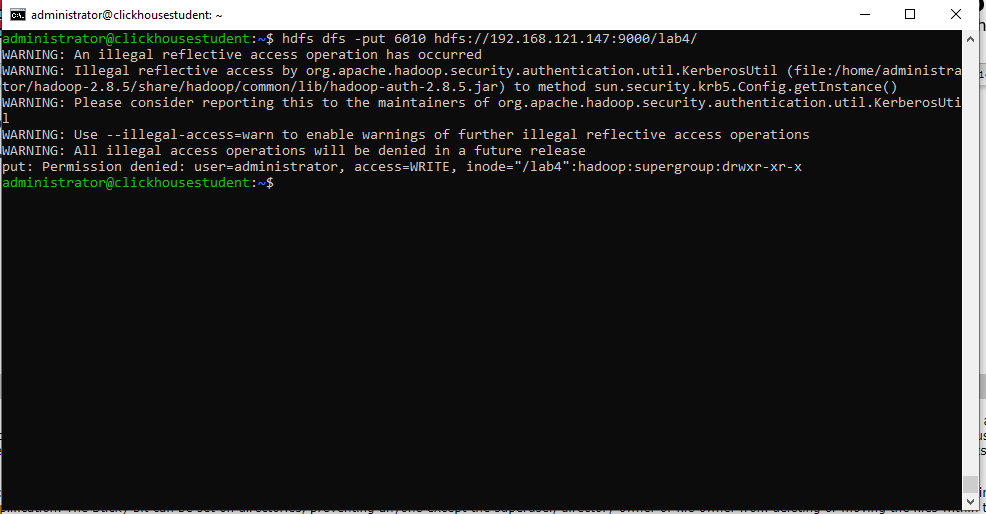
Создадим на диске файл и запишем туда состав группы



Переместим файл в каталог exam и проверим его наличие



Попробуем поместить файл в каталог lab4



В доступе отказано, так как administrator не входит в группу supergroup. При обращении к хадупу от имени пользователя hadoop файл удается записать.



