МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования

«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»

КАФЕДРА № 43

ОТЧЁТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Старший преподаватель | |  |  | | | |  | | М. Д. Поляк |
| должность, уч. степень, звание | |  | подпись, дата | | | |  | | инициалы, фамилия |
| ОТЧЁТ О ЛАБОРАТОРНОЙ РАБОТЕ №2 | | | | | | | | | | |
| работа с текстовыми потоками в командном интерпретаторе Bash | | | | | | | | | | |
| по дисциплине: Операционные системы | | | | | | | | | | |
|  | | | | | | | | | | |
| РАБОТУ ВЫПОЛНИЛА | | | | | | | | | | |
| СТУДЕНТКА ГР. | 4931 | | |  | 04.03.2022 |  | | Е.Ю. Ильченко | | |
|  |  | | |  | подпись, дата |  | | инициалы, фамилия | | |
|  |  | | |  |  |  | |  | | |

Санкт-Петербург 2022

## Вариант 11

## Цель:

#### Изучение принципов работы с командным интерпретатором GNU/Linux и основ обработки текстовых файлов с помощью команд grep, awk, sed.

## Вариант 11:

Подсчитать количество уникальных пользователей, использующих DNS-туннели (см. предыдущий вариант). Пользователей можно идентифицировать по их IP-адресам. Определить количество DNS–запросов, выполненных каждым из найденных пользователей. В файл results.txt вывести таблицу, в которой каждая строка имеет вид: <IP-адрес><символ табуляции><количество запросов с использованием DNS-туннелей, выполненных с указанного IP-адреса> Строки в файле отсортировать в порядке убывания числа запросов к доменам DNS-туннелей. Пользователей с одинаковым количеством запросов дополнительно отсортировать по возрастанию IP-адреса. Для этого можно, например, использовать повторную сортировку с ключом -s или --stable. В переменную VAR\_2 записать количество строк в файле results.txt.

#### Описание входных данных:

Файл dns-tunneling.log содержит логи [DNS-сервера](https://ru.wikipedia.org/wiki/DNS), представленные в виде текстового файла, в котором каждая строка соответствует записи о поступившем на вход сервера запросе. В логах сохраняются следующие параметры запроса, разделенные символом табуляции:

1. Название провайдера телекоммуникационных услуг: character array,
2. Название узла, на котором хранятся данные: character array,
3. Порядковый номер запроса: long,
4. Отметка времени, когда поступил запрос: два числа long, разделенных точкой; первое число – количество секунд, прошедших с 1 января 1970 года; второе число – количество микросекунд; т.е. фактически это тип данных float,
5. IP-адрес пользователя: character array,
6. Порт пользователя: int,
7. Локальный IP-адрес, на который поступил запрос: character array,
8. Локальный порт: int,
9. Название оборудования DNS-сервера: character array,
10. Класс запроса: int,
11. [Тип запроса](https://ru.wikipedia.org/wiki/%D0%A2%D0%B8%D0%BF%D1%8B_%D1%80%D0%B5%D1%81%D1%83%D1%80%D1%81%D0%BD%D1%8B%D1%85_%D0%B7%D0%B0%D0%BF%D0%B8%D1%81%D0%B5%D0%B9_DNS): int,
12. Код возвращаемого значения: int,
13. Флаги: int,
14. Вспомогательный идентификатор: int,
15. Запрашиваемый URL: character array,
16. Зона: character array,
17. Вспомогательное поле 1: character array,
18. Вспомогательное поле 2: character array,
19. Вспомогательное поле 3: character array,
20. Вспомогательное поле 4: character array,
21. Ответ сервера: character array,
22. Вспомогательное поле 5: character array,
23. Вспомогательное поле 6: character array,
24. Длина ответа: int

**Результат:**

10.114.54.213 20558

10.100.40.92 15705

10.17.128.209 14470

10.17.71.70 13942

10.18.31.27 12883

10.17.246.221 12422

10.55.96.230 12337

10.13.6.44 11141

10.117.181.62 10144

10.1.178.208 7699

10.19.25.188 7598

10.119.76.168 5973

10.116.74.176 5719

10.46.147.115 5489

10.37.183.0 5152

10.3.66.214 4789

10.21.32.22 4068

10.25.221.43 3705

10.37.149.47 3399

10.115.60.216 3164

10.20.21.144 3051

10.32.88.226 2498

10.114.2.25 2486

10.28.238.95 2373

10.26.63.12 2229

10.27.228.227 2094

10.111.186.66 1624

10.115.95.162 1478

10.10.150.255 1218

10.94.112.46 1092

10.115.205.160 1069

10.19.169.52 965

10.115.34.104 931

10.15.220.154 826

10.100.160.104 681

10.16.242.146 652

10.47.144.132 565

10.9.83.18 552

10.38.214.32 485

10.9.13.199 467

10.15.133.112 445

10.7.133.249 378

10.17.82.32 367

10.30.52.97 346

10.37.81.64 272

10.111.174.73 226

10.15.221.249 135

10.112.191.146 114

10.1.208.241 89

10.7.227.42 88

10.21.19.205 75

10.28.80.84 67

10.116.92.145 47

10.25.124.1 40

10.54.111.5 38

10.30.135.249 36

10.38.136.233 28

10.98.4.244 26

10.16.187.246 24

10.115.244.135 23

10.16.158.53 20

10.116.91.127 20

10.98.131.43 17

10.36.207.183 13

10.10.57.39 12

10.98.191.141 8

10.35.29.120 7

10.28.149.104 4

10.34.164.149 4

10.116.100.212 4

10.30.179.196 1

10.37.25.38 1

10.97.218.106 1

10.117.118.174 1

**Исходный код программы с комментариями:**

#!/usr/bin/env bash

# edit the code below and add your code

# отредактируйте код ниже и добавьте свой

cat dns-tunneling.log |egrep '(1yf.de.|2yf.de.)'|awk '{print $5}'|sort|uniq -c| awk '{print $2 "\t" $1}'|sort -k2nr -Vk1 > results.txt

# Переменная с номером варианта (константа):

TASKID=11

# Дополнительные переменные (должны вычисляться динамически):

VAR\_1=$(cat dns-tunneling.log |wc -l)

VAR\_2=$(cat results.txt |wc -l)

**Выводы:**

* Изучили принципы работы с командным интерпретатором GNU/Linux и основы обработки текстовых файлов с помощью команд grep, awk, sed.
* egrep '(1yf.de.|2yf.de.)' – используется для поиска строк, содержащих 1yf.de. и 2yf.de.
* awk '{print $5} – из найденных строк получаем столбец с IP адресами.
* sort – сортируем IP адреса
* uniq –c – оставляем только уникальные строчки и считаем их повторение.
* awk '{print $2 "\t" $1}' – после сортировки у нас получилась таблица с двумя столбцами: 1 – повторения; 2 – IP. Меняем их местами (необходимо по заданию).
* sort -k2nr -Vk1 – Сортируем первый столбец в порядке убывания. Если есть повторения, IP адреса сортируются в порядке возрастания.