МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»

КАФЕДРА КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ

И ПРОГРАММНОЙ ИНЖЕНЕРИИ (КАФЕДРА №43)

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Старший преподаватель |  |  |  | М.Д. Поляк |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ОТЧЁТ О ЛАБОРАТОРНОЙ РАБОТЕ № 2 | | | | | |
| **ЛР2. Работа с текстовыми потоками в командном интерпретаторе Bash** | | | | | |
| по дисциплине: ОПЕРАЦИОННЫЕ СИСТЕМЫ | | | | | |
|  | | | | | |
| РАБОТУ ВЫПОЛНИЛ | | | | | |
| СТУДЕНТ ГР. | 4936 |  |  |  | Е.А. Цыганкова |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург  
2022

**Цель работы:**

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

**Индивидуальное задание**

Во всех вариантах предполагается, что в конце каждой строки с текстом стоит символ перевода строки \n, а пустые строки отсутствуют.

20. Найти все доменные имена второго уровня, находящиеся в доменной зоне net.. Подсчитать количество запросов к каждому из найденных доменов второго уровня, включая его поддомены. В файл results.txt вывести таблицу, в которой каждая строка имеет вид:

<доменное имя><символ табуляции><количество запросов этого доменного имени>

Строки в файле отсортировать в порядке убывания числа запросов к доменам. Если число запросов к нескольким доменам одинаковое, отсортировать их по имени домена в алфавитном порядке. Предварительно все доменные имена привести к нижнему регистру.

В переменную VAR\_2 записать количество запросов к домену второго уровня, оказавшемуся на третьей строке в файле results.txt.

**Описание входных данных**

Файл dns-tunneling.log содержит логи DNS-сервера, 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: 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 выполнения работы**

Содержимое файла results.txt:

whatsapp.net. 2270

akamaihd.net. 1108

doubleclick.net. 330

gpsonextra.net. 267

fbcdn.net. 255

akamai.net. 217

ksmobile.net. 187

cloudfront.net. 164

miisolutions.net. 160

cap-mii.net. 80

kingroot.net. 68

duba.net. 67

edgesuite.net. 58

live.net. 37

opera-mini.net. 33

akadns.net. 29

surpax.net. 29

akamaiedge.net. 23

ribob01.net. 22

swiftkey.net. 19

pubnative.net. 17

your-freedom.net. 15

gpstream.net. 14

playstation.net. 14

izatcloud.net. 13

fastly.net. 10

msecnd.net. 10

online-metrix.net. 9

acompli.net. 7

adnitro.net. 6

ospserver.net. 6

facebook.net. 5

googlezip.net. 5

supercell.net. 5

apple-dns.net. 4

att.net. 4

fotoable.net. 4

ftpub.net. 4

nflximg.net. 3

thetrafficstat.net. 3

tubemate.net. 3

adecosystems.net. 2

demdex.net. 2

fingersoft.net. 2

tapas.net. 2

adfection.net. 1

apxadtracking.net. 1

azurewebsites.net. 1

deviantart.net. 1

entrust.net. 1

everesttech.net. 1

getui.net. 1

hockeyapp.net. 1

inmobicdn.net. 1

ltkcdn.net. 1

mozaws.net. 1

netai.net. 1

netau.net. 1

netne.net. 1

omtrdc.net. 1

onclickads.net. 1

site40.net. 1

site90.net. 1

tapsong.net. 1

via.net. 1

**Исходный код программы с комментариями**

#! /usr/bin/env bash

# edit the code below and add your code

# отредактируйте код ниже и добавьте свой

# очищаем файлы вывода

echo -n > results.srt

echo -n > results.txt

sep=$'\t'

# в цикле перебираем все домены, подходящие условию

# записываем в файл имя домена и считаем количество строчек в файле логов

# с данным доменом. выводим в файл результатов

for i in `cat dns-tunneling.log |awk '{print $15}'| grep "\.net\.$" | sed 's/^.\*\.\(.\+\.net\.\)$/\1/' | sort -u`; do

echo -e "$i\t`grep "$i$sep" dns-tunneling.log | wc -l`" | tr [:upper:] [:lower:] >> results.srt

done

# сортируем файл

sort -rnk2 -k1d results.srt > results.txt

# Переменная с номером варианта (константа):

TASKID=20

# Дополнительные переменные (должны вычисляться динамически):

# количество строк в файле

VAR\_1=`wc -l dns-tunneling.log | awk '{print $1}'`

# количество обращений (2 колонка) к домену на третьей строчке в файле results.txt

VAR\_2=`sed -n 3p results.txt | awk '{print $2}'`

**Выводы**

В ходе выполнения лабораторной работы были изучены принципы работы с командным интерпретатором GNU/Linux, освоены основы обработки текстовых файлов с помощью команд grep, awk, sed.