

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский университет ИТМО»

ФАКУЛЬТЕТ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Лабораторная работа № 2
по теме «Обработка и тарификация трафика NetFlow»
по дисциплине «Управление мобильными устройствами»

Работу выполнил
студент группы №3351
очного отделения
Меренков Даниил



Проверил

Федоров Иван Романович

Дата сдачи: 11.04.2020

Санкт-Петербург
2020



Цель работы(задача):

Цель данной работы – изучение протокола NetFlow, его выходных файлов данных, парсинг этих файлов, составление диаграммы зависимости трафика от времени, а так же тарификация абонента, в соответствии с заданными условиями.

Описание выбранных средств реализации и обоснование выбора:

Исходные данные: Протарифицировать абонента с IP-адресом 192.168.250.62 с коэффициентом k: 0,5руб/Мб первые 100Мб, далее 1руб/Мб

Для реализации программного модуля в данной лабораторной работе мною был выбран язык Python3 из-за его легкости чтения, гибкости и практичности в задачах, касающихся парсинга.

При реализации данной программы была задействована библиотека matplotlib для рисования графика по переданным параметрам и сохранения его в файл .png, а так же утилита nfdump для составления нормализованного вида исходных данных.

Пример использования утилиты nfdump.

```
su-adm@su-main:~/trash$ nfdump -r nfcapd.202002251200 -q -o "fmt:%tsr,%sap,%dap,%byt" > output.csv
su-adm@su-main:~/trash$
```

Где ключ -o позволяет задать параметр вывода в нужном формате fmt

%sap	Source Address:Port
%dap	Destination Address:Port
%byt	Bytes - default input
%tsr	Start Time, but in fractional seconds since the epoch (1970-01-01)

На выход получаем csv файл с исключительно необходимой информацией – time,srcIP,dstIP и суммарное кол-во байт при данном обращении.

Вывод программы:

```

27 ip_addr = "192.168.250.62"
28 pivot = {}
29 first_b = 100
30 first_b_cost = 0.5
31 later_b_cost = 1
32 input_file = open("traf.csv", "r")
33 cost = 0
34 total_bytes = 0
35 for line in input_file:
36     line = line.replace("\n", "").split(",")
37     if ip_addr in line[1] or ip_addr in line[2]:
38         if "M" in line[3]:
39             mb = int(float(line[3][:-1])*1024*1024)
40             total_bytes += mb
41             if int(float(line[0])) in pivot:
42                 pivot[int(float(line[0]))] = pivot[int(float(line[0]))] + mb
43             else:
44                 pivot[int(float(line[0]))] = mb
45         else:
46             total_bytes += int(line[3])
47             if int(float(line[0])) in pivot:
48                 pivot[int(float(line[0]))] = pivot[int(float(line[0]))] + int(line[3])
49             else:
50                 pivot[int(float(line[0]))] = int(line[3])
51
52     # Проверяем распространяется ли Примечание 2 на наши исходные данные.
53     if total_bytes/1024 > first_b:
54         if total_bytes/(1024*1024) > first_b:
55             cost = first_b*first_b_cost + (int(total_bytes/(1024*1024))-first_b)*later_b_cost
56             # считаем в мб
57         else:
58             cost = first_b*first_b_cost + (int(total_bytes/1024)-first_b)*later_b_cost
59             # считаем в кб
60     elif total_bytes > first_b:
61         cost = first_b*first_b_cost + (total_bytes-first_b)*later_b_cost
62         # считаем в байтах
63     else:
64         cost = 0
65         print("Невозможно рассчитать абонента в связи с малым кол-вом потраченного трафика")
66         exit(99)
67
68     print(f"Итоговая сумма: {cost} рублей")
69

```

DEBUG CONSOLE PROBLEMS 1 OUTPUT TERMINAL

```

PS C:\Users\mi\Documents\WORK\3 Курс\Мобилки\Лаба_2>
PS C:\Users\mi\Documents\WORK\3 Курс\Мобилки\Лаба_2>
PS C:\Users\mi\Documents\WORK\3 Курс\Мобилки\Лаба_2>
PS C:\Users\mi\Documents\WORK\3 Курс\Мобилки\Лаба_2> & C:/Users/mi/AppData/Local/Programs/Python/Python37-32/python.exe
Итоговая сумма: 33998.0 рублей
PS C:\Users\mi\Documents\WORK\3 Курс\Мобилки\Лаба_2>

```

Согласно Примечанию 2, так как в наших исходных данных кол-во трафика абонента слишком мало, мы уменьшаем единицу учета до байт -> исходное задание меняется на

«Протарифицировать абонента с IP-адресом 192.168.250.62 с коэффициентом k: 0,5руб/байт первые 100 байт, далее 1руб/байт».

Ответом на данные условия является: 33998 рублей

Так же, собираем статистику о зависимости трафика от времени для построения графика и визуализации данных. Для этого суммируем общий объем трафика для нужного абонента на каждую зарегистрированную секунду, согласно UNIX-времени(%tsr) и с помощью библиотеки matplotlib строим график, где по оси x – время , а по оси y – трафик. Сохраняем данную статистику в виде графика в файле.

```
import matplotlib.pyplot as plt

def lineplot(x_data, y_data, x_label="", y_label="", title=""):
    # Создаю холст
    fig, ax = plt.subplots()

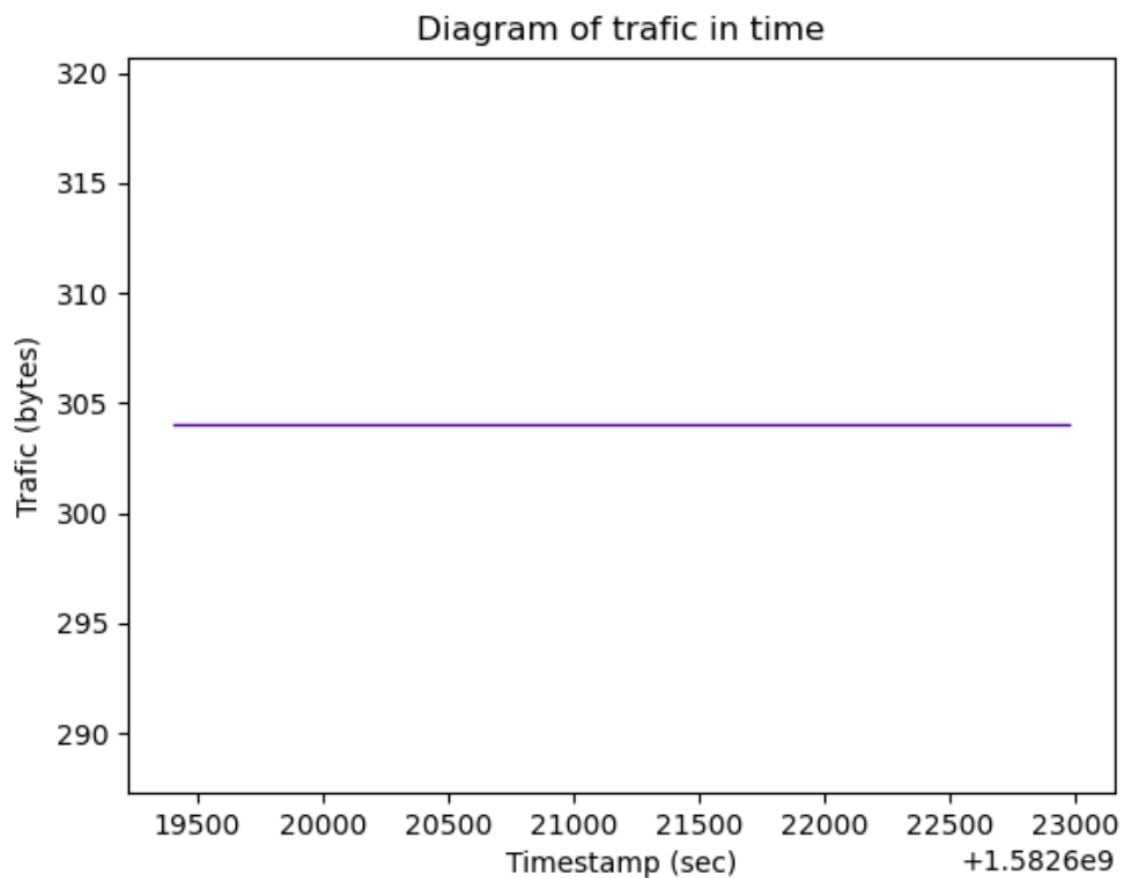
    ax.plot(x_data, y_data, lw = 1, color = '#5A009D', alpha = 1)

    ax.set_title(title)
    ax.set_xlabel(x_label)
    ax.set_ylabel(y_label)
    fig.savefig('diagram.png')
```

```
x_line = []
y_line = []
sorted_income = {k: pivot[k] for k in sorted(pivot)}
for key,value in sorted_income.items():
    x_line.append(key)
    y_line.append(value)

lineplot(x_line,y_line,"Timestamp (sec)","Trafic (bytes)","Diagram of trafic in time")
```

В итоге, получаю следующую зависимость для моих исходных данных:



Исходный код:

```
import matplotlib.pyplot as plt
import os

def lineplot(x_data, y_data, x_label="", y_label="", title=""):
    # Создаю холст
    fig, ax = plt.subplots()

    ax.plot(x_data, y_data, lw = 1, color = '#5A009D', alpha = 1)

    ax.set_title(title)
    ax.set_xlabel(x_label)
    ax.set_ylabel(y_label)
    fig.savefig('diagram.png')

def main_2():

    ip_addr = "192.168.250.62"
    pivot = {}
    first_b = 100
    first_b_cost = 0.5
    later_b_cost = 1
    input_file = open("traf.csv", "r")
    cost = 0
    total_bytes = 0
    for line in input_file:
        line = line.replace("\n", "").split(",")
        if ip_addr in line[1] or ip_addr in line[2]:
            if "M" in line[3]:
                mb = int(float(line[3][:-1])*1024*1024)
                total_bytes += mb
                if int(float(line[0])) in pivot:
                    pivot[int(float(line[0]))] = pivot[int(float(line[0]))] + mb
                else:
                    pivot[int(float(line[0]))] = mb
            else:
                total_bytes += int(line[3])
                if int(float(line[0])) in pivot:
                    pivot[int(float(line[0]))] = pivot[int(float(line[0]))] + int
(line[3])
```

```

        else:
            pivot[int(float(line[0]))] = int(line[3])

# Проверяем распространяется ли Примечание 2 на наши исходные данные.
if total_bytes/1024 > first_b:
    if total_bytes/(1024*1024) > first_b:
        cost = first_b*first_b_cost + (int(total_bytes/(1024*1024))-
first_b)*later_b_cost
        # считаем в мб
    else:
        cost = first_b*first_b_cost + (int(total_bytes/1024-
first_b))*later_b_cost
        # считаем в кб
    elif total_bytes > first_b:
        cost = first_b*first_b_cost + (total_bytes-first_b)*later_b_cost
        # считаем в байтах
    else:
        cost = 0
        print("Невозможно рассчитать абонента в связи с малым кол-
вом потраченного трафика")
        exit(99)

x_line = []
y_line = []
sorted_income = {k: pivot[k] for k in sorted(pivot)}
for key,value in sorted_income.items():
    x_line.append(key)
    y_line.append(value)

lineplot(x_line,y_line,"Timestamp (sec)","Trafic (bytes)","Diagram of trafic
in time")

return cost

if __name__ == '__main__':
    os.chdir('/home/Mobile/Lab_2')
    cost = main_2()
    print(f"Итоговая стоимость: {cost} рублей")
    print("Диаграмма сформированна, она находится в директории Lab_2 под именем d
iagram.png")

```

Выводы:

В рамках данной лабораторной работы, я научился обрабатывать файлы, порождаемые протоколом NetFlow, научился работать с библиотекой matplotlib и генерировать с ее помощью графики зависимостей с выводом в заданный файл. Более того, мной были закреплены такие практические навыки, как парсинг файлов, тарификация заданного абонента и создание диаграмм для данных этого абонента.