

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

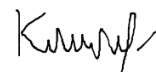
ФАКУЛЬТЕТ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

**Управление мобильными устройствами**  
Лабораторная работа №2  
«Обработка и тарификация трафика NetFlow»

**Выполнила:**

Студентка гр. N3352

Климова Д. А.



Дата защиты: 12.05.2020

**Проверил:**

Федоров И. Р.

Санкт-Петербург

2020 г.

**Цель работы:** обработка трафика NetFlow v5 из данного файла и реализация простейшего правила тарификации этого трафика.

**Задачи:**

1. Привести данный файл в читабельный вид (например, с помощью утилиты nfdump)
2. Сформировать собственный файл для тарификации любого формата, с которым удобно работать (в соответствии с вариантом работы);
3. Построить график зависимости объема трафика от времени;
4. Протарифицировать трафик в соответствии с вариантом задания.

**Ход работы**

Для реализации программного модуля был выбран язык программирования Python. Выбор обусловлен тем, что данный язык программирования является высокоуровневым и простым в работе.

Правила тарификации услуг “Интернет”:

$X=Q*k$ , где  $X$  - итоговая стоимость,  $Q$  - общий объем трафика NetFlow за отчетный период,  $k$  - множитель тарифного плана

В программном модуле использовался модуль pandas для обработки данных csv. Также использовался модуль matplotlib для построения графика зависимости и корректного отображения времени.

Файл nfcapd.202002251200 был преобразован в файл формата CSV при помощи команды `nfdump -r nfcapd.202002251200 -o "fmt:%ts,%sa,%da,%ibyt,%oby" | sed s/"s"/g | head -n -4 > nfdump.csv`

Сначала программа обрабатывает данные файла nfdump.csv. Выбираются строки, содержащие заданный вариант IP-адрес, при необходимости мегабайты переводятся в байты, и в колонке времени отбрасывается дата. Все полученные данные находятся в структуре DataFrame (tr).

После идет подсчет входящего и исходящего трафика, их сложение, а потом умножение на заданный в варианте коэффициент для получения итогового счета.

Для построения графика зависимости объема трафика от времени сортируются уникальные значения в колонке времени. Для каждого значения подсчитывается объем трафика и с помощью полученных значений строится график.

## Вариант 8

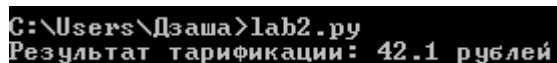
Протарифицировать абонента с IP-адресом 192.168.250.3  
с коэффициентом k: 3руб/Мб

Листинг программы (рис.1):

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.dates import DateFormatter
#обработка данных и тарификация
IP = '192.168.250.3'
kf = 3
tr = pd.read_csv('nfcapd.csv', skiprows=1, header=None)
tr.columns = ['time', 'src', 'dst', 'inbyte', 'outbyte']
tr = tr[np.logical_or(tr.src == IP, tr.dst == IP)]
tr.inbyte = tr.inbyte.apply(lambda row: int(row) if 'M' not in row else (int(float(row[:-1])*10**6)))
tr.time = tr.time.apply(lambda row: row[10:18])
outgoing = tr[tr.src == IP].inbyte.sum() / 10**6
ingoing = tr[tr.dst == IP].inbyte.sum() / 10**6
bill = kf*(outgoing + ingoing)
print(f'Результат тарификации: {bill:0.1f} рублей')
#построение графика зависимости
def graphic(times, size, format=None, save_to=None):
    fig, ax = plt.subplots(figsize=[10,5])
    ax.plot(times, size)
    Format = DateFormatter("%H:%M")
    ax.xaxis.set_major_formatter(Format)
    plt.ylabel('Байт')
    plt.xlabel('Время')
    plt.show()
times = np.sort(tr.time.unique())
size = []
for time in times:
    size.append(tr.loc[tr.time == time, ['inbyte', 'outbyte']].sum().sum())
graphic(pd.to_datetime(times), size)
```

Рис. 1 Листинг кода

Вывод программного модуля (рис.2,3):



```
C:\Users\Дзаша>lab2.py
Результат тарификации: 42.1 рублей
```

Рис.2 Результат тарификации

График зависимости объема трафика от времени:

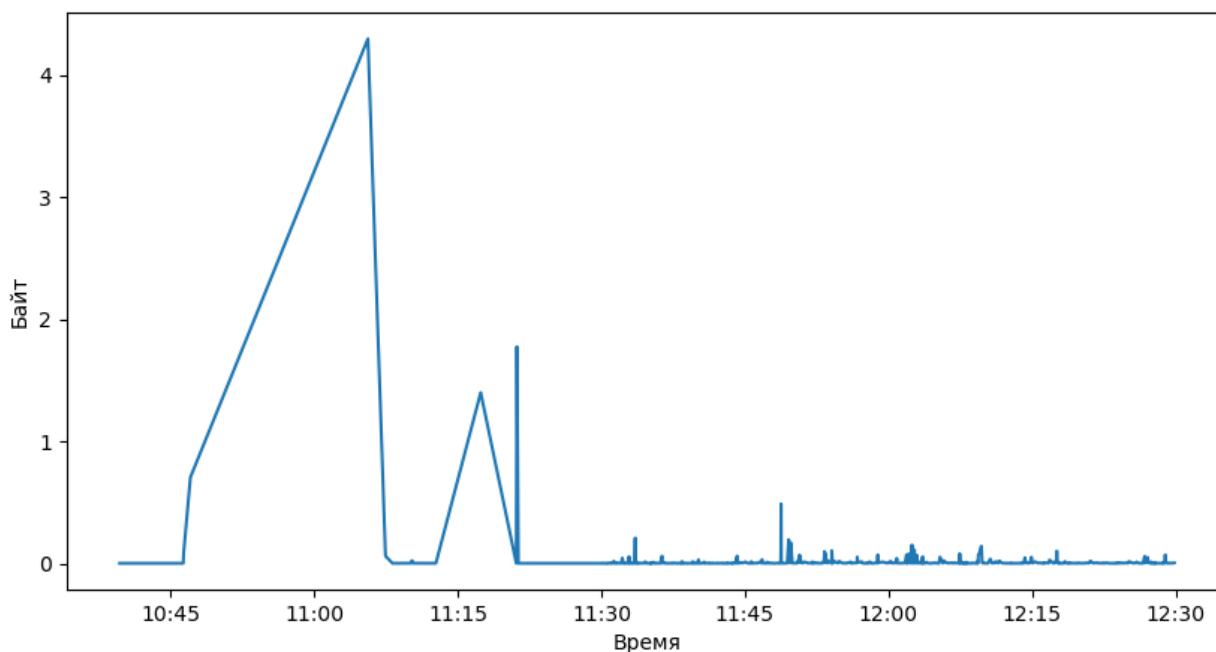


Рис.3 График зависимости

**Вывод:** В ходе данной работы были изучены технологии работы протокола NetFlow, а также разработан и реализован программный модуль обработки трафика NetFlow v5 и тарификации абонента.