ELEC-E7130 - Internet Traffic Measurements and Analysis Assignment 1 Learning to use tools for measurement and analysis Haibi Peng 875552

Task 1: Process CSV data on command line

- 1. How you can peek on file if it is too large to fit into memory?
- a. Use head/tail -n number filename to view the first/last n lines in the file

```
pengh1@vdlubuntu097 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.out
% head -n 10 ./log_tcp_complete
#15#c_ip:1 c_port:2 c_pkts_all:3 c_rst_cnt:4 c_ack_cnt:5 c_ack_cnt_p:6 c_bytes_uniq:7 c_pkts_data:8 c_bytes_all:9 c_
pkts_retx:10 c_bytes_retx:11 c_pkts_ooo:12 c_syn_cnt:13 c_fin_cnt:14 s_ip:15 s_port:16 s_pkts_all:17 s_rst_cnt:18 s_
ack_cnt:19 s_ack_cnt_p:20 s_bytes_uniq:21 s_pkts_data:22 s_bytes_all:23 s_pkts_retx:24 s_bytes_retx:25 s_pkts_ooo:26
s_syn_cnt:27 s_fin_cnt:28 first:29 last:30 durat:31 c_first:32 s_first:33 c_last:34 s_last:35 c_first_ack:36 s_firs
t_ack:37 c_isint:38 s_isint:39 c_iscrypto:40 s_iscrypto:41 con_t:42 p2p_t:43 http_t:44 c_rtt_avg:45 c_rtt_min:46 c_r
tt_max:47 c_rtt_std:48 c_rtt_cnt:49 c_ttl_min:50 c_ttl_max:51 s_rtt_avg:52 s_rtt_min:53 s_rtt_max:54 s_rtt_std:55 s_
rtt_cnt:56 s_ttl_min:57 s_ttl_max:58 p2p_st:59 ed2k_data:60 ed2k_sig:61 ed2k_c2s:62 ed2k_c2c:63 ed2k_chat:64 c_f1323
opt:65 c_tm_opt:66 c_win_scl:67 c_sack_opt:68 c_sack_cnt:69 c_mss:70 c_mss_max:71 c_mss_min:72 c_win_max:73 c_win_m
in:74 c_win_0:75 c_cwin_max:76 c_cwin_min:77 c_cwin_in:78 c_pkts_rto:79 c_pkts_fs:80 c_pkts_reor:81 c_pkts_dup:82 c
_pkts_unk:83 c_pkts_fc:84 c_pkts_unrto:85 c_pkts_unfs:86 c_syn_retx:87 s_f1323_opt:88 s_tm_opt:89 s_win_scl:90 s_sack_kopt:91 s_sack_cnt:92 s_mss:93 s_mss_max:94 s_mss_min:95 s_win_max:96 s_win_min:97 s_win_0:98 s_cwin_max:99 s_cwin_min:100 s_cwin_in:101 s_pkts_rto:102 s_pkts_fs:103 s_pkts_reor:104 s_pkts_dup:105 s_pkts_unk:106 s_pkts_fc:107 s_pk
ts_unrto:108 s_pkts_unfs:109 s_syn_retx:110 http_req_cnt:111 http_res_cnt:112 http_res:113 c_pkts_push:114 s_pkts_pu
sh:115 c_tls_SNI:116 s_tls_SCN:117 c_npnalpn:118 s_npnalpn:119 c_tls_sesid:120 c_last_handshaker:121 s_last_handshake
et:122 c_appdataT:123 s_appdataT:124 c_appdataB:125 s_appdataB:126 fqdn:127 dns_rslv:128 req_tm:129 res_tm:130
163.33.158.67 64548 17 0 16 9 714 6 714 0 0 0 1 1 17.114.183.199 443 12 1 11 2 5597 7 5597 0 0 0 1 1 1491922800473.7
82959 1491922800609.770020 135.987000 33.231000 41.022000 1
```

b. Use less [OPTIONS] filename to peek on the file one page at a time

pengh1@vdiubuntu097 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.out % less ./log tcp complete

2. Averages of columns 3, 7, 10, 17, 21, 24 *awk*

```
'\{c[0]=3;c[1]=7;c[2]=10;c[3]=17;c[4]=21;c[5]=24;for(i=0;i<=5;++i)\{b[i]+=$c[i]\}\}END\{for(i=0;i<=5;++i)\{printf("Average of column %s: %s", c[i], b[i]/NR);printf("\n")\}\}'./log_tcp_complete
```

3. Percentage of records where c10/c7 exceeds a) 0.01, b) 0.10, c) 0.20 (c10/c7 means for each line value in column 10 is divided by value in column 7) $awk \quad FNR>1\&\&\$7>0\&\&\$10/\$7>0.01\{c+=\$3;n++\}END\{printf("Percentage of records where c10/c7 exceeds 0.01:%s",c/n)\}' log_tcp_complete$

```
pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.o
ut
   % awk 'FNR>1&&$7>0&&$10/$7>0.01{c+=$3;n++}END{printf("Percentage of records where c10/c7
exceeds 0.01:%s",c/n)}' log_tcp_complete
Percentage of records where c10/c7 exceeds 0.01:20.4996

pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.o
ut
   % awk 'FNR>1&&$7>0&&$10/$7>0.1{c+=$3;n++}END{printf("Percentage of records where c10/c7 e
xceeds 0.1:%s",c/n)}' log_tcp_complete
```

Percentage of records where c10/c7 exceeds 0.1:10.0888%

```
pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017 04 11 18 00.o
% awk 'FNR>1&&$7>0&&$10/$7>0.2{c+=$3;n++}END{printf("Percentage of records where c10/c7 e
xceeds 0.2:%s",c/n)}' log_tcp_complete
Percentage of records where c10/c7 exceeds 0.2:11.0624
```

```
4. Percentage of records where c24/c21 exceeds a) 0.01, b) 0.10, c) 0.20
     pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.o
     % awk 'FNR>1&&$21>0&&$24/$21>0.01{c+=$3;n++}END{printf("Percentage of records where c24/c
21 exceeds 0.01:%s",c/n)}' log_tcp_complete
Percentage of records where c24/c21 exceeds 0.01:17.5581%
     pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.o
     % awk 'FNR>1&&$21>0&&$24/$21>0.1{c+=$3;n++}END{printf("Percentage of records where c24/c2
1 exceeds 0.1:%s",c/n)}' log_tcp_complete
Percentage of records where c24/c21 exceeds 0.1:5.09839%
     pengh1@vdiubuntu016 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.o
     % awk 'FNR>1&&$21>0&&$24/$21>0.2{c+=$3;n++}END{printf("Percentage of records where c24/c2
1 exceeds 0.2:%s",c/n)}' log_tcp_complete
     Percentage of records where c24/c21 exceeds 0.2:4.81658%
```

5. Largest values in columns 3, 9, 17, 23, 31

awk

```
|c| = 3; c[1] = 9; c[2] = 17; c[3] = 23; c[4] = 31; for(i=0; i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4; ++i) | max[i] = 0; if($c[i] > i < =4;
max[i]{max[i]=$c[i]}}END{for(i=0;i<=4;++i){printf("Largest
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           value
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           of
colunm \%s : \%s'', c[i], max[i]); printf("\n")}\}' ./log_tcp_complete
```

```
pengnlgvotubuntu097 /work/courses/unix/T/ELEC/E7130/general/trace/tstat/2017_04_11_18_00.out
% awk '{c[0]=3;c[1]=9;c[2]=17;c[3]=23;c[4]=31;for(i=0;i<=4;++i){max[i]=0;if($c[i]=max[i]){max[i]=$c[i]}}}END{for(i=
;i<=4;++i){printf("largest value of colunm %s : %s", c[i], max[i]);printf("\n")}}' ./log_tcp_complete
.argest value of colunm 3 : 9
.argest value of colunm 9 : 213
.argest value of colunm 17 : 7
.argest value of colunm 23 : 4858
.argest value of colunm 31 : 947.208000</pre>
```

Task 2: Make a script to produce ping statistics

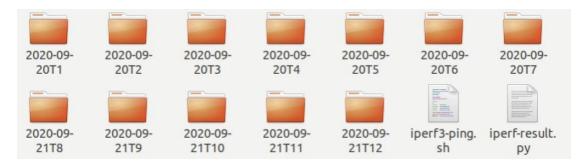
Solution:

```
Use crontab to control the running frequency of script: Crontab -e
SHELL=/bin/bash
* * * * * /bin/sh /u/88/pengh1/unix/Desktop/linux intro-code/ping/ping-kill.sh >> ~/log 2>&1
```

ping-kill.sh:

```
host=${1:-ok1.iperf.comnet-student.eu}
secs=${2:-20}
d=$(date -Isec | tr -d : | sed s/+.*//)
mkdir $d
ping -c $(($secs)) $host >> $d/ping-$(date +%s).txt &
iperf3 -c $host -t $secs -J >> $d/iperf3-$(date +%s).json
wait
```

The output data:



```
ping-1600608001.txt
   Open ▼
                Ð
PING iperf.netlab.hut.fi (195.148.124.36) 56(84) bytes of data.
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=1 ttl=60 time=0.658 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=2 ttl=60 time=1.55 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=3 ttl=60 time=1.38 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=4 ttl=60 time=1.51 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=5 ttl=60 time=1.55 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=6 ttl=60 time=1.38 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=7 ttl=60 time=1.41 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=8 ttl=60 time=1.57 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=9 ttl=60 time=1.43 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=10 ttl=60 time=1.41 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=11 ttl=60 time=1.46 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=12 ttl=60 time=1.56 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=13 ttl=60 time=1.53 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp seq=14 ttl=60 time=1.48 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=15 ttl=60 time=1.71 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=16 ttl=60 time=1.46 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=17 ttl=60 time=0.682 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=18 ttl=60 time=1.47 ms 64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=19 ttl=60 time=1.30 ms
64 bytes from iperf.netlab.hut.fi (195.148.124.36): icmp_seq=20 ttl=60 time=1.55 ms
--- iperf.netlab.hut.fi ping statistics ---
20 packets transmitted, 20 received, 0% packet loss, time 19045ms
rtt min/avg/max/mdev = 0.658/1.407/1.715/0.261 ms
```

1. Percentage of lost packets.

```
#percentage of lost packets
pofp=[]
for i in range(len(SENT)):
    pofp.append((np.sum(SENT[i])-np.sum(RECIEVED[i]))/np.sum(SENT[i]))
print('Percentage of lost packets for each hour(%):',pofp)
print('\n')

In [8]: runfile('/u/88/pengh1/unix/Desktop/linux_intro-code/ping/parse-re.py', wdir='/u/88/pengh1/unix/Desktop/linux_intro-code/ping')
Percentage of lost packets for each hour(%): [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

2. Average of successful round-trip-time (RTT) measurements.

3. Median of RTT. Lost packet is counted RTT as inf.

```
#median of RTT
median=[]
for i in range(len(RTT)):
    RTTS = sorted(RTT[i], key=float)
    #print(RTTS)
    #RTT[i].sort(cmp=my_cmp, reverse=True)
    #print(RTT[i])
    median.append(RTTS[59])
print('Median of RTT for each hour(ms):',median)
Median of RTT for each hour(ms): [1.3, 1.44, 2.52, 1.03, 0.796, 1.29, 1.43, 1.06, 1.42, 1.2, 1.36, 0.705]
```

Task 3: Measure network throughput performance

Solution:

875552%60=32

Use crontab to control the running frequency of script: Crontab -e

```
SHELL=/bin/bash

1 * * * /bin/bash /u/88/pengh1/unix/Desktop/linux_intro-code/ping/Task3/iperf3-send.sh >> ~/log1 2>&1

32 * * * * /bin/bash /u/88/pengh1/unix/Desktop/linux_intro-code/ping/Task3/iperf3-receive.sh >> ~/log2 2>&1
```

iperf3-send.sh:

```
function rand(){
    min=$1
    max=$(($2-$min+1))
    num=$(date +%s%N)
    echo $(($num%$max+$min))
}

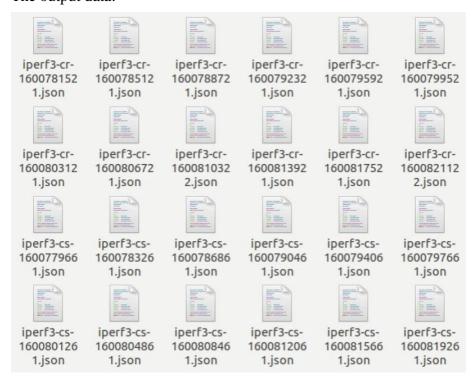
host=${1:-ok1.iperf.comnet-student.eu}
secs=${2:-10}
d=$(date -Isec | tr -d : | sed s/+.*//)
port=$(rand 5200 5210)
mkdir $d
iperf3 -s -p $port -1 -J >> $d/iperf3-ss-$(date +%s).json &
iperfid=$!
iperf3 -c $host -t $secs -p $port -J >> $d/iperf3-cs-$(date +%s).json kill -INT $iperfid
```

iperf3-receive.sh:

```
function rand(){
    min=$1
    max=$(($2-$min+1))
    num=$(date +%s%N)
    echo $(($num%$max+$min))
}

host=${1:-ok1.iperf.comnet-student.eu}
secs=${2:-10}
d=$(date -Isec | tr -d : | sed s/+.*//)
port=$(rand 5200 5210)
mkdir $d
iperf3 -s -p $port -1 -J >> $d/iperf3-sr-$(date +%s).json &
iperfid=$!
iperf3 -c $host -t $secs -p $port -R -J >> $d/iperf3-cr-$(date +%s).json
kill -INT $iperfid
```

The output data:



Sending record:

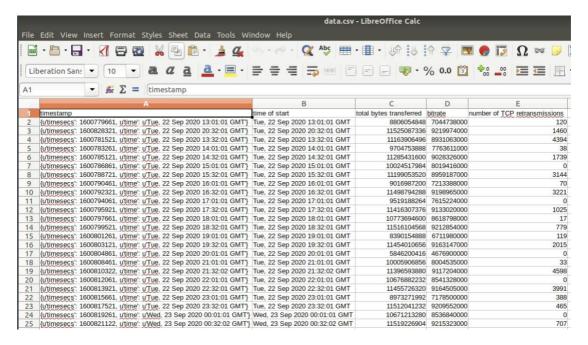
```
"start": {
    "connected":
                                          [{
"socket":
                                                               4,
"130.233.71.166",
                                          "local_host": "130.233.71.166",
"local_port": 37628,
"remote_host": "195.148.124.36",
"remote_port": 5209
                     }],
"version":
"timesecs":
                                                    1600779661
                    },
"connecting_to": {
    "host": "ok1.iperf.comnet-student.eu",
    "port": 5209
                     },
"cookie":
                     "cookie": "vdiubuntu166.1600779661.518077.48b49",
"tcp_mss_default": 1448,
                     "test_start": {
    "protocol": "Ti
    "num_streams": 1,
    "blksize": 13
                                                     "TCP",
                                                    131072.
                                "omit": 0,
"duration":
                                                     10,
                                "bytes":
"blocks":
                                "reverse":
           },
```

Receiving record:

```
"start": {
    "connected":
                                             [{
"socket":
                                                                   4,
"130.233.71.166",
                                             "local_host":
                                             "local_nost": "130.233.71.166",
"local_port": 53172,
"remote_host": "195.148.124.36",
"remote_port": 5210
                                 }],
                      "version": "iperf 3.1.3",
"system_info": "Linux vdiubuntu166 4.15.0-99-generic #100-Ubuntu SMP
Wed Apr 22 20:32:56 UTC 2020 x86_64",
                      "timestamp": {
    "time": "Tue, 22 Sep 2020 13:32:01 GMT",
                                 "timesecs":
                                                       1600781521
                       'connecting_to": {
    "host": "ok1.iperf.comnet-student.eu",
    "port": 5210
                      },
"cookie":
                                             "vdiubuntu166.1600781521.873819.78b46",
                      "tcp_mss_default":
"test_start": {
                                  art": {
"protocol":
                                                        "TCP",
                                  "num_streams":
                                  "blksize":
                                                        131072,
                                 "omit": 0,
"duration":
                                                        10,
                                 "bytes":
"blocks":
                                  "reverse":
                      }
           },
```

To produce .csv file:

```
dataList=[]
for f in datalist:
 with open(f) as jf:
    trv:
      ip=json.load(jf)
      print('\t'.join(map(str, [ip['start']['timestamp'],
                                 ip['start']['timestamp']['time'],
                                 ip['end']['sum_sent']['bytes'],
ip['end']['sum_sent']['bits_per_second'],
                                 ip['end']['sum_sent']['retransmits']])))
      ip['end']['sum_sent']['bytes'],
                                 ip['end']['sum_sent']['bits_per_second'],
ip['end']['sum_sent']['retransmits']]))
    except json.decoder.JSONDecodeError:
                               # invalid json file, just ignore
      pass
                                 # tried to read non existent value
    except KevError:
      if 'start' in ip:
        print('\t'.join(map(str, [ip['start']['timestamp']['timesecs'], 0, 0])))
#print(len(dataList))
with open("data.csv", "w") as csvfile:
    writer = csv.writer(csvfile)
   writer.writerow(["timestamp", "time of start", "total bytes transferred",
                      bitrate", "number of TCP retransmissions"])
    writer.writerows(dataList)
```



1. Average bitrate for all measurement.

```
#Average bitrate for all measurement
sum=0
for i in dataList:
    sum+=float(i[3])
avgbr=sum/len(dataList)
print('Average bitrate for all measurement:', avgbr)
```

```
In [52]: runfile('/u/88/pengh1/unix/Desktop/linux_intro-code/Task3/iperf-
result.py', wdir='/u/88/pengh1/unix/Desktop/linux_intro-code/Task3')
Average bitrate for all measurement: 8311599083.33
```

2. Minimum, median and maximum bitrate.

```
#Minimum, median and maximum bitrate.
 br=[]
  for i in dataList:
                          br.append(float(i[3]))
  BR=sorted(br, key=float)
  #print(BR
Bitrate={'Minimum bitrate':BR[0], 'Median bitrate':BR[11], 'Maximum bitrate':BR[23]} #List = [('Minimum bitrate', BR[0]), ('Median bitrate', BR[11]), ('Maximum bitrate', BR[0]), ('Median bitrate', BR[11]), ('Maximum bitrate', BR[11]), 'Maximum bitr
  #Bitrate = dict(List)
 print(Bitrate)
    {'Minimum bitrate': 4676900000.0, 'Maximum bitrate': 9219974000.0, 'Median
    bitrate': 8618798000.0}
```

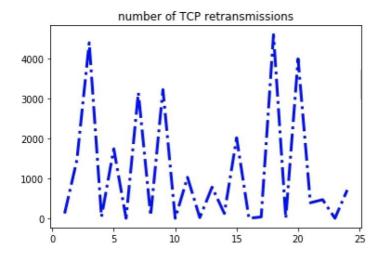
3. Graph comparing bitrate and number of TCP retransmissions.

```
#Graph comparing bitrate and number of TCP retransmissions
rt=[]
for i in dataList:
    rt.append(float(i[4]))
x1 = list(range(1,24+1))
x2 = list(range(1,24+1))
plt.plot(x1,br,color='red',linewidth=2.0,linestyle='--')
plt.title('bitrate')
plt.figure()
plt.plot(x2,rt,color='blue',linewidth=3.0,linestyle='-.')
plt.title('number of TCP retransmissions')
                           bitrate
0.9
0.8
0.7
0.6
0.5
                                             20
```

10

15

25



4. If you used your own computer to run iperf, describe network connectivity. Can you make any conclusions of stability based on data?

The network connectivity remained stable in most of the time but sometimes had fluctuation since the bitrate was comparatively lower and the number of TCP retransmissions were higher, as is seen in the graph.