Internet Measurements and Forensics

Exercise 2

Task 2.1: Ping Experiment (cont'd)

In this problem, we will be using your own data from the ping experiment in the previous exercise.

a) Create a scatterplot in which you plot the RTT against the distance into a plot showing distance on the x axis and the corresponding time on the y axis. Add a line to the plot that shows the speed of light as a function of the distance.

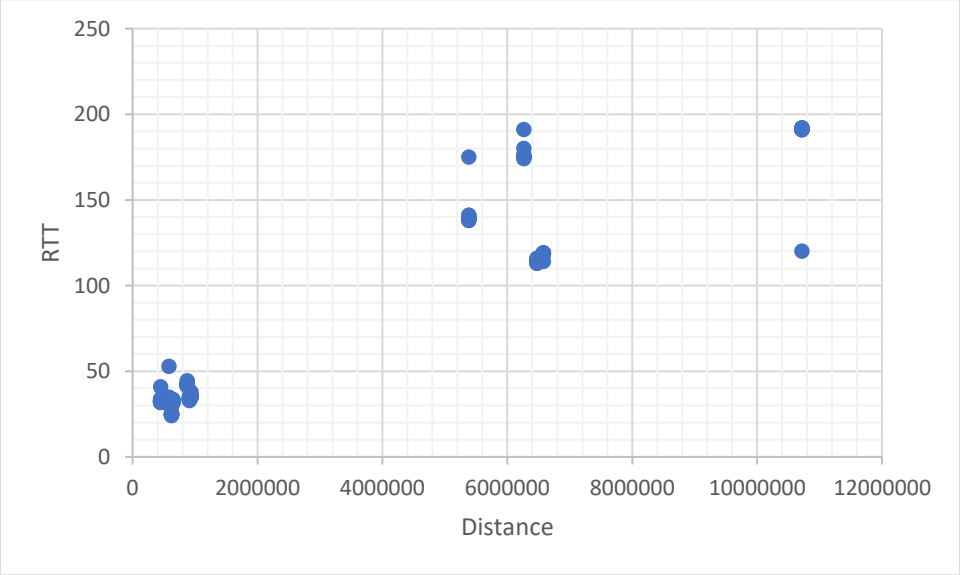
Solution:

school website	Distance	RTT
rwth-aachen.de	**583200**	31
rwth-aachen.de	**583200**	32.2
rwth-aachen.de	**583200**	31.3
rwth-aachen.de	**583200**	34.7
rwth-aachen.de	**583200**	34.7
rwth-aachen.de	**583200**	52.7
rwth-aachen.de	**583200**	31.8
rwth-aachen.de	**583200**	31.7
rwth-aachen.de	**583200**	30.6
rwth-aachen.de	**583200**	33.1
lmu.de	**447600**	32.6
lmu.de	**447600**	31.7
lmu.de	**447600**	40.8
lmu.de	**447600**	32.9
lmu.de	**447600**	32.7
lmu.de	**447600**	32.3
lmu.de	**447600**	32.1
lmu.de	**447600**	31.8
lmu.de	**447600**	32.4
lmu.de	**447600**	33.5
tum.de	**447600**	32.4

tum.de	**447600**	31.9
tum.de	**447600**	31.9
tum.de	**447600**	33.2
tum.de	**447600**	31.8
tum.de	**447600**	32.7
tum.de	**447600**	33.8
tum.de	**447600**	33.1
tum.de	**447600**	31.8
tum.de	**447600**	32.8
wwwen.uni.lu	**625100**	24.7
wwwen.uni.lu	**625100**	29.3
wwwen.uni.lu	**625100**	24.1
wwwen.uni.lu	**625100**	24.9
wwwen.uni.lu	**625100**	24.7
wwwen.uni.lu	**625100**	24.8
wwwen.uni.lu	**625100**	24.4
wwwen.uni.lu	**625100**	25.2
wwwen.uni.lu	**625100**	24.7
wwwen.uni.lu	**625100**	25
aup.edu	**909500**	33.6
aup.edu	**909500**	35.5
aup.edu	**909500**	33.3
aup.edu	**909500**	34.2
aup.edu	**909500**	33.4
aup.edu	**909500**	33.8
aup.edu	**909500**	32.8
aup.edu	**909500**	33.1
aup.edu	**909500**	34.7
aup.edu	**909500**	34.5
ethz.ch	**642500**	33.1
ethz.ch	**642500**	31.8
ethz.ch	**642500**	31.8
ethz.ch	**642500**	31.3
ethz.ch	**642500**	32.9
ethz.ch	**642500**	33.5
ethz.ch	**642500**	32.9
ethz.ch	**642500**	32.8
ethz.ch	**642500**	31.5
ethz.ch	**642500**	32.1
uu.se	**872800**	41.6
uu.se	**872800**	41.6
uu.se	**872800**	41.8
uu.se	**872800**	41.2

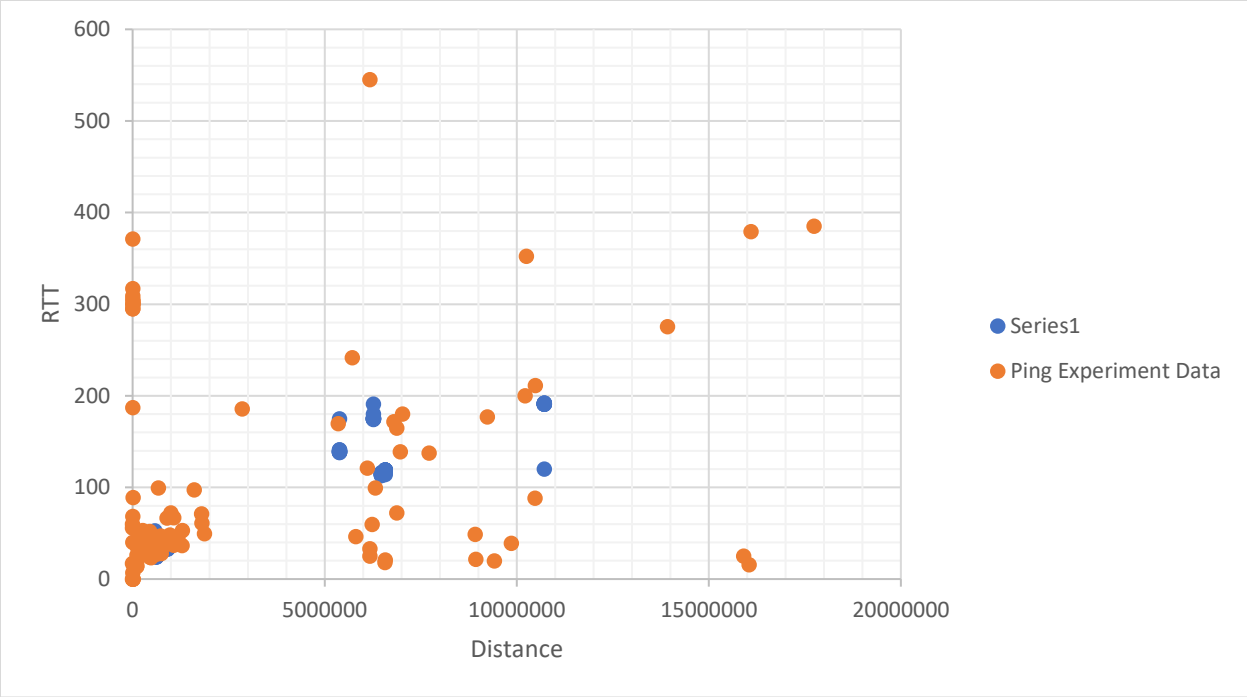
uu.se	**872800**	41.5
uu.se	**872800**	43
uu.se	**872800**	42
uu.se	**872800**	42.3
uu.se	**872800**	44.4
uu.se	**872800**	43.3
uio.no	**933100**	35.9
uio.no	**933100**	34.8
uio.no	**933100**	35.3
uio.no	**933100**	37.8
uio.no	**933100**	35.8
uio.no	**933100**	35.1
uio.no	**933100**	35.7
uio.no	**933100**	35.6
uio.no	**933100**	35.3
uio.no	**933100**	34.6
ucc.edu.gh	**5384800**	139
ucc.edu.gh	**5384800**	139
ucc.edu.gh	**5384800**	141
ucc.edu.gh	**5384800**	138
ucc.edu.gh	**5384800**	140
ucc.edu.gh	**5384800**	139
ucc.edu.gh	**5384800**	138
ucc.edu.gh	**5384800**	141
ucc.edu.gh	**5384800**	139
ucc.edu.gh	**5384800**	175
uonbi.ac.ke	**6266100**	175
uonbi.ac.ke	**6266100**	175
uonbi.ac.ke	**6266100**	175
uonbi.ac.ke	**6266100**	176
uonbi.ac.ke	**6266100**	175
uonbi.ac.ke	**6266100**	180
uonbi.ac.ke	**6266100**	174
uonbi.ac.ke	**6266100**	175
uonbi.ac.ke	**6266100**	176
uonbi.ac.ke	**6266100**	191
ugm.ac.id	**10715300**	191
ugm.ac.id	**10715300**	192
ugm.ac.id	**10715300**	192
ugm.ac.id	**10715300**	192
ugm.ac.id	**10715300**	191
ugm.ac.id	**10715300**	191
ugm.ac.id	**10715300**	191

ugm.ac.id	**10715300**	191
ugm.ac.id	**10715300**	191
ugm.ac.id	**10715300**	120
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	118
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	119
utoronto.ca	**6575000**	114
columbia.edu	**6475000**	115
columbia.edu	**6475000**	115
columbia.edu	**6475000**	116
columbia.edu	**6475000**	114
columbia.edu	**6475000**	114
columbia.edu	**6475000**	114
columbia.edu	**6475000**	114
columbia.edu	**6475000**	113
columbia.edu	**6475000**	115
columbia.edu	**6475000**	113



b) Download other ping results “Ping Experiment Data” from Moodle. Add the data downloaded in the previous step to your scatter plot, but in a different color.

Solution:



c) How can you inspect the data to detect data quality issues? Which data quality issues do you need to look for?

Solution:

1. **Packet Loss – You can check the relationship between the packet sent and the packet received. You can inspect this from looking at packet loss percentage in the response.**
2. **Time – The time it took for the sample packets to be sent and returned. You can detect this from the time for packets in the return dataset.**

d) Can you find targets that have a smaller RTT but a higher distance than other targets? How can this happen?

Solution:

There are several factors that can influence this such as:

Factor	Reason
Application Hosting Location	**With the onset of the internet, its is possible to host an application outside the location of the of the organization hence, the ping might contact a hosting server closer to the requesting cost whiles the organization location is far.**
Internet Connection speed	**Internet connection speed can influence how a packet is quickly forwarded and received over the transmission medium; this can have a huge influence on the RTT.**

Task 2.2: Passive Measurements: Trace Analysis

a) How many TCP connections are contained in the trace, at least in part? How do you get this information?

Solution:

I observed a total of **2429 connections** in the trace, representing **94.6%** of the captured data. This information was retrieved by filtering with "**tcp**" and fetch the data from status toolbar of Wireshark.

b) Look at the TCP connection in the trace which starts first. Use display filters to filter for only this connection. Which display filter do you use?

Hint: Right click on a packet and follow the TCP stream.

Fill in the following table (only a single row!) with the IP addresses of the hosts that are communicating. and with the start and end time of the connection.

Solution

Filter: frame.time_relative == 0.000194

Connection Identifier	**IP or initiating host**	**IP of peer**	**conn. start time**	**Conn. end time**
236254517	**192.168.100.200**	**192.168.100.100**	**2009-12-02 22:07:59**	**2009-12-02 22:12:55**

c) Fill Table 2 below for all TCP connections in the trace, one row per connection (including the one in the previous table). Sort the connections in increasing order of Wireshark TCP stream identifier.

How do you get the information? Explain which Wireshark analyzers or display filters you use.

Connection Identifier	IP of initiating host	IP of peer	Conn. start time	Conn. end time

d) How do you detect the start / end of a TCP connection? Does your approach work in every situation?

Solution:

The connection is usually started with the 3-way handshake where the client first sends a SYN to the server followed by an ACK from the server to the client. The client also sends an ACK to the server then the connection is started, and communication can begin. After the communication is done, the client will send a FIN flag to the server followed by an ACK from the server to the client. The server then sends a FIN again followed by an ACK from the client. The communication is then terminated afterwards.

e) How many UDP flows are in the trace? Explain how you find this information.

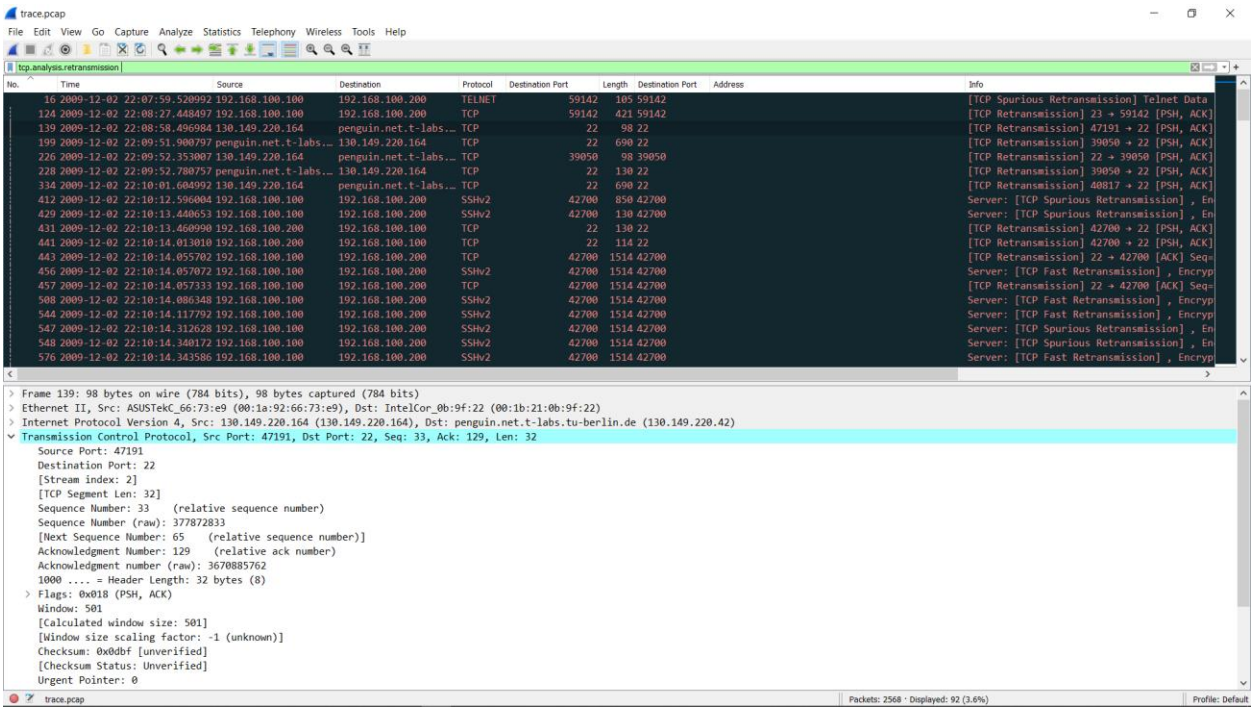
Solution:

There are 137 udp packet trace in the captured data, it can be located by filtering with 'udp' and fetch the data from status toolbar of Wireshark.

f) Give an example of a TCP connection exhibiting a packet loss, specified by its Wireshark TCP stream identifier. How do you find this information? E.g., what display filter can help you find packet losses in TCP?

Solution:

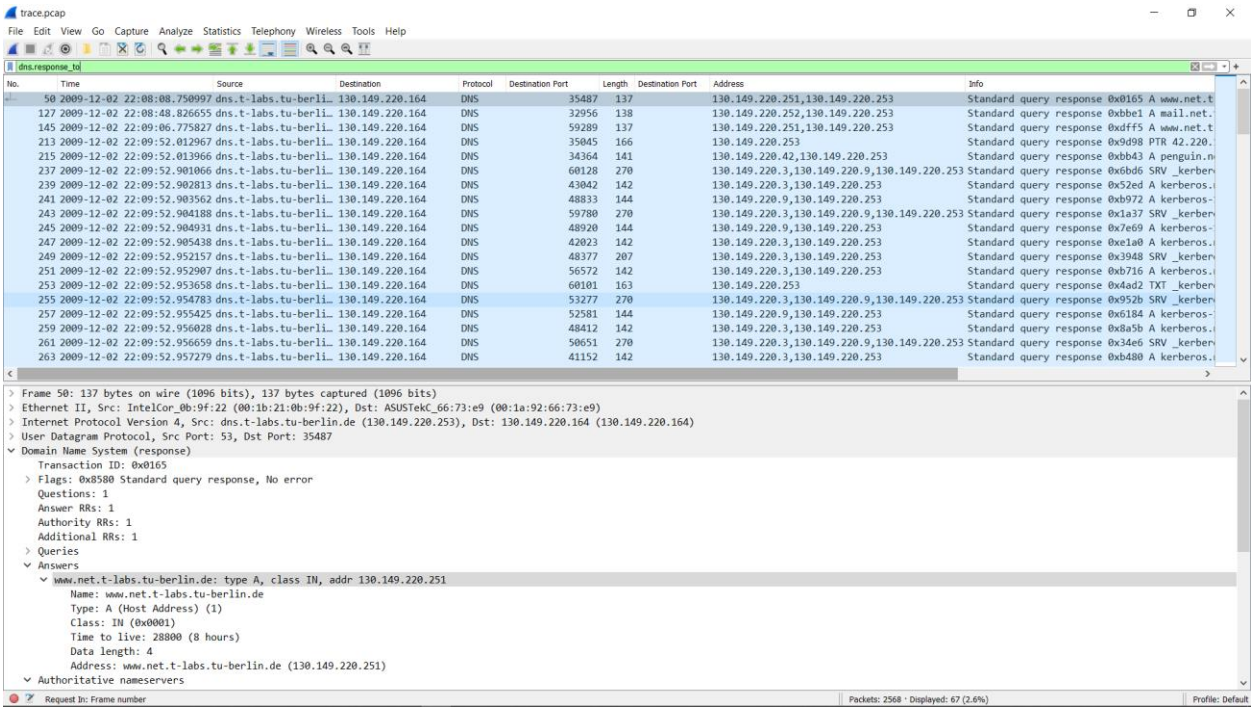
You can use this filter: **tcp.analysis.retransmission** to find this information.



g) Obtain the host name of a single host, as queried via DNS in the trace, and specify its IP address. Explain your approach.

Solution:

Using the ‘dns’ filter, I resolve the IP through view to look for the response from the DNS server and checked the response in the packet display to check for the resolved IP.



Name: www.net.t-labs.tu-berlin.de
Resolved IP: 130.149.220.251

h) Use automatic analyzers of Wireshark to provide the host names of all the hosts (as resolved using DNS, including the host from the previous question). Present your results in Table 3 below.
What analyzer do you use?
Solution:

Host Ip	DNS name
130.149.220.253	**dns.t-labs.tu-berlin.de**
130.149.220.2	**intserv.net.t-labs.tu-berlin.de**
130.149.220.9	**kerberos-1.net.t-labs.tu-berlin.de**
130.149.220.3	**kerberos.net.t-labs.tu-berlin.de**
130.149.220.252	**mail.net.t-labs.tu-berlin.de**
130.149.220.42	**penguin.net.t-labs.tu-berlin.de**
130.149.220.251	**www.net.t-labs.tu-berlin.de**

I used the Resolved the Statistics: resolved Addresses Analyzer and changed the entry to host.

i) Look again at the TCP connections in increasing order of Wireshark identifier.
For each TCP connection, answer all of the following questions in one paragraph per connection.
If you cannot find this information, justify why it is not possible.

a) What protocol is being used on the application layer?
Solution:
Protocols being used in the application layer are:

- Telnet
- SSH

- SMTP
- HTTP

b) What is the user doing? Is the client requesting anything, sending anything, ...?
Solution:

- Telnet – User uses a command line interface to communicate with a server and also manage the network device.
- SSH – Protocol used to provide a secure communication tunnel between user computer and any network enabled responsive device over a secure communication path.
- SMTP – This is a protocol that the user uses to send electronic mail, hence user is sending an email here.
- HTTP – Protocol used to send and receive data / communicate between web client and servers usually by use of a browser.

c) Which information is disclosed (passwords, content, etc.)?
Solution:

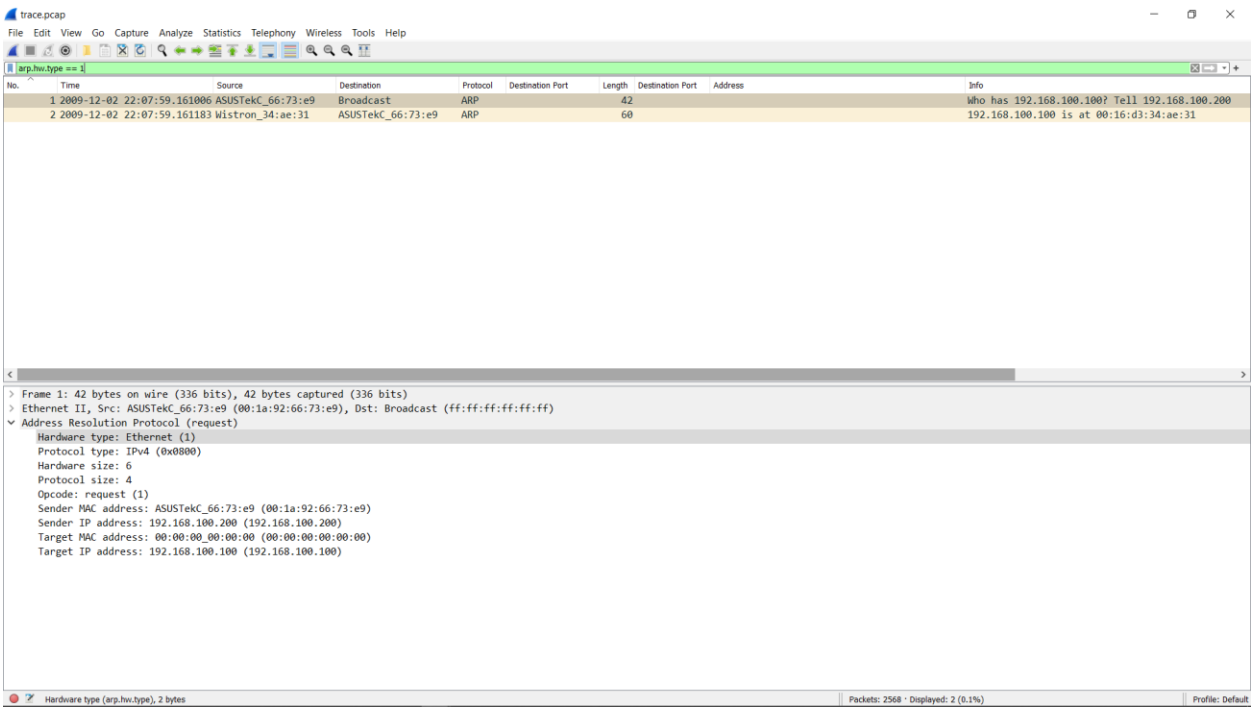
1. **Ethernet and IP resolutions**
2. **Login and logout time**
3. **Password**
4. **Username**
5. **Operating System**
6. **Source and destination of messages**
7. **Email timestamp**

d) Which host is the server, which is the client?
Solution:

Although each host can play the role of a client and server simultaneously, usually the source column is the host and destination column are the server.

Host	Server
130.149.220.42	130.149.220.14
130.149.220.164	130.149.220.23
130.149.220.164	130.149.220.21
130.149.220.164	130.149.220.22
192.168.100.100	192.168.100.20
192.168.100.100	224.0.0.251

j) What can you infer about the network topology by considering layer 2 information: The traffic was captured in one LAN, which other hosts are located in the same LAN? Hint: filter by MAC addresses and analyse their usage. Remember how MAC addresses are used for addressing and when they are rewritten.
Which hosts have multiple IP addresses?
Solution:



From the filter, it was observed that there are two hosts in one LAN:

- **ASUSTekC**
- **Wistron**

ASUSekC sent a ARP broadcast and Wistron responded with the MAC address.