1. No. of TCP Flows in the PCap file

# of TCP flows: 3

During the second step in the TCP handshake, the sender receives a SYN-ACK packet from the receiver. This is essentially used to identify connections in a packet trace.
On traversing the packet list, a packet with SYN-ACK flag set implies a connection is being setup. You can register the end-point addresses of this connection.
Now traverse the packet list again and for every packet find the connection addresses pair to which the packet source and destination addresses match. The packet belongs to the connection to which its addresses are matched.

2.a Transactions after TCP connection is set up

We see for the first two transactions the window size is some garbage value and it is set to 3 once the setup is complete.

Transaction 1 is SYN
Transaction 2 is SYN-ACK
Transaction 3 and 4 are ACKs with same sequence and ack numbers. Here the sender and receiver are acknowledging the decided seq and ack values at the sender and receiver respectively.

CONNECTION 1

Sr.No	SEQ #	ACK #	WIN-SIZE	SYN	ACK
1	705669102	0	42340	1	0
2	1921750143	705669103	43440	1	1
3	705669103	1921750144	3	0	1
4	705669103	1921750144	3	0	1
5	705669127	1921750144	3	0	1

CONNECTION 2

Sr.No	SEQ #	ACK #	WIN-SIZE	SYN	ACK
1	3636173851	0	42340	1	0
2	2335809727	3636173852	43440	1	1
3	3636173852	2335809728	3	0	1
4	3636173852	2335809728	3	0	1
5	3636173876	2335809728	3	0	1

CONNECTION 3

Sr.No	SEQ #	ACK #	WIN-SIZE	SYN	ACK
1	2558634629	0	42340	1	0
2	3429921722	2558634630	43440	1	1
3	2558634630	3429921723	3	0	1
4	2558634630	3429921723	3	0	1
5	2558634654	3429921723	3	0	1

2.b, 2.c, 2.d Compute empirical and theoretical throughput, loss rate and average RTT for each TCP flow

The empirical throughput disregards retransmissions and calculates throughput by merely dividing the total payload size by the time taken. The theoretical throughput on the other hand considers retransmissions. Hence the values differ.

Empirical throughput: The empirical throughput for a connection can be found by simply summing up the payload sizes for all the packets and dividing it by the total number of packets sent.

Loss rate: Number of lost packets is equivalent to the number of retransmissions in the tcp flow. Loss rate can then be defined as a ratio of number of lost packets to the total number of packets sent.

RTT estimation: According to Karne's algorithm, RTT estimation is not done for retransmitted packets. Thus if the packets are indexed by seq number and ack number, then both the indexings should have packets with unique ack number and seq number.
For every unique ack number, we find a unique seq number which is one less than the ack number.

TCP FLOW ANALYSIS

CONN #	EMP THROUGHPUT (MBPS)	THROUGHPUT (MBPS)	LOSS RATE	AVG RTT (MS)
1	4.8529	99.9341	0.0004	0.9126
2	1.1887	22.4394	0.008	0.9088
3	1.3692	70.1873	0.0008	0.9188

TCP segment contents

```
TCP SEGMENT
src-ip: 130.245.145.12
dest-ip: 128.208.2.198
src-port: 43498
dest-port: 80
sequence-num: 705669102
ack: 0
data-offset: 10
reserved: 0
flags:
- ns: 0
- cwr: 0
- ece: 0
- urg: 0
- ack: 0
- psh: 0
- rst: 0
- syn: 1
- fin: 0
window-size: 42340 bytes
checksum: 63936
urgent-ptr: 0
payload-size: 20 bytes
timestamp: 1487361393.534537
base64-encoded-payload: AgQFtAEBCAoObomWAAAAAAEDAw4=
```