PART A

(1) Number of TCP flows initiated from the sender: 3

Here, I have checked for the syn and ack for each packet. If both the values are 1, the TCP flow is incremented by 1. This means that the syn-ack handshaking is done.

(2) (a) For each TCP flow:

Explanation:

- Here, each connection is segregated and packets are classified for each connection based on src and dest ports.
- Acknowledgement number tells the other side of the connection about the next packet it is expecting. Therefore, I have matched the sequence number of the next packet with ack_no.
- For each flow, I'm checking the sequence number and acknowledgement number of the sender's first transaction.
- Then, the acknowledgment number of the sender is matched with the next sequence number at the receiver.
- Last, the ack_no at receiver is matched with the seq_no at sender for the second transaction.
- This process is followed for each of the 3 flows and below are the results printed.

Printing for Flow number: 1

FIRST TRANSACTION:

At sender:

seq_no:705669103, ack_no:1921750144, window_size:3

At receiver:

seq_no:1921750144, ack_no:705669127, window_size:3

SECOND TRANSACTION:

At sender:

seq_no:705669127, ack_no:1921750144, window_size:3

At receiver:

seq_no:1921750144, ack_no:705670575, window_size:3

Printing for Flow number: 2

FIRST TRANSACTION:

At sender:

seq_no:3636173852, ack_no:2335809728, window_size:3

At receiver:

seq_no:2335809728, ack_no:3636173876, window_size:3

SECOND TRANSACTION:

At sender:

seq_no:3636173876, ack_no:2335809728, window_size:3

At receiver:

seq_no:2335809728, ack_no:3636175324, window_size:3

Printing for Flow number: 3

FIRST TRANSACTION:

At sender:

seq_no:2558634630, ack_no:3429921723, window_size:3

At receiver:

seq_no:3429921723, ack_no:2558634654, window_size:3

SECOND TRANSACTION:

At sender:

seq_no:2558634654, ack_no:3429921723, window_size:3

At receiver:

seq_no:3429921723, ack_no:2558636102, window_size:3

b) Explanation for Throughput calculation:

- Here, I have taken the packet_size for each packet and added them together to calculate the overall payload.
- Additionally, I have computed the time taken through the time_stamp field for each packet. Elapsed time is the difference between time_stamps of the last and first packet.
- Finally, throughput is calculated as payload_sum divided by the elapsed time.
- Please note, I have converted the values to Mbps.

Code Results:

Throughput for flow 1: 41.099145149333516 Mbps
Throughput for flow 2: 10.095422506647271 Mbps
Throughput for flow 3: 11.613253469151397 Mbps

c) Explanation for Loss Rate for each flow:

- Here, I have fetched the packets with the same sequence number with the help of a dictionary.
- If the sequence number is assigned to multiple packets, then it means that the sender has to resend the packet due to an earlier loss of packet and the total number of these retransmitted packets is the value of the total number of packets not received.
- Finally, loss rate is calculated by dividing the total lost packets with the total number for packets.

Code Results:

Loss rate for flow 1: 0.0005733123118818977
Loss rate for flow 2: 0.013440860215053764
Loss rate for flow 3: 0.0013717421124828531

(d) Explanation for the average RTT calculation:

- Here, I have taken the sequence number and time_stamp of all packets in a dictionary as key, value pairs.
- I have calculated the difference between the same sequence number and the same ack_number which will form the denominator of the RTT calculation.
- Finally, RTT is estimated by dividing the total time taken by the total number of transactions.

Code Results:

Estimated RTT for flow 1: 0.0731438426165122
Estimated RTT for flow 2: 0.16715163803478528
Estimated RTT for flow 3: 0.07205969513508312

- **Comparison with theoretical throughput:**
  Theoretical throughput for flow 1: 7.789599920396022 Mbps
  Theoretical throughput for flow 2: 0.7039865955140537 Mbps
  Theoretical throughput for flow 3: 5.111635795994693 Mbps

  Theoretical throughput Formula used:
  (sqroot(3/2) * MSS) / (RTT * sqroot(loss_rate))

  where , MSS (maximum segment size) = 1460
  , RTT = Calculated in above part
  , loss_rate = Calculated in above part

  Here, the empirical throughput is greater than the theoretical value as this formula was based on configurations that were present at that time. Now, we have higher bandwidths available due to which the empirical throughput values have exceeded as compared to theoretical.