Part 1 (Traceroute and Ping)

(a)

Host	Average RTT (ms)	RTT Variance (ms)	Average Hops	Hop Variance
Sydney	45.07143	2.587302	16.14286	1.142857
Bangkok	150.2857	3.322751	11	0
Riverside	164.1429	0.126984	16	0
Austin	211	0.148148	17	0
Boston	220.6071	1.136243	19	0
Kyoto	254.2143	1.21164	16.42857	0.952381
Guna	331.25	26.71296	10	0
Lausanne	334.2857	28.50794	20.85714	0.47619
Novosibirsk	339.1071	1093.655	24	14
Berlin	341.4286	16.40212	20	0
Ilorin	348.0714	17.99471	20	0
Tel Aviv	403.4286	266.7725	20	0
Linkoping	409.9286	144.6614	19.71429	0.238095

Average RTT/hops are higher for some cities compared to other cities due to:

- Greater physical distance as well as greater number of nodes (routers) between the source and the destination.
- Greater number of other requests being handled by intermediate nodes and the remote server.
- Slower intermediate nodes and the remote server function.
- Greater presence of interference in the circuit.

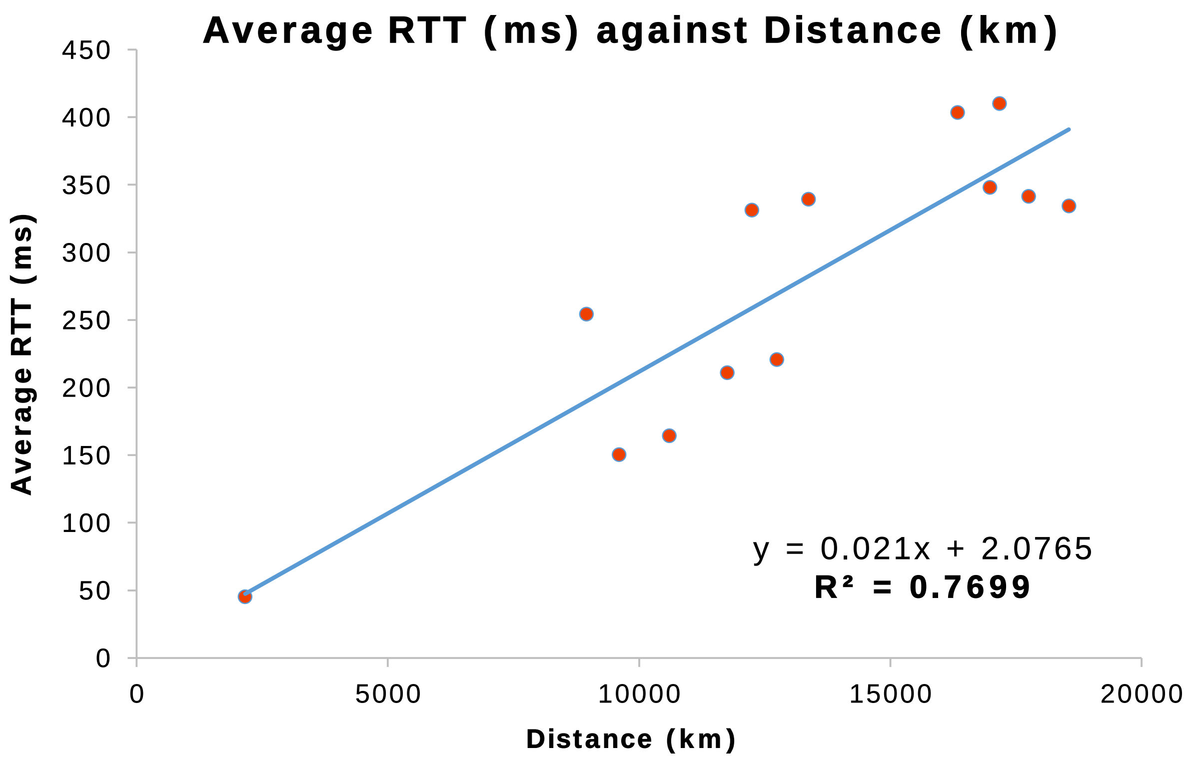
There are variances in RTT values and hops because each ping and hop measurement cannot be the same. The greater the range between the maximum value and the minimum value of ping/hop, the greater the variance.

Trace route shows the RTT to each hop and when there is a huge difference between one RTT to the next, that is when it is going through an undersea link. I could observe some huge jumps between my RTT values hence there are some major undersea links in the traceroutes.

(b)

Host	R/T	D (km)
Sydney	3.1302502	2159.8
Kyoto	4.2604318	8950.3
Bangkok	2.347163	9604.3
Riverside	2.322601	10600.8
Austin	2.693044	11752.5
Guna	4.0585079	12242.8
Boston	2.5974153	12740
Novosibirsk	3.803924	13372
Tel Aviv	3.703694	16338.9
Ilorin	3.0744557	16982.1
Linkoping	3.5809128	17171.4
Berlin	2.8850843	17751.4
Lausanne	2.7024343	18554.7

R/T represents the ratio between the actual time and the theoretical time. In other word, the ratio of round trip between the average RTT (ms) and the light. R/T is always greater than 1 because the speed that the packets travel across many routers (which reduce speed) must be slower than the speed of light traveling in a straight line.



According to the graph above, we can observe a positive linear line going across the graph with the R^2 value of 0.7699. This demonstrates that the greater the distance (km), the greater the average RTT (ms). Therefore there is a positive correlation between the distance and the average RTT.

Part 2 (IP Geolocation)

The host is located in the United Kingdom.

With the IP address of 212.162.63.104, Europe had an average ping of 18ms (shown in the screenshot below).

Location		Loss	Min	Avg	Max	Deviation	IP
Europe	order by country, state, city						
Roubaix, France NEW	traceroute dns	0%	13.4	13.5	13.6	0.16	212.162.63.104 (United Kingdom)
Frankfurt, Germany	traceroute dns	0%	8.7	8.7	8.8	0.11	212.162.63.104 (United Kingdom)
Fareham, GB	traceroute dns	0%	14.5	14.7	14.7	0.09	212.162.63.104 (United Kingdom)
London, GB	traceroute dns	0%	12	12.2	12.5	0.25	212.162.63.104 (United Kingdom)
Milan, Italy	traceroute dns	0%	12.9	13	13.2	0.19	212.162.63.104 (United Kingdom)
Amsterdam, Netherlands	-						
Lausanne, Switzerland	traceroute dns	0%	20.8	20.9	21	0.16	212.162.63.104 (United Kingdom)
Istanbul, Turkey NEW	traceroute dns	0%	41.2	44.4	46.1	2.29	212.162.63.104 (United Kingdom)
Kharkiv, Ukraine	traceroute dns	0%	50.4	50.5	50.6	0.11	212.162.63.104 (United Kingdom)
Europe Ping Averages	**18ms**						

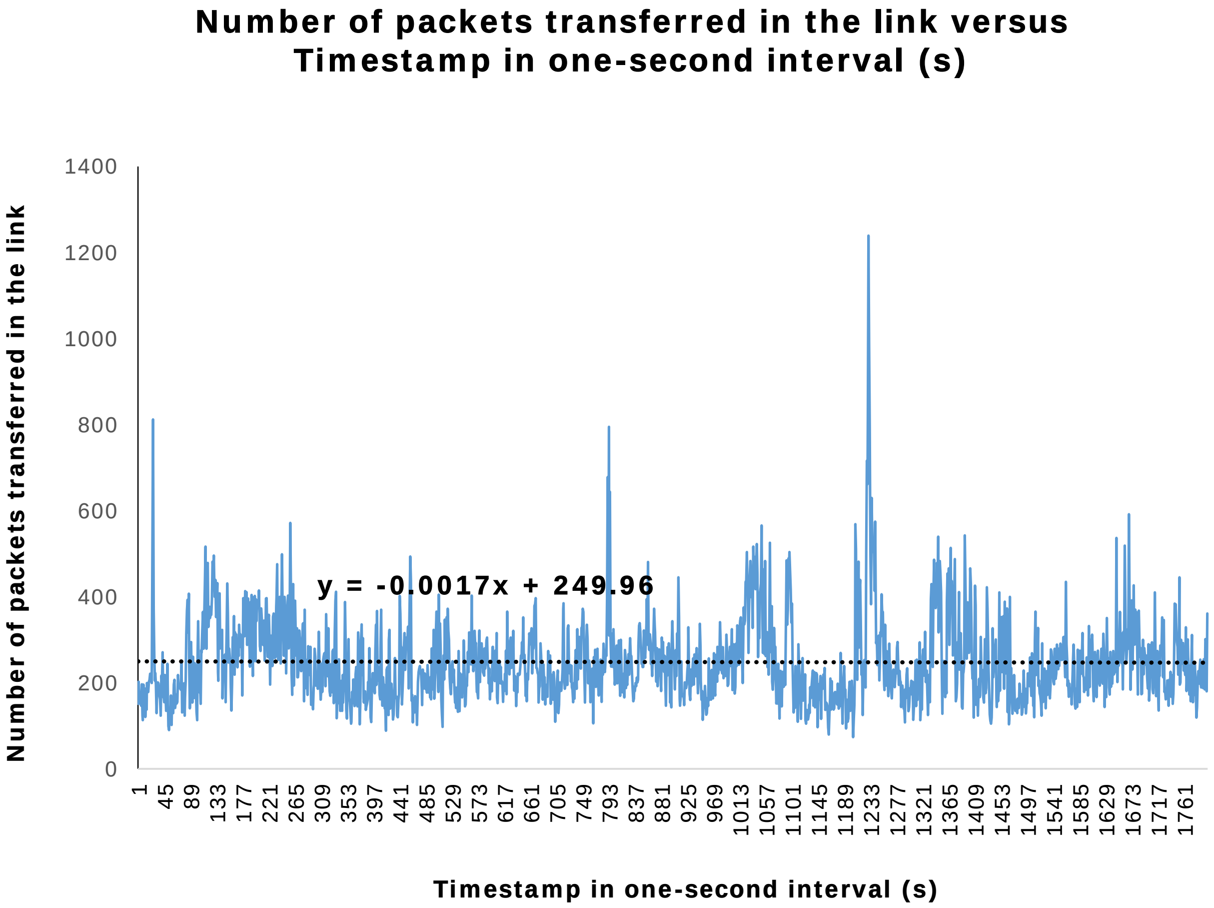
Next to the IP address is the location of the host, the United Kingdom.

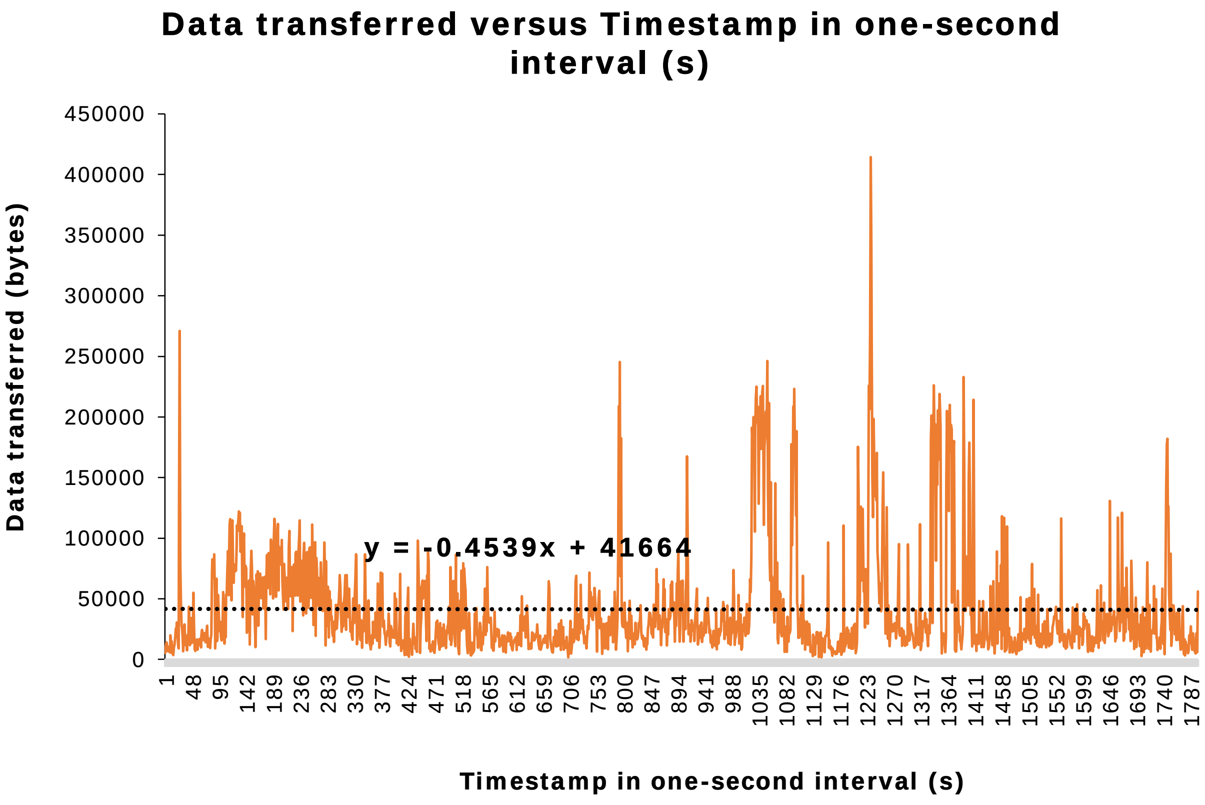
Part 3 (Internet Traffic Characterisation)

(a)

Number of packets transferred	447148
Average packet arrival rate	248.4158727
Average data rate (bits/second)	330045.2836
Trace duration (seconds)	1799.997702
Number of distinct sender hosts	777
Number of distinct receiver hosts	827
Number of distinct source ports	1569
Number of distinct destination ports	1518

(b)


Number of packets transferred in the link versus Timestamp in one-second interval (s)
Number of packets transferred in the link
y = -0.0017x + 249.96
0
200
400
600
800
1000
1200
1400
1
45
89
133
177
221
265
309
353
397
441
485
529
573
617
661
705
749
793
837
881
925
969
1013
1057
1101
1145
1189
1233
1277
1321
1365
1409
1453
1497
1541
1585
1629
1673
1717
1761
Timestamp in one-second interval (s)

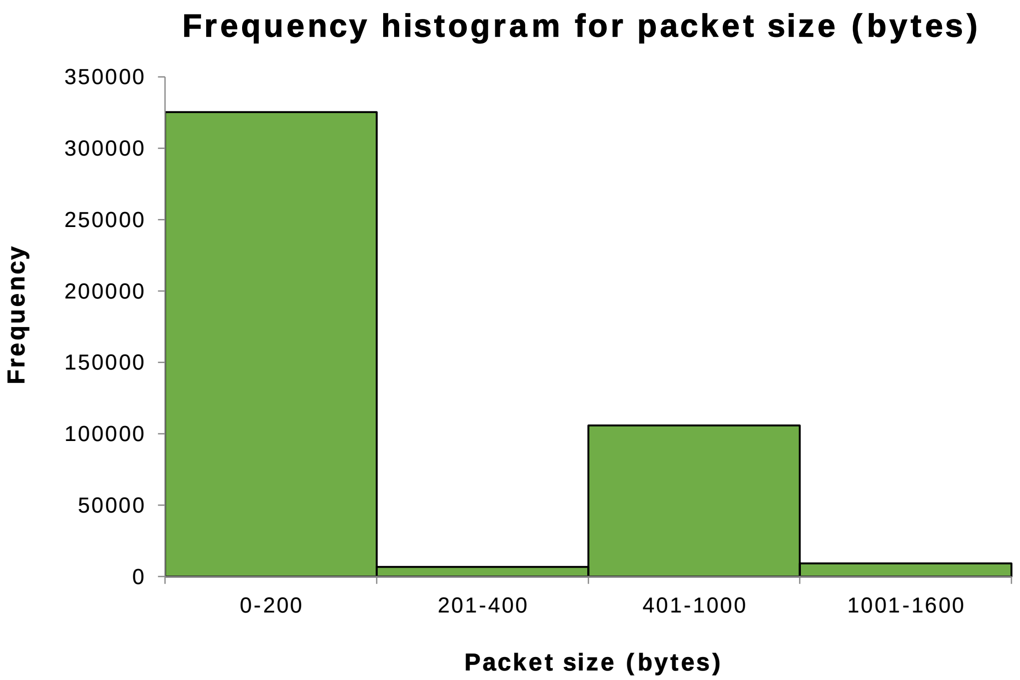

Data transferred versus Timestamp in one-second interval (s)
Data transferred (bytes)
y = -0.4539x + 41664
0
50000
100000
150000
200000
250000
300000
350000
400000
450000
1
48
95
142
189
236
283
330
377
424
471
518
565
612
659
706
753
800
847
894
941
988
1035
1082
1129
1176
1223
1270
1317
1364
1411
1458
1505
1552
1599
1646
1693
1740
1787
Timestamp in one-second interval (s)

The line graph in blue demonstrates the relationship between the packets transferred in the link and the timestamp in one-second interval in seconds. The line is generally moving up and down at an average of 250 across the timestamp.

The line graph in orange demonstrates the relationship between the data transferred in bytes and the timestamp in one-second interval in seconds. The line is generally moving up and down at an average of 41664 bytes across the timestamp.

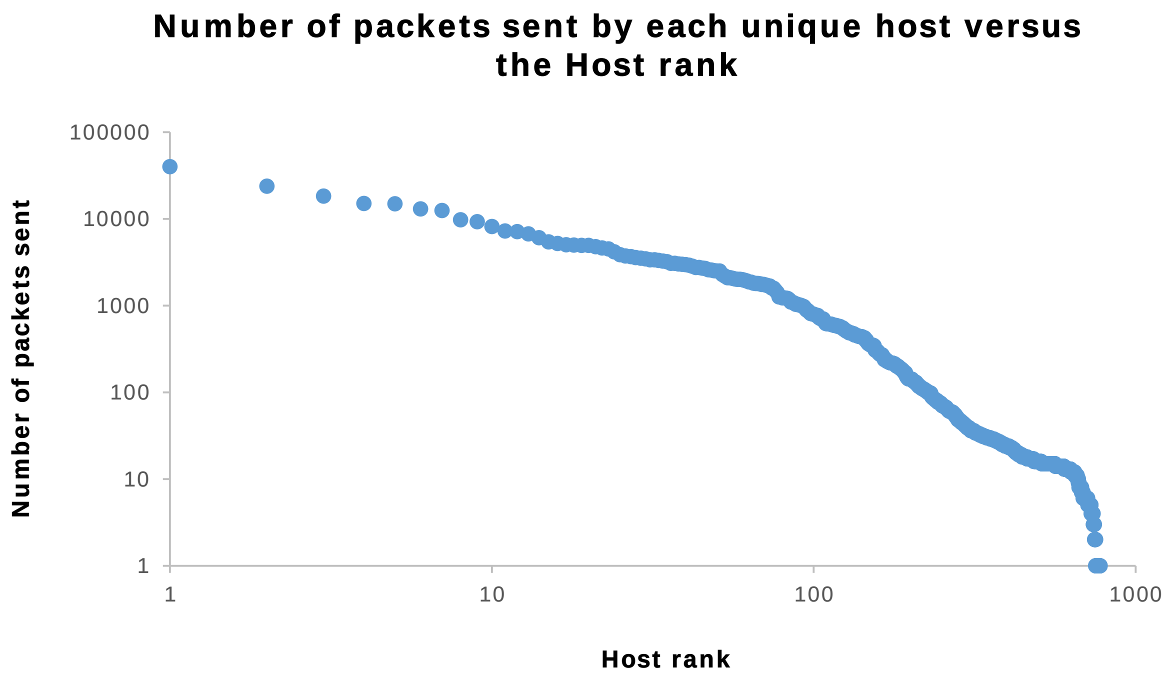
Among the up and downs, there are some bursty packet arrivals which are shown by the huge peaks. We can assume that at these huge peaks, the link is busier than other times hence a great amount of data could be transferred over the link.

(c)



The above frequency histogram is to show the packet size distribution. It is a bimodal distribution where there are two central clusters, at '0-200' and '401-1000'. Majority of packets are between 0 and 200 bytes, and the next most packets are in between 401 and 1000 bytes. Very few packets are in between 201-400 bytes and 1001-1600 bytes, compared to other categories.

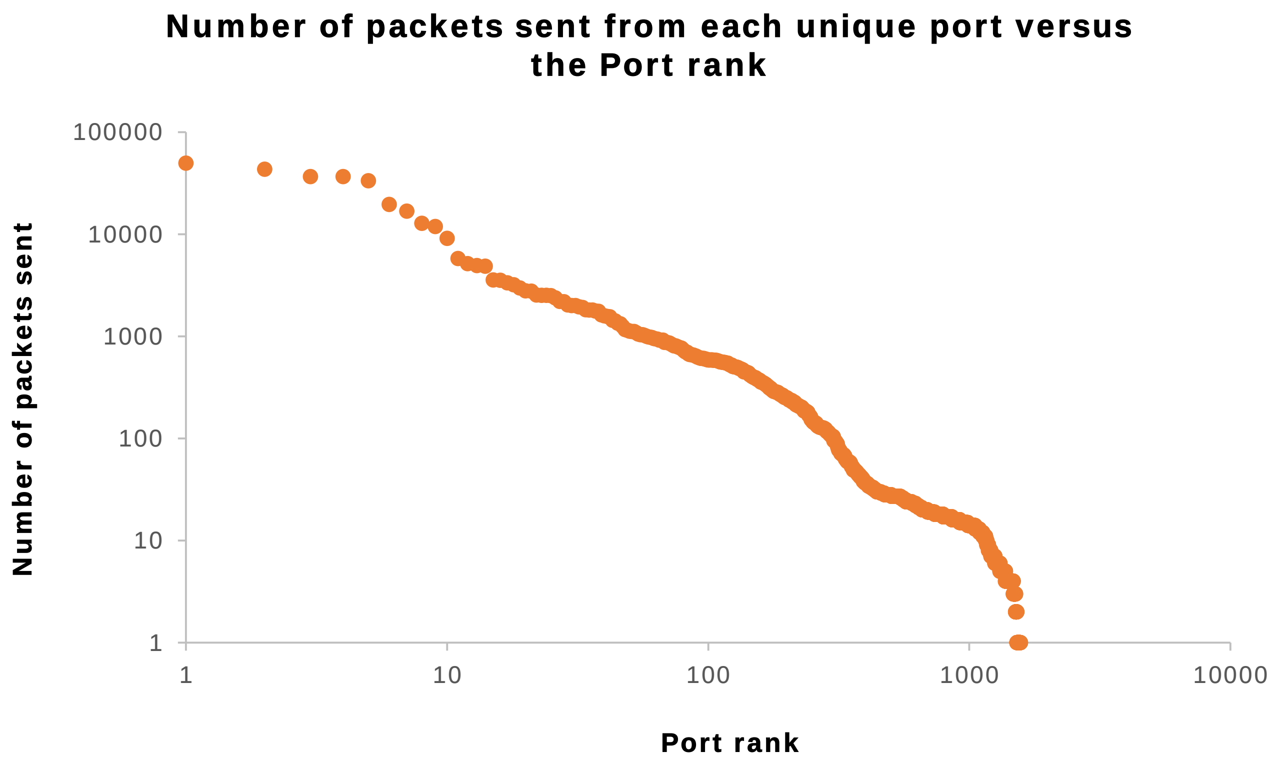
(d)



This scatter plot shows the relationship between the number of packets sent by each unique host and the rank of the hosts. Both axis are log scaled. The graph generally tells us that the lower the host rank, the less the number of packets sent.

Here, the plot exemplifies power-law behaviour in internet traffic. This is also known as Pareto law or 80-20 rule, where approximately "80% of the effects comes from the 20% of the causes". By looking at the shape and the distribution of the plots from the above graph, This demonstrates the power law where 80% of the usage comes from the 20% of the users.

(e)



Above graph is very similar to the graph in question (d), except it is a relationship between the number of packets sent from each unique port and the port rank.

According to the raw data, majority comes from the registered port (268082 bytes) and then the well-known port (176044 bytes) and very small usage of dynamic ports (3022 bytes). Port numbers are used to distinguish between different services that run over transport protocols such as TCP, UDP, DCCP, and SCTP. TCP protocol or Network News Transfer Protocol (NNTP) is represented as Port 119. It is an application protocol used for transporting Usenet news articles between news servers. The X Window System protocol, Port 6000 is the protocol used between an X client and server over the network. This demonstrates how applications can be identified using port numbers, and the method is fairly accurate.