COMP90007 Internet Technologies

Network Analysis Project, Project 1

Semester 2, 2020

Due Date: Wednesday September 23, 4:00pm

1 Introduction

This project forms 10% of your final mark. The key output of this project is a report which has to follow a certain format (refer to Section 5 of this document). The project is about measuring bandwidth, delay, and jitter in networks. These tasks will be similar to those you have performed in the laboratory sessions held in the tutorials. It is recommended that you perform these tasks in a consistent networking environment to reduce the variance in your report.

Important Note: As evidence of your work, when you run the following commands, please remember to take screenshots of the results obtained and place it in the appendix of your report. Reports failing to do so will be penalized. All the plots needed to answer questions should be placed in the main body of your report where you will explain the observations being derived.

2 Measuring the hop count

In this section, we will be observing the number of intermediate hosts in the route taken to communicate with a remote server and its relation to the physical geographical distance.

To count the number of hops taken to reach a destination host, the command `tracert` will be used (or its corresponding equivalent, depending on your operating system). This utility should be pre-installed on your operating system.

The utility can be invoked by launching a command line terminal and typing in the command. An example output of the `traceroute` command (on OS X) and `tracert` command (on Windows) is as follows:

```
$ traceroute -nw1 cis.unimelb.edu.au
traceroute to cis.unimelb.edu.au (128.250.37.164), 64 hops max, 52
     byte packets
1  10.0.0.254  533.676 ms  1.063 ms  0.940 ms
2  58.96.2.205  27.872 ms  28.137 ms  28.293 ms
3  58.96.2.129  28.647 ms  28.577 ms  28.085 ms
4  218.100.78.33  28.299 ms  28.469 ms  28.332 ms
5  202.158.200.9  29.626 ms  28.871 ms  29.841 ms
6  202.158.210.26  31.320 ms  28.722 ms  29.135 ms
7  202.158.200.250  29.668 ms  29.096 ms  28.660 ms
8  * * *
9  * * *
10  * * *
11  128.250.37.130  957.521 ms  33.475 ms  29.891 ms
12  128.250.37.164  29.940 ms  29.260 ms  30.020 ms
```

In this section of the project, you are interested in the number of hops it takes to reach the destination server. In the example above, the number of hops to reach cis.unimelb.edu.au is 12.

Based on the number of measurements you will be taking, there are some useful command line parameters you may wish to take advantage of, to speed up the time it takes to gather

results. The help documentation for the `traceroute` utility can be accessed by running `man traceroute` or `tracert /?` on Windows. You may also wish to investigate shell scripting to automate the collection of results, but this is not required for the project. Any scripts (Shell, Python, etc) you do choose to write, however, **must be included in the Appendix**.

2.1 Specific task description (2 marks)

Please include all raw measurements in the Appendix.

2.1 What do the command line parameters `-n w 1` (equivalently `-d -w 1` on Windows) mean in the example given above and what is the importance in using them?

2.2 Determine the hop count for the following hosts given in Table 1. It is recommended that students find one more public iperf server other than the ones listed here to gather their results.

Table 1: List of public `iperf` hosts

Host	Location
`iperf.he.net`	USA
`bouygues.testdebit.info`	France
`iperf.comneonext.de`	Germany
`ikoula.testdebit.info`	France
`st2.nn.ertelecom.ru`	Russia
`iperf.biznetnetworks.com`	Indonesia
`iperf.scottlinux.com`	USA
`speedtest.serverius.net`	Netherlands
`iperf.volia.net`	Ukraine

Determine the approximate geographical distance for the above hosts and plot the hop count versus the approximate geographical distance from the city you are currently in. Do you observe a correlation or not? Please explain your rationale with respect to networking concepts.

You may use any scientific computing package or spreadsheet software to do your plotting, **for example, Microsoft Excel**.

For finding out the physical grographical distance you may use any tool or application available online, for example, you may use a combination of: https://db-ip.com and https://www.freemaptools.com/how-far-is-it-between.htm or https://www.site24x7.com/find-website-location.html and https://www.distancecalculator.net/ or anything of your choice. However, do make sure to **document it and provide the appropriate reference** to that application/ tool/ software used.

Note: The servers listed in Table 1 are public servers and are not maintained by the University of Melbourne, hence they are likely to go down at any point in time. Based on past experience, it would be advisable to **conduct your tests on these as soon as possible (ASAP) rather than leave it till the end** as there is a high probability that these servers might not be available and this cannot be used as an excuse for a late submission. If these servers stop responding then please visit the link: https://iperf.fr/iperf-servers.php and find your own servers (anything that is responding) or feel free to find any public iperf server from the internet. Some alternate strategies worth exploring also include changing port numbers and trying to get the iperf metrics.

3 Measuring delay and jitter

In this section, you will be measuring the delay and jitter of the hosts used in Section 2, located in different geographical locations.

We will be using the `ping` utility, to measure the round-trip delay of packets. The `ping` utility should be pre-installed on all major operating systems. The standard deviation of the round-trip delay time will be taken as the value for *jitter* for this project.

The standard deviation measures the variation in a set of data. It is defined as the square root of the variance and **is expressed as follows:**

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

where $\bar{x}$ is the mean of the set of data. Details of this simple statistical measure can be found in many sites online if you do not remember this from high school years.

A sample output of the `ping` utility is shown below, but this output will vary depending on your operating system.

```
$ ping unimelb.edu.au
PING unimelb.edu.au (172.22.44.10): 56 data bytes
64 bytes from 172.22.44.10: icmp_seq=0 ttl=124 time=3.364 ms
64 bytes from 172.22.44.10: icmp_seq=1 ttl=124 time=3.416 ms
64 bytes from 172.22.44.10: icmp_seq=2 ttl=124 time=3.730 ms
^C
--- unimelb.edu.au ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 3.364/3.503/3.730/0.162 ms
```

In the output above, various statistics, including the mean and standard deviation, were calculated for you by the utility. It is expected that you record all the values of delay and perform your own calculations to find the mean values and their standard deviation, to confirm the result.

3.1 Specific task description (2 marks)

For this section, you may want to consider the user-facing implications of high delay and high jitter in networking applications, especially for applications sensitive to the affect of high delay and high jitter. Please include all raw measurements in the Appendix.

3.1 Measure the round-trip delay for the following hosts. Make **three** delay measurements (run this command 3 times - not 1 command gathering 9-10 rows of ICMP responses) of each host and find the average round-trip delay and jitter by calculating the standard deviation, for all the hosts used in Section 2.

For each of the above hosts, plot the average round-trip delay versus the approximate physical geographical distance to the server. Do the same with the jitter (i.e. jitter vs geo distance).

3.2 From the two plots above, do you observe any correlation between delay and jitter as a function of distance? Why? Why not? Explain your results with reference to the network environment in which you were collecting your results (this includes metrics like your download/ upload speed, users sharing the network, load on network through other apps, etc) and how does your networking environment influence your results obtained (examples required)?

4 Measuring the bandwidth-delay product

In this section, we will be measuring the bandwidth of different hosts in order to determine the bandwidth-delay product, using the results from the previous sections.

The utility that will be used to perform bandwidth measurements will be the `iperf` utility. This command line utility is available for download for all operating systems from https://iperf.fr. Alternatively you may choose to use the package manager for your relevant operating system.

There are two modes of operation in `iperf`. The server mode will host a server which will listen to incoming requests from a client. An `iperf` instance running in client mode will connect to the server, and packets will be exchanged and timed between the two hosts to calculate the bandwidth. In this project, we will be running `iperf` in client mode.

A sample output of `iperf` in client mode is shown below, noting the `-c` flag to designate operating in client mode.

```
$ iperf -c iperf.eenet.ee
------------------------------------------------------------
Client connecting to iperf.eenet.ee, TCP port 5001
TCP window size:  129 KByte (default)
------------------------------------------------------------
[  5] local 10.0.0.5 port 51878 connected with 193.40.100.7 port 5001
[ ID] Interval       Transfer     Bandwidth
[  5]  0.0-106.7 sec   128 KBytes  9.83 Kbits/sec
```

(Note: speedtest.serverius.net server) For speedtest.serverius.net in Table 1, we may need to use the port 5002, so the command can be: "iperf3 -c speedtest.serverius.net -p 5002" for this host. Also, **some iperf servers respond to iperf2, rest to iperf3** so please try to use both iperf2 and iperf3 to verify if the server is responsive to either version. If you have response from both iperf2 and iperf3 then you can pick iperf2 amongst them.

4.1 Specific task description (6 marks)

Please include all raw measurements in the Appendix as usual.

4.1 What does the bandwidth-delay product tell us about the networks? Collect **three** set of measurements (run this command 3 times) measuring the bandwidth of the public `iperf` hosts in Section 2 and find the mean bandwidth for each host.

4.2 Take the mean bandwidth and calculate the bandwidth-delay product in kilobits. You may use the mean round-trip delay time from your `ping` experiments to use as the delay time. Plot a bar chart for each host showing your results. You may wish to use a logarithmic scale, if appropriate.

Explain your results making reference to your networking environment in which you performed your measurements. How do your results reflect upon your actual internet link speed and how does your network environment influence your results obtained (provide examples)? Are there outliers in your data? If yes, point out the outliers and explain why they are marked as outliers in your data?

4.3 Plot the bandwidth-delay product versus the hop count. Do you observe any correlation?

4.4 When running your tests for bandwidth, delay, and jitter, were there any variables which may have affected the accuracy or reliability of your results? How might you improve upon these (explain your rationale with examples from your experience)?

5 Project Administration

This project is to be performed individually and is worth 10% of your overall mark in the subject.

5.1 Getting help

If you have any questions, the Canvas discussion board will be a useful resource in resolving any issues. If your concern is a personal matter, then you should email the subject coordinator.

Any answers posted by the subject coordinator or the academic staff on the Canvas discussion board will be considered as part of the project specification. Any announcements made about the assignment in the lectures will be considered part of the project specification. In addition, please keep an eye on any Canvas announcements to any changes made to the project specification.

5.2 Report submission

The deadline for the final report submission is as specified at the start of this project document.

The report will consist of all relevant discussion, graphs, data and answers from the experimentation conducted in this project. You must place the raw data as screens or copy paste them to the Appendix of the document, however, the diagrams like charts, flow diagrams (if any) and so on relevant to the discussion will be placed in the main document and not the appendix. Every diagram and/or raw measure used for a specific question must be referred to (using a designated reference scheme) for us to verify the result. All plots and figures must be appropriately labelled. Any information obtained that is not of your own work must be cited.

The report must be submitted as a PDF file via Turnitin on Canvas (will be available for submission soon). Please include your name, student Id and login user name on the top of the first page. The report is to be formatted on A4 sized paper in 10 pt text, 1.5 line spacing, single column. It is highly recommended that students use the respective formatting scheme outlined by us in the **Format Guideline Document** and here as massive deviation from it may incur relevant penalties.The report should not exceed a maximum of 10 pages (excluding appendix) else relevant penalties will apply.

Late submissions will attract a penalty of 10% per day (or part thereof). No submissions will be allowed passed 5 days after the deadline.