Chaehyeon Kim

cxk445

Report for CSDS 325 Project 4 Extra Credit (425) Portion

(Output files for the different modes can be found here)

Program Outputs:

```
time: first: 1431367242.907936 last: 1431368162.619505 duration: 919.711569
pkts: total: 5000000 ip: 5000000
```

From running the **summary mode** (-s) (screenshot of the output above), we know that

a. All packets are IP packets, and there are 5,000,000 of them.
    i. The number of total packets equals the number of IP packets.
b. All the packets recorded in the packet trace run over the span of 919.711569 seconds, which is equivalent to 15 minutes and 19 seconds.

The main observation is performed by running the **length mode** (-l) program on the given trace file. Using the resulting output file, the mean and the median of the following values across all packets are calculated:

a. Packet/captured length
b. IP header length
c. TCP header length (if applicable)
d. UDP header length (if applicable)
e. Application payload length (if applicable)

For the last three values, only packets containing the appropriate header/payload lengths are evaluated when calculating the mean and the median values; i.e. if a packet does not have a TCP header, it will be skipped over instead of assigning its length as a 0.

To obtain a table containing the values above, ***packetTraceAnalysis.py*** program was written and ran. The resulting table is shown below (screenshot taken from the terminal output):

	Packet	IP header	TCP header	UDP header	Payload
Mean	58.941564	20.0	30.270254	8.0	1194.375646
Median	66.000000	20.0	32.000000	8.0	151.000000

- Each column name specifies which length is measured; i.e. Packet = packet lengths.
- To calculate the mean across the specified type, the total length was added, then the sum was divided by the total number of specified packets.
    - ex) Mean of TCP header length = (total sum of all TCP headers' lengths) / (total number of packets containing a TCP header)
- To calculate the median across the specified type, a list containing all appropriate values of the specified type was created, then python's statistics.median() method was used.

Observations:

From the output above, we can infer a few things about the trends across the packets. From the -s mode's output, we observed that all packets recorded in the trace file include IP headers and are IP packets. Here, it can be seen that the mean and the median of the IP header lengths is 20 bytes, implying that almost all, if not all, packets have IP headers that are 20 bytes in length. From this, we can infer that most packets likely have IPv4 headers, as that is usually the standard length of IPv4 headers. We can also see that the packet was around 59 bytes long on average (considering all included headers; ethernet, IP, TCP, UDP, etc.) with a median that is higher (66 bytes). The higher median implies that there are a few exceptional cases of very big packets with most being around the median length. By comparing the lengths of the IP header and the packets, we can deduce that most packets had (median packet length) - (average IP header length) - (ethernet header length) = 59 - 20 - 14 = 25 bytes reserved for the remaining TCP/UDP/other transport layer header after the ethernet and the IP header. This also implies that most packets include the header for the transport layer. There are a few points that can be taken away from looking at the mean/median lengths of the transport layer headers, too. The table shows that both the mean and the median lengths of UDP headers are 8 bytes, which is consistent with the fact that all UDP headers are 8 bytes long. The TCP headers' mean and median lengths are around/at 30 and 32 bytes respectively. This could mean that most packets with the TCP header had options on top of the required information/fields as the minimum is only 20 bytes. Lastly, with the payload's value having a mean around 1194 bytes and a median at 151 bytes, we can infer that there were a few packets with very high payload lengths, skewing the distribution as a whole. It can also mean that those packets that included the payload information usually had around 151 bytes worth of payload. Overall, we can see that the packets mostly likely had header lengths that are consistent with the norm and that most of them included options and headers up to the transport layer and more.