Programming Assignment 2: TCP and Wireshark
CSE 310, Spring 2020
Instructor: Aruna Balasubramanian
Due date: March 13 2020, 9.00pm

The goal of this assignment is to dissect the TCP protocol using the Wireshark tool. To do this, you should be familiar with the packet formats, PCAP files, TCPDump, and Wireshark.

The goal here is to capture packets that are going on the wire---both packets from the computer and to the computer. TCPdump is the command-line tool that analyzes the packets captured on the wire. Wireshark is the graphical version of TCPDump.

Your goal is to write a version of TCPDump. To do this, you need to parse PCAP files. PCAP is the file format used to store packets captured on the wire. PCAP files are in binary format and cannot be read easily. A PCAP library is used to parse the binary packet. Your goal is to write a parser, similar to TCPDump that parses the packet with the help of the PCAP library.

Part A PCAP Programming Task and flow-level information (70 points)

Your task is to write a program analysis_pcap_tcp that analyzes a PCAP file to characterize the TCP flows in the trace. A TCP flow starts with a TCP "SYN" and ends at a TCP "FIN" between two hosts with fixed IP address and ports. There can be multiple TCP flows at the same time between the two hosts, on different ports.

You may use a PCAP library to analyze this file. Example PCAP libraries are provided in the end of this assignment. A PCAP library helps convert a PCAP packet from binary to byte format. You need to then write code to analyze the bytes to get the information about the packet.

[Important: You can create your own packet structures and read the bytes into the structure. This will let you easily parse the bytes rather than doing byte operations. However, you cannot convert the PCAP file into text for analysis.]

In the resources section, you will find a pcap file assignment2.pcap. In this file, we have captured packets sent between 130.245.145.12 and 128.208.2.198. Node 130.245.145.12 establishes the connection (let's call it sender) with 128.208.2.198 (let's call is receiver) and then sends data. The trace was captured at the sender. Your `` analysis_pcap_tcp" code should take as input assignment2.pcap and output the answers to these questions on the screen (Ignore non-TCP traffic):

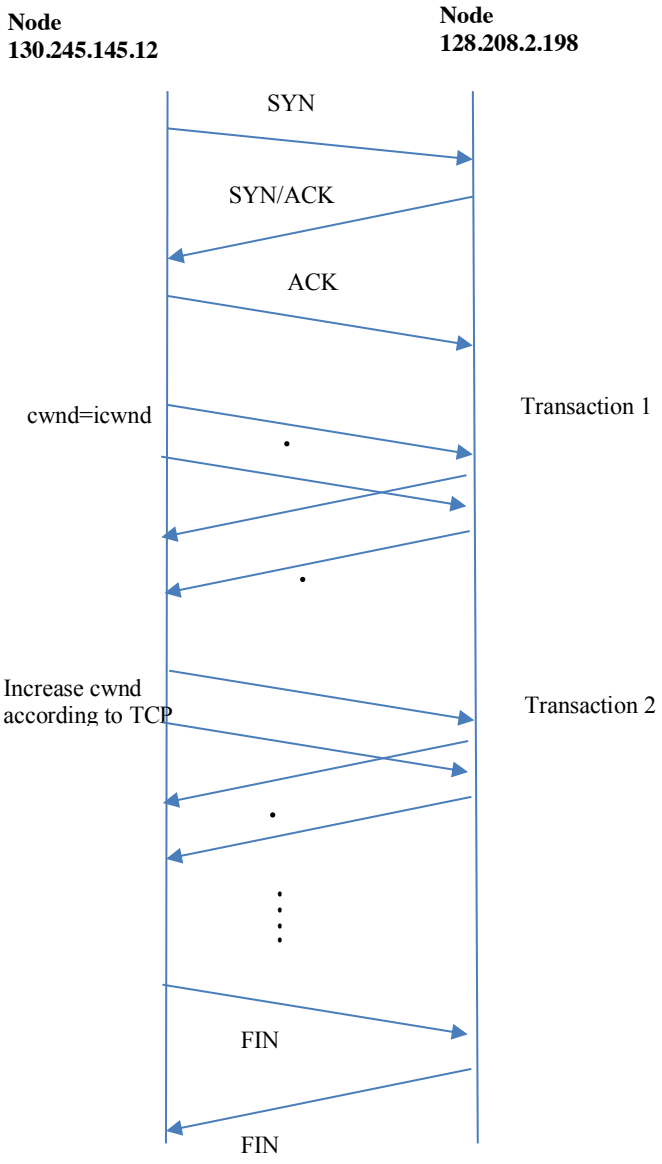
- The number of TCP flows initiated from the sender

- For each TCP flow

(a) For the first 2 transactions after the TCP connection is set up (from sender to receiver), the values of the Sequence number, Ack number, and Receive Window size.

(b) The sender throughput. The throughput is the total amount of data sent by the sender over the period of time. The period is the time between sending the first byte to receiving the last acknowledgement. For throughput, only consider the packets at the TCP level (including the header). You can ignore all other headers and acks.

Below is the example of a flow. This is only an example, your flow may look different compared.



Part B Congestion control (30 points)

Now extend your program and output the answer to the questions below along with the answers to Part A. For each TCP flow:

(1) Print the first 5 congestion window sizes (or till the end of the flow, if there are less than 3 congestion windows). The congestion window is estimated at the sender. You need to estimate the congestion window size empirically since the information is not available in the packet. Comment on how the congestion window size grows. Remember that your estimation may not be perfect, but that is ok. Congestion window sizes at roughly RTT-intervals.

(2) The number of times a retransmission occurred due to triple duplicate ack and the number of time a retransmission occurred due to timeout.

Submission Instruction

As before, you may write your programs in Python or C/C++. If you want to write in any other language, please talk to me. Viewing these traces on Wireshark can be helpful.

You need to submit your homework in a single zip file as follows:

- The zip file and (the root folder inside) should be named using your last name, first name, and the assignment number, all separated by a dash (‘-‘) e.g. lastname-firstname-assignment3.zip
- The zip file should contain (i) the high-level summary of the analysis_pcap_tcp code including how you estimated the answers to the questions in Part A and Part B, (ii) the analysis_pcap_tcp program, and (iii) instructions on how to run your code

Some example pcap libraries that you can use:
C/C++ - libpcap
Python - dpkt