1. [8+2 = 10 pts] Consider a modification to TCP’s additive-increase, multiplicative-decrease congestion control algorithm wherein a connection reduces the threshold by a constant amount upon detecting packet loss rather than by reducing the threshold to one-half of the current congestion window. Would the resulting additive-increase multiplicative-decrease algorithm retain the property that a set of two connections sharing a link would receive roughly equal shares of the link’s capacity? Does your answer depend on the initial rates (starting point) of the two connections? Explain.

2. [10 pts] The analysis for dynamic congestion windows that we presented in class implicitly assumed that there was only a single network link between the sender and the receiver. Redo the latency analysis for sending an object of $O$ bytes assuming that there are $T$ links between sender and receiver. To simplify the analysis, assume all links have the same transmission speed and that there is no congestion in the network (i.e., that there are no queuing delays at any network interconnection point). Assume that $RTT$ represents simply the total propagation latency on the round-trip path. You may also ignore the processing overhead at each network interconnection point.

3. [10+6+7 = 23 pts] Consider a web browsing session on a network where a user requests a page with 10 embedded, 5 Kbyte images (assume that the base page is also 5 Kbytes long). Assume a TCP implementation with an MSS of 536 bytes (a common MSS in many TCP implementations). For each of the questions below, consider networks with bandwidths of 28 Kbps, 100 Kbps, 1 Mbps, and 10 Mbps.

   a) Assuming a round-trip-time of 100 ms, construct a chart that compares the response time for retrieving the web page and its embedded object under a persistent and a non-persistent HTTP connection.

   b) Construct a second response time chart assuming an $RTT$ of 1 second.

   c) Consider now a browser using parallel non-persistent HTTP connections (recall that this is the most common use of HTTP). Let $x$ denote the maximum number of parallel connections the browser is permitted to open to a given server. A browser will first open an HTTP connection to retrieve the base page and then upon receiving and parsing this file, the browser will initiates $x$ parallel connections to retrieve the embedded objects. Show that the response time for the download of the web page and its contents is of the form

   $$(M+1)O/R + 2(M/x + 1)RTT + SSL$$

   where $M$ is the number of embedded objects, $O$ is the size of an embedded objects (assumed to be a constant for all embedded objects), $R$ is the capacity of the link connecting the browser to the server, and $SSL$ is the latency of TCP slow-start due to server stalling.
4. [3+3+2 = 8 pts] Consider the procedure used in TCP for estimating the RTT of a connection. Suppose that \( x = 0.1 \). Let \( \text{SampleRTT}_1 \) be the most recent RTT sample, and let \( \text{SampleRTT}_2 \) through \( \text{SampleRTT}_4 \) be the next 3 most recent samples. For a given connection, suppose four ACKs have been returned and that the four RTT samples previously mentioned have been generated. Express \( \text{EstimatedRTT} \) in terms of the four sample RTTs. Next, generalize your expression for \( n \) sample RTT values. If you then let \( n \) approach infinity in the general formula, comment on why this averaging procedure is called an “exponential moving average.”

5. [5+5+5 = 15 pts] Suppose that TCP increased its window size by two segments rather than one for each ACK received during slow-start. In this scenario the first window would consist of 1 segment, the second window would consist of 3 segments, the third window would consist of 9 segments, etc. For the variables \( K, P, \) and \( Q \), used in the analysis of TCP latency:

a) Express \( K \) in terms of \( O \) and \( S \).

b) Express \( Q \) in terms of RTT, \( S \), and \( R \).

c) Express the latency of a TCP connection in terms of \( P, O, R \), and RTT.