CS 372 Introduction to Computer Networks
Self-Check Exercises: Lecture 22 **Solutions**

1) What are the two major approaches to network congestion control?
Network core assistance: The network core senses when congestion is an issue – directly – and sends messages (either to the destination host, which bounces it back to the source host, or directly back to the source host) indicating congestion in the core, and possibly advising on a course of action.
Host inference: A source host attempts to infer congestion in the network core based on observables (e.g. round trip time, dropped packets)

2) In basic terms, how is congestion control implemented in TCP? What is this method called?
When no congestion is perceived, the sliding window size gradually increases (additive increase). When there is perceived congestion, the sliding window size is drastically reduced (multiplicative decrease). Together this makes up an additive-increase multiplicative-decrease (AIMD) scheme.

3) What is an MSS?
Maximum Segment Size. This is the maximum amount of transport-layer data which can be sent, such that it will fit within one link-layer data frame. This ranges from the 500s for a dial-up modem to around 1460 for cable internet.

4) Describe TCP's "Slow Start" mode. What is its purpose?
The purpose of TCP Slow Start is that it allows for a very slow start, but a rapid increase, in the size of the congestion window. By doubling the CongWin every time the transmitted packets ACK within the timeout interval, there is an exponential increase in the rate of transmission (until a threshold is reached).

5) Describe TCP's "Fast Retransmit" technique. What is its purpose? What affect does this have on congestion control?
TCP's fast retransmit (from RDT) is intended to more quickly infer the loss of a packet, and retransmit it, to increase data throughput. If a host receives three duplicate (so four total) ACKs "requesting" the same segment number, it assumes the segment starting with that sequence number was dropped, and retransmits the segment, even if the timer for that segment has not expired. In terms of congestion control, this assumed loss is perceived as congestion, but given less weight than a full timout. The CongWin is cut in half, rather than set to 1 MSS (as in a timeout).

6) A host has started a TCP transmission with MSS = 1460 bytes. It uses “slow-start”, with a limit of 11680 bytes.

a. What is the congestion window size after sending 15 packets? Assume all packets have length = MSS. Show your work. 13140 bytes

Slow-start begins with 1 MSS, then doubles each round until limit is reached. Then it increases linearly until loss or 3 duplicate ACK’s.

1) CW = 1 MSS (1460 bytes), send packet #1

2) CW = 2 MSS (2920 bytes), send packets #2 and #3

3) CW = 4 MSS (5840 bytes), send packets #4 - #7

4) CW = 8 MSS (11680 bytes), send packets #8 - #15

Since limit is reached, next increase is linear: 11680 + 1460 = 13140

b. Suppose that after sending packet #8, we get 3 duplicate ACK’s. What is the new congestion window size? 7 MSS (10220 bytes)

c. Suppose that after sending packet #7, we get a timeout on packet #4. What is the new congestion window size? 1 MSS (1460 bytes)