Problem 1

1.

Method 1: Please refer to figure 3.51 of the textbook.

Suppose, that the plot starts at A, with AIAD, the plot can only move back and forth along line AB, which is parallel to the Equal bandwidth share line. There is no movement of the plot in the direction of the Full bandwidth utilization line. So, the AIAD method generally will not converge to fairness.

With MIMD, The plot can only move back and forth along line BC, so MIMD will not converge to fairness either.

Method2:

In the paper, convergence to fairness is defined as moving towards the fairness index of one as time goes to positive infinity.(page 8)

$$F(x(t)) \to 1 \; as \; t \to \infty$$

Based on linear control policies, the fairness equation can be stated as follows:

$$F(x(t+1)) = F(x(t)) + (1 - F(x(t))) \times \left(1 - \frac{\sum x_i^2(t)}{\sum (c + x_i(t))^2}\right)$$

$$\begin{aligned} & where: \\ & \quad c = a / b \\ & \quad a = additive \\ & \quad b = multiplicative \end{aligned}$$

Since the above equation is an increasing function of c, it is vital to ensure that c ≥ 0 to guarantee non-decrease of fairness. Hence, the following conditions must be fulfilled.

$$\begin{aligned} a_I &\geq 0, & b_I &\geq 0, \\ a_D &\geq 0, & 0 \leq b_D &< 1 \end{aligned}$$

For AIAD, the x is stated as follows (page 4):

$$x_i(t+1) = \begin{cases} a_I + x_i(t) & if\ y(t) = 0 \Rightarrow Increase \\ a_D + x_i(t) & if\ y(t) = 1 \Rightarrow Decrease \end{cases}$$

$$\begin{aligned} &where: \\ &\quad a_I > 0 \\ &\quad a_D < 0 \end{aligned}$$

As we can see, the reason why AIAD doesn't converge to fairness is because of the a_D term. It doesn't satisfy the requirement of being greater than and equal to 0.

For MIMD, the x is stated as follows (page 4):

$$x_i(t+1) = \begin{cases} b_I x_i(t) & if\ y(t) = 0 \Rightarrow Increase \\ b_D x_i(t) & if\ y(t) = 1 \Rightarrow Decrease \end{cases}$$

$$\begin{aligned} &where: \\ &\quad b_I > 1 \\ &\quad 0 < b_D < 1 \end{aligned}$$

Although the variables satisfy the non-negative requirement, it does not contain the 'a' term. From our equation on fairness, we know that there is the c term which requires both 'a' and 'b' terms. In MIMD, though the 'b' term is bounded, the 'a' term is not. Hence, the MIMD doesn't change the fairness.

// 10 points each for correct analysis of the AIAD and MIMD //

Problem2.

2. Concepts: RTT estimation

One example:

Lets suppose that the 4 sample RTTs are 1 sec, 6 sec, 20sec, 20 sec. The initial AverageRTT is 4 sec. The initial deviation is 0. The EWMA is 0.125.

We are using the following formulas:

AverageRTT(k)=AverageRTT(k-1)*(1-x)+x*SampleRTT(k);

Deviation(k)=(1-x)*Deviation(k-1)+x*|SampleRTT(k)-AverageRTT(k-1)|;

Using method1:

Timeout=2*AverageRTT(k);

Using method2:

Timeout=AverageRTT(k)+4*Deviation(k);

Then we have,

SampleRTT(1) 1 sec	AverageRTT(1) 3.625 sec	Deviation(1) 0.375 sec
SampleRTT(2) 6 sec	AverageRTT(2) 3.922 sec	Deviation(2) 0.625 sec
SampleRTT(3) 20 sec	AverageRTT(3) 5.932 sec	Deviation(3) 2.558 sec
SampleRTT(4) 20 sec	AverageRTT(4) 7.691 sec	Deviation(4) 3.997 sec

With Method 1, we have the timeout=2*AverageRTT(4)= 15.382sec

Because SampleRTT(4) is 20 sec, then we have a spurious timeout.

With Method2, we have the timeout =AverageRTT(4) +4*Deviation(4)

= 23.6749sec

There is no spurious timeout using method2.

// Correct procedure 10 points; correct example 5 points.

Method 1 leads to more spurious timeouts in general than method 2 using a deviation. That is because method 2 includes the information of how much the sample RTT deviates from the estimate RTT. Furthermore, method 2 puts more weight on the deviation of the most recent sample RTT from the estimate RTT, because the most recent sample RTT have a better reflection about the behaviors of the future connections.

In other words, the estimation methods using only the average can only reflect the static nature of the samples. Deviation provides information about some additional dynamic natures of the samples such as the estimation errors. So, method 2 can have a better performance than method 1 in estimation. //10 points//

Problem 3.

Here we are assuming that the T links are in series between the client and the server. In this case, a packet will have (T-1) additional transmission times. Therefore,
The time for the server to send out the first segment until it receives the acknowledgement is:

TS/R+RTT

The latency will have four components:

1. 2 RTT: time to set up the TCP connection and requesting the file;
2. O/R: time to transmit the object at the first link;
3. (T-1)*(S/R), the additional transmission times for the last packet;
4. The sum of the stalled times.

If students missed the third component and only have component 1,2 and 4 for the latency, DO NOT DEDUCT ANY POINTS. *

Let K be the number of windows that cover the object.

We have

K=ceiling of $\{ \log_2(\frac{O}{S}+1) \}$

Let Q be the number of times the server would stall if the object contained an infinite number of segments.

Then,

Q=integer part of $\{ \log_2(T+\frac{RTT}{S/R}) \}$ +1

Define P=min{Q, K-1},

We have

The stalled time= $\sum_{k=1}^{p}(RTT + T*(S/R) - (S/R)*2^{(k-1)})$

Therefore,

$$\text{Latency} = \frac{O}{R} + 2*RTT + (T-1)*(S/R) + \sum_{k=1}^{p}(RTT + T*(S/R) - (S/R)*2^{(k-1)})$$

$$= \frac{O}{R} + 2*RTT + (T-1)*(S/R) + P*RTT + (PT - (2^{p} - 1))*(S/R)$$

Equation 3.1

//15 points for equation 3.1, partial credits for correct logic//.

$$\text{Minimum Latency} = \frac{O}{R} + 2*RTT + (T-1)*(S/R)$$

** If students missed the third component, (T-1)*(S/R), in the minimum latency, DO NOT DEDUCT NY POINTS. **

Equation 3.2

//5 points for equation 3.2//

Table 3.1 gives the values of both the minimum latency and the latency with slow start.

O	R	RTT	Minimum latency	P	Latency with slow start
100K byte	28K bps	0.1 sec	29.384 0 sec	3	30.90914 sec

100K byte	10M bps	0.1 sec	0.28172 sec	7	0.94227 sec
5K byte	28K bps	1 sec	4.041142sec	3	8.26629 sec
5K byte	1M bps	1 sec	2.05715 sec	3	5.09146 sec

Here we have T=5, S= 536 bytes.

//10 points for getting the table//

We can find that:

- For a large object, slow-start adds big delay when the transmission rate is high.
- For a small object, slow-start adds appreciable delay when the transmission rate is high.

// 5 points for the analysis of the results//

Additional questions for 600 level students.

Problem 1.

(a):

In the period of time from when the connection's window varies from (W. Mss)/2 to W.Mss,

The total number packet sent is

$$N = \sum_{k=0}^{W/2} (W/2 + k) \qquad \text{//10 points//}$$

$$= W/2 * (W/2 + 1) + \frac{(W/2) * (W/2 + 1)}{2}$$

$$= \frac{3}{8}W^2 + \frac{3}{4}W \qquad \text{//5 points//}$$

Therefore,

$$L = \frac{1}{\frac{3}{8}W^2 + \frac{3}{4}W}$$

(b).

The total time to transmit all the N packets (including the last one that is lost) is $(W/2 + 1) * RTT$. Recall that here we are analyzing the idealized model for the steady-state dynamics of the TCP connection. // 5 points//

Therefore, the average bandwidth of connection is

$$B = \frac{Mss}{L * (W/2 + 1) * RTT}$$

$$= \frac{Mss}{sqrt(L) * RTT * sqrt(L * (W/2 + 1)^2)}$$

$$= m * \frac{Mss}{sqrt(L) * RTT}$$

where $m = 1 / sqrt(L * (W/2 + 1)^2)$

$$= \frac{1}{sqrt\left(\frac{\frac{1}{4}W^2 + W + 1}{\frac{3}{8}W^2 + \frac{3}{4}W}\right)}$$

If W is large, we can neglect the first and zero order components of W in equation 3.1.

Therefore,

$$m \approx \frac{1}{sqrt(\frac{\frac{1}{4}W^2}{\frac{3}{8}W^2})}$$

$=1.22$ //10 points//

Therefore,

$$B \approx 1.22 * \frac{Mss}{sqrt(L) * RTT}$$