CSEE 4119 Fall 2022
Homework 1 ANSWER KEY

Assigned: 2022-09-23. **Due: 11:59 PM, 2022-10-05**

Uploading course materials, including questions from this homework, to sites such as CourseHero, Chegg or Github is academic misconduct at Columbia (see pg 10).

Question 1: Academic Honesty [1 Point]

[1 point] Please read the Academic Integrity policy on a version of the syllabus dated 2022-09-06 or after (see the section Updates to this Document). Then please answer this question with one of the following:

- I agree to abide by the policy. I do not have any questions about the policy at this time.
- I agree to abide by the policy. I have posted my questions on Ed Discussion before submitting the homework.

Question 2: Internet Speed Measurements [15 Points, Parts a-e]

1. [10 points] Follow this Google Form Survey and answer all the questions. During the survey, you will be asked to perform two Internet speed tests on the Speakeasy Internet Speed Test, one through your WiFi/cable network, and one through your cellular carrier. If you don't have a cellular plan, perform the second measurement through a different WiFi/cable network. You may be asked to perform some calculations based on your speed test results. Fill all of your answers in part (a) directly into the survey, and do not submit anything in your homework PDF. We will look at your Google Form submissions and score your responses accordingly.

Grading Criteria: the theoretical RTT question is worth 1 point. For the rest of the 18 blanks, each of the questions is worth 0.5 point. Total is 10 points. We used a script to check (1) if the students' theoretical RTT makes sense and (2) if the student's actual RTT is greater than the theoretical RTT. The script could be found here.

2. [2 points] How does your measured download bandwidth in part (a) compare to your measured upload bandwidth? Are they roughly the same, or is one is faster than the other? Explain the results you observed.

The question will depend on the student's actual speed test results. Most of the students should report that their download is faster than upload. This is because human users download much more data than their uploads, so the ISP designed the speeds to be asymmetrical, with the download being much faster usually.

However, some ISPs, such as Google Fiber, promise symmetrical speeds to the user. If the student is on a campus network, that campus network might also reserve a large upload capacity, and therefore make upload speed similar to or sometimes even faster than the download speed in a speed test.

The students should receive full points on these questions as long as their explanation makes sense.
+0.5 point for correct comparison.
+1.5 point for students who explained that because ISPs usually provision more download capacity than upload capacity, so their download speed appears to be higher than upload speed.
+1.5 point for students who explained that their ISP's download link might be oversubscribed, so the upload speed is faster than download speed.
+1.5 point for students who explained that the asymmetrical physical layers in download vs upload might have resulted in differences in upload vs download speeds, and gave a specific example that is relevant to their observations. Examples include, but not limited to: carrier aggregation for wireless networks; wireless AP usually has more transmission power than the mobile device, making the download generally faster, etc.
+0.5 point for students that explained that the asymmetrical physical layers in download vs upload might have resulted in differences in upload vs download speeds, but did not offer any specific and relevant examples.
+0 point for no answer/incorrect answer.

3. [1 point] Speedtest results roughly work by calculating the time it takes to download/upload a large file. Some speed test providers, such as fast.net, provide you with an "unloaded" and "loaded" latency. Unloaded refers to latency before sending lots of traffic into the network, whereas loaded refers to latency while downloading/uploading the file. Why are these different?

Loading the network may cause queues to build up in routers, adding queueing delay thus increasing end to end latency.
+1 point for answers that explained both (1) loading the network will cause queues to build up in the switch and routers and (2) longer queues will lead to larger latency.
+0.5 point for answers that explained (1) loading the network will cause queues to build up in switches and routers, but failed to correctly explain longer queues lead to larger latency.
+0.5 point for answers that explained (2) longer queues will lead to larger latency, but failed to correctly explain loading the network will cause a longer queue.
+0.25 point for answers that explained what "loaded" and "unloaded" latency are, but not why the results are different.
+0 point for incorrect answers/no submission.

4. [1 point] Assume that, to implement *reliable* delivery, some transport layer protocols lower sending/receiving speeds when they detect loss. Given this information, why might ISPs have "deep" buffers (buffers that can hold lots of packets)? (Hint: ISPs want their speed test numbers to look good!)

Loss occurs when buffers fill up and routers drop packets. ISPs may have deep buffers to avoid loss so that transport layers don't slow down, making speedtest numbers look better.
+1 point for answers that explained (1) loss occurs when buffers fill up and routers drop packets, (2) transport layers will slow down when loss occurs, and (3) ISPs would like to avoid transport layer slow down to make their speed test numbers look good
+1 point for answers that explained (1) loss occurs when buffers fill up and routers drop packets and (2) transport layers will slow down when loss occurs.
+0.5 point for answers that only explained (1) loss occurs when buffers fill up and routers drop packets.

+0.5 point for answers that only explained (2) transport layers will slow down when loss occurs.
+0 point for incorrect answers/no submission.

5. [1 point] What are some possible negative performance effects for users of ISPs having deep buffers?

Latency-sensitive applications like zoom or games might experience reduced quality due to higher RTT and more Jitter.
+1 point for mentioning any of the following: higher RTT or more Jitter.

Question 3: Protocol Layers and Service Models [20 Points, Parts a-i]

1. [1 points] What is an advantage of a layered Internet architecture?
2. [1 points] What is another advantage of a layered Internet architecture? (different from your answer to the previous question)
3. [1 point] What are the main disadvantages of layered architecture? Please provide one answer to this question.
4. [4 points] There are two versions of the Internet Protocol used on the Internet today. The newer version tried to overcome limitations of the older version by introducing new features—in particular longer addresses so that we never run out of addresses to assign end hosts. (i) What are these two protocols? What years were they standardized?

(ii) To what layer do these two protocols belong?
(iii) Though the adoption is increasing for the new version, the older version of Internet Protocol is still most widely used in terms of its share of Internet traffic. For example, only 3% of a large cloud/content provider's Internet traffic uses the newer version. Why do you think that is, given the benefits of the newer version?

5. [2 points] Our 5-layer model of the network stack lacks two layers that are present in the OSI model, Presentation and Session. These layers provide services such as data compression, encryption and exchange synchronization, etc. How does today's Internet provide these services?
6. [2 points] The two main Transport layer protocols are TCP and UDP. How does the selection of the transport layer protocol affect the behavior of a switch between two users?
7. [5 points] Alice and Bob's computers are connected by 2 intermediate routers (Topology: Alice --- Router1 --- Router2 --- Bob). How many times total is data processed by a host and/or router at each of the 5 layers, assuming one physical packet propagates between Alice and Bob? The answer should be 5 numbers, one for each layer. Leave your answer in the form {Number of times processed by application layer} - {Number of times processed by transport layer} - {Number of times processed by network layer} - {Number of times processed by link layer} - {Number of times processed by physical layer}

Answers:

1. Advantages(only give the credit when the answer is mentioned below):
   - Using a layered architecture has great compatibility between devices, systems and networks because they are following the same model in certain layer. For example, protocols at network layer for one device is compatible with protocols at network layer for device from another manufacturer.
   - Keep each layer isolated so that allow people to change the implementation of a service without affecting other components of the system.

- From a developer's perspective, working on a certain layer does not require the developer to know all the information about the whole model. Related, layers make it simpler to discuss a complex system.
- Ease of Network troubleshooting, giving full explanation.

+1 point total for any of the above answers
+0.5 point if reason only mentions that "it simplifies implementation" or any of above answer without substantial reason.

2. Same possible answers as first question, must not be repeated/restated

+1 point total for any of the above answers
+0.5 point if reason only mentions that "it simplifies implementation" or any of above answer without substantial reason.

3. Disadvantages (only give the credit when the answer is mentioned below):
    - Each layer will create more overhead because they all add data to the packet.
    - A layered model can make useful information in a certain layer invisible and inaccessible to other layers.
    - One layer may duplicate lower-layer functionality.
    - A layer is stuck with the behavior of layers below it, some of which it might not want.

+1 point total for any of the above answers.
+0.5 point if alluding to any of above answer without substantial reason.

4. (i) For IPv4 (0.5 points), both 1980 (RFC 760) and 1981 (RFC 791) are acceptable.

For IPv6 (0.5 points), any of 1995 (RFC 1883), 1998 (when RFC 2460 became a draft IPv6 standard), or 2017 (when IPv6 was formally ratified by IETF as RFC 8200) are acceptable.
If the student gives 1982 or 1983 for IPv4, only award 0.25 points. These are when IPv4 were first deployed, not when they became a standard.
1 point total
(ii) Network Layer.
1 point
(iii) Answers could include that it's expensive to upgrade to the newer version, the new features in IPv6 turned out to not be so useful or not useful enough to warrant the cost, people like working with something they already know/understand, network effects.
2 points, 1 point for something alluding to some "real-world" (as opposed to technical/logical) challenge even if the "challenge" the student mentions is not actually a barrier to IPv6 deployment . Like IPV6 is more resource intensive.
+3.5 if Either IPV4 or IPV6 year Incorrect
+3.75 point. Deducting 0.25 for using 1982 or 1983 for IPv4. These are when IPv4 was first deployed, not when they became a standard. Full points for d(i) and d(iii)
+3 point. For something alluding to some "real-world" (as opposed to technical/logical) challenge even if the "challenge" mentioned is not actually a barrier to IPv6 deployment. Like just mentioning IPV6 is more resource intensive.
+2 point. When Incorrect d(iii). Full points for d(i) and d(ii)

5. TCP/IP model does not provide services such as data compression, encryption and exchange synchronization. If the application developer thinks these functions are important, they can add them to the application layer. Some examples: TLS for encryption; optional compression in HTTP, gRPC for exchange synchronization, etc.

+2 point, If the answer mentions
- If the application developer thinks these functions are important, they can add them to the application layer.
- Can be entirely handled by the application layer.

+1 point, If the reason is unclear but alludes to getting handled by the application layer.

6. The switch's behavior will not be influenced by the choice of transport layer protocol because the header of transport layer protocol is in the payload of a link layer frame,

the switch will ignore the payload while switching the dataframe.

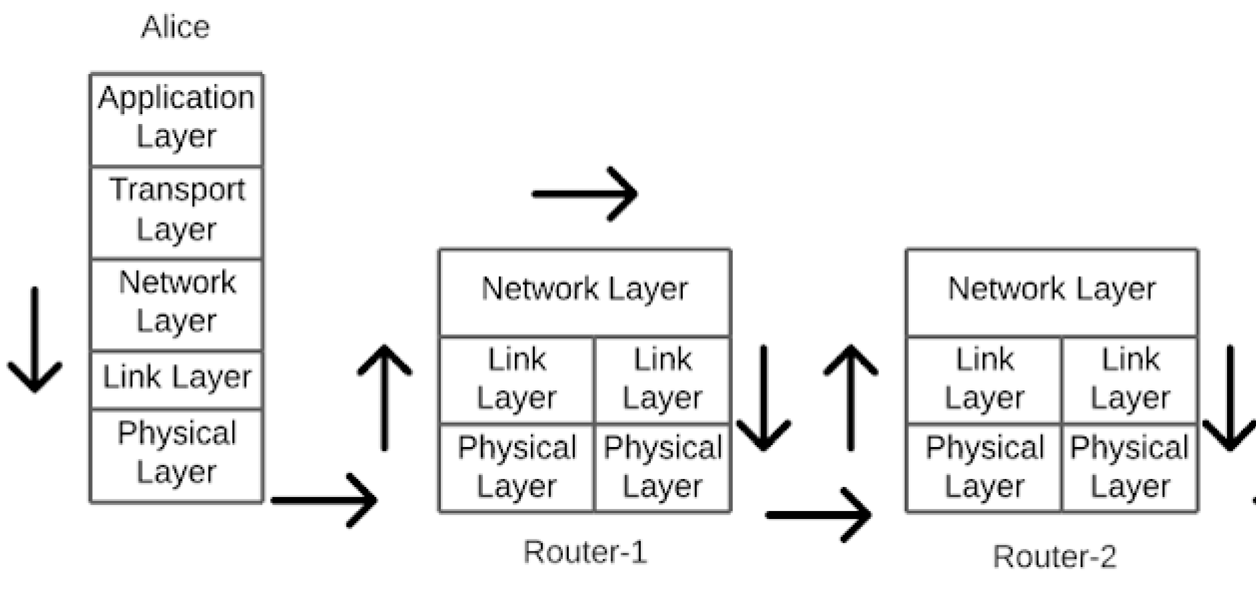
+2 point if answer mentions "Switch remains unaffected irrespective of the protocol"

+0.5 point if answer mentions In the case of TCP switch will have to process more packets than in the case of UDP.

7. From app to physical, 2-2-4-6-6

+1 point for each layer correctly counted and identified.

+0.25 if student correctly recognizes processing in that layer but with incorrect count.



Question 4: Circuit vs Packet Switching [20 Points, Parts a-h]

HD video is 1080p=1920x1080 pixels. Netflix recommends 5.0 Mbps for HD quality. Ethan sets up his phone as a hotspot so that everyone in class can share his Verizon 5G connection (at the speed he reported in class, which is 146 Mbps) to stream video.

1. [2 points] If FDM is used to divide the connection into circuits, assuming no traffic other than Netflix, how many students can watch HD video? Show your work.

146 Mbps / (5.0 Mbps/user) = 29 users

Explanation: to stream at HD quality, each user needs 5Mbps bandwidth.

+1 correct equation

+1 correct answer

-.5 no units

Netflix videos are (or at least were) encoded in 4 second chunks (see this paper). Netflix recommends 5.0 Mbps for HD, but it actually decides the encoding bitrate (bits per second) for different video resolutions (pixels per image) based on properties of the video. For example, BoJack Horseman is simply drawn, so it begins streaming 1080p video at 1.5 Mbps.

2. [2 points] Continue to assume that circuits were provisioned for 5.0 Mbps (which more complex HD videos might need), but the class is watching BoJack, with chunks encoded at 1.5 Mb (assume that includes all headers and overhead). What is the download time for a 4 second chunk of video? Show the math.

((4 seconds) * (1.5 Mbps)) / (5.0 Mbps) = 1.2 seconds
Explanation: a 4-sec video chunk is (4 s)*(1.5 Mb/s)=6.0Mb, therefore needing (6Mb)/(5.0Mb/s)=1.2 sec to download.
+1 correct equation
+1 correct answer
-.5 no units

3. [2 points] What fraction of time would the connection be idle for a provisioned 5.0 Mbps circuit when streaming BoJack (assuming the only traffic is the chunk downloads)?

Explanation: From b), downloading a 4-sec video chunk requires 1.2 seconds, so for 4-1.2=2.8 sec the connection is idle. That is to say, 2.8 / 4 seconds = 0.7 fraction of time the connection is idle.
+1 correct equation
+.5 if incorrect equation, but instead uses 1 - 1.5 Mbps / 5.0 Mbps to get answer
+.5 if calculated active time, but not idle time
+1 correct answer
-.5 no units
(4 seconds - 1.2 seconds) / 4 seconds = 0.7

4. [4 points] Suppose we switch to using *packet switching*. Use the idle time calculated in the last question (for streaming 1.5 Mbps on a 5.0 Mbps circuit). How many students can watch BoJack at once, with every student downloading chunks at 5.0 Mbps at a random time within a few minute window, while keeping the chance of a collision below 1%? Show your work.

Explanation:
As mentioned in lecture slides, the underlying distribution is binomial.
From question a), we know that at most 29 users can simultaneously download at 5.0 Mbps.
So to tackle this problem, we need to derive the probability that AT MOST 29 users are active at the same time.
From question c), we know that Pr[a user is active] = 0.3, Pr[a user is inactive] = 0.7
Therefore, we can model the probability that n users are active using a binomial distribution. Pr[out of k users, n users are active whereas (k-n) users are inactive]=[(k choose n) 0.3^n x 0.7^(k-n)], where (k choose n) = k!/[n!(k-n)!] Is the number of ways of assigning a group of k users into a set of n active users and k-n inactive users, without regarding their order.
We can thus derive, the probability that out of k users, up to 29 active users are active simultaneously:
Pr[out of k users, UPTO 29 users are active whereas other users are inactive]=sum_{from n=0 to 29}[(k choose n) 0.3^n x 0.7^(k-n)]
The reason why we can sum them up: **Pr[out of k users, 0 user is active k inactive] +Pr[out of k users, 1 user is active, k-1 inactive]+Pr[out of k users, 2 user are active, k-2 inactive]+...+Pr[out of k users, 29 user are**

active k-29 inactive], is because these events are DISJOINT (they can never happen at the same time. e.g. "exactly 3 users are active" can't happen at the same time as "exactly 4 users are active"), therefore, the sum of them is simply the probability that "out of k users, UPTO 29 active users are active", which is precisely the probability that there is no collision (recall that if more than 29 users are active simultaneously, collision happens)
Hence, the probability that collision happens is just its complement:
Pr[collision when there are k users] = 1 - sum_{from n=0 to 29}[(k choose n) 0.3^n x 0.7^(k-n)]
To find out the maximum k such that Pr[collision when there are k users] < 0.1. By the convexity of binomial coefficient, we can simply plug in different k's until we find such k', under which,
Pr[collision when there are k' user] <= 0.1 and Pr[collision when there are k'+1 user] > 0.1
(You can do this efficiently using binary search over, say, from k'=1 to k'=100)
68 users:
100*(1- sum [(68 choose n) .3^n .7^(68-n)] from n=0 to 29) = 0.9551%

$$100\left(1-\sum_{n=0}^{29}\binom{68}{n}\times 0.3^n\times 0.7^{68-n}\right)$$

69 users:
100*(1- sum [(69 choose n) .3^n .7^(69-n)] from n=0 to 29) = 1.1216%

$$100\left(1-\sum_{n=0}^{29}\binom{69}{n}\times 0.3^n\times 0.7^{69-n}\right)$$

Need to show both for full credit, because you need to establish that 68 can fit and 69 can’t fit.
+1 Correct formula
+1 Derivation of formula or justification based on knowledge of common distributions
+1 Show that 68 users can fit, but not 69 (either by solving the inequality or showing that P(collision) w/ 69 users > 0.01, and P(collision) w/ 68 users < 0.01
+1 Correct answer

5. [2 points] The network operator notices that everyone is only using 1.5 Mbps, so she decides to switch to provisioning circuits at 1.5 Mbps. With that change, how many students can watch simultaneously? Show your work.

146 Mbps / (1.5 Mbps/user) = 97 users
+1 Correct formula
+1 Correct answer
-0.5 for lack of units.

6. [4 points] Suppose that many students try to watch via packet switching, with downloads at 5.0 Mbps. What is the chance there will be a collision? Show your work.

Explanation: Like in part (d), the probability of a given user downloading is 0.3. The network can still support a maximum of 29 people downloading at the same time, so we need to find the chance of collision with 97 people periodically attempting to download at the same time. We can adapt the same formula derived in part (d).
+1 correct formula
+1 explanation (okay if explained in part (d))
+1 correct use of formula
+1 numerical answer

-0.5 for lack of units.
45.8875%:
100*(1- sum [(97 choose n) .3^n .7^(97-n)] from n=0 to 29)

$$100\left(1-\sum_{n=0}^{29}\binom{97}{n}\times 0.3^{n}\times 0.7^{97-n}\right)$$

7. [2 points] Consider the various answers you've found. Which download rate (1.5 Mbps vs 5 Mbps) and network design (circuit vs packet switching) is best when everyone is watching BoJack Horseman? Explain your answer in a few sentences.

Circuit switching at 1.5 Mbps is the best. Since we know everybody will be downloading at 1.5 Mbps, we don't need to worry about any users having an insufficient download speed. Trying to do packet switching at 1.5 Mbps would not allow any additional users without guaranteeing collisions, so it offers no benefit. Packet switching at 5 Mbps has a high probability of collision for the same number of users as calculated in (f). And circuit switching at 5 Mbps would support fewer users and waste bandwidth.
+1 Answer for circuit switching at 1.5 Mbps
+1 Logical justification for answer, even if answer is incorrect
-0.5 no justification for either using circuit switching or the bandwidth
Circuit switching at 1.5 Mbps. (Could try to do 5 Mbps with TDM.) We know exactly the demand, so we can provision it.

8. [2 points] Now suppose different students are watching different videos. Which approach should we use? Why is the same/different than in the last question? Explain in a few sentences.

Packet switching may be better, because we do not know the exact bitrate needed by various videos, so a fixed allocation could provide too much or too little, and it would be hard to do the time assignments if different users may start at any time.
Since some users might watch a video encoded at more than 1.5 Mbps, having a 1.5 Mbps connection would prevent those users from watching videos without constant buffering, so limiting to 1.5 Mbps connections is off the table. Circuit switching at 5 Mbps would have a lot of wasted bandwidth in the case in which a substantial portion of the users are watching BoJack (or a similarly low-bitrate encoding) anyway. Since we don't know what the usage will be, packet switching is the better option since it can support different needed data rates without wasting bandwidth.
+1 Answer for packet switching at 5 Mbps
+1 Logical justification for answer, even if answer is incorrect
+0.5 for saying to use packet switching (without providing bandwidth)
+0.5 for logical justification of using packet switching (but not bandwidth)

Question 5: To Segment or Not to Segment [31 Points, Parts a-k]

Normally large messages are broken down into smaller packets, a process called segmentation. In this question, we will compare segmenting to not segmenting (sending the message as one large packet). Suppose Bob wants to send a message that is M bits long to Alice.
With segmenting, the message is sent as k packets. Without segmenting, it is sent as a single packet of M bits. There are total N routers between Alice and Bob, and each link has bandwidth R bps. Ignore propagation delay and the time to break a message up and reassemble. Until mentioned, ignore headers and loss.

1. [2 points] With segmentation, what is the time to deliver the full message?

$\left(\frac{N}{k} + 1\right) \times \frac{M}{R}$

+2 points for correct answer

2. [2 points] Without segmenting, what is the time to deliver the full message?

$(N + 1) \times \frac{M}{R}$

+2 points for correct answer

3. [3 points] Suppose now each datagram delivered has an h(bits) header, both with and without segmenting. Without segmenting, the whole message is one datagram. With segmenting, each packet is one datagram. In which circumstances do the two approaches have the same end-to-end delay?

Transmission delay for message switching: $(N + 1) \times \frac{h + M}{R}$

Transmission delay for packet switching: $\left(\frac{N}{k} + 1\right) \times \frac{kh + M}{R}$

$(N + 1) \times \frac{h + M}{R} = \left(\frac{N}{k} + 1\right) \times \frac{kh + M}{R}$

$h = \frac{NM}{k}$

+1 point for correct message switching
+1 point for correct packet switching
+1 point for final answer

4. [1 point] Based on your answer for question a)~c), if there are many routers along the path, will delivery be faster with or without segmentation?

With segmentation

+1 point for answer

5. [2 points] Which mechanism (or behavior) is the root cause of the difference we see in d)? Explain in a few sentences.

Lower transmission delay allows for better pipelining. Because the switch uses **store-and-forward**, for message switching, the switch cannot begin to transmit the file until it receives it fully over the 1st link, resulting in a long delay before the switch can begin transmitting even the beginning of the file. But for packet switching, after the first packet arrives at the switch, both of the links could be busy transmitting segments of the files at the same time, thus decreasing the delay.

+2 points if store-and-forward is mentioned, or has a closely related expression (such as the router has to wait for the entire packet to arrive before forwarding it)

Next, we will explore segmentation in a lossy link. Consider transferring a file of length L=25.0 MB from Host A to Host B. The path from A to B is a single fiber link with transmission rate R=1 Gbps on which packets can be dropped randomly (not due to congestion). If a packet loss happens, the sender can sense the loss at the end of the transmission period for the lost packet and retransmit immediately. We will assume each packet (regardless of size) has a 2% chance of being dropped on each attempt. We will compare segmentation into packets of length 1.6 KB (assume no headers need to be added), vs no segmentation (single transfer of 25.0 MB).

6. [3 points] With segmentation, what is the expected time for the whole file to be transmitted? You don't need to give the exact result, but

please write down a formula from which the expectation can be calculated.

# of packets = 25.0 MB / 1.6 KB = 15,625 packets
1.6 KB / 1 Gbps = 12.8 μs to transmit one packet one time
Expected # of transmissions of a packet = E[n] =

$$\sum_{n=1}^{\infty} n(0.98)(0.02)^{(n-1)} = 1/.98 = 1.0204081632...$$

Expected time to transmit one packet = 80 μs * E[n] = 12.8 μs * 1/.98 = 13.0612245 ms
The chance of each packet experiencing loss is independent, so total delay is the # of packets times the time to transmit one packet:
15,625 packets * 12.8 μs * 1/.98 = 0.2 seconds/.98 = 0.204081632653 seconds
+0.5 point for correct # of packets
+0.5 point for time for transmitting 1 packet
+1 point for getting the correct first or second try of the total delay
+1 point for getting the full summation of the total delay correct
Can receive full credits giving the correct formula and no need to show the exact result.

7. [3 point] If segmentation is not used, will the expected transmission time be the same as with segmentation? (Same/Different) Explain your answer (either give a brief argument or a calculation).

Same
25.0 MB / 1 Gbps = 0.2 seconds to transmit the message one time
Expected # of transmission = 1/.98 still
Expected time to transmit the message = 0.2/.98 = 0.20408163265 seconds
+1 point for correct result
+2 points reason/calculation, will give full points if there is correct formula to calculate the transmission time or the argument clearly explains why the expected time will not change with or without transmission.

8. [6 points] Users don't have a straightforward way to assess the "expected" delay of their network. Instead, let's assume that a user does not complain at all if the transmission can be completed in less than 200% of the “ideal” delay (file size/transmission rate), and becomes very unhappy if it is 200% or more. For this question, we can refer to 200% as *unacceptable delay.* What is the probability of unacceptable delay with and without segmentation (using the same parameters as in the previous questions)? In this question (but not the previous one), we can assume that the 1st retransmission (the 2nd attempt) for any packet is guaranteed to be successful.

Please give your calculations for:

i) Segmented, delay at least 200%

Since we are assuming every retransmission works, this only occurs if every packet is retransmitted.
So 0.02^15,625, a very very small number.
+1 point correct unacceptable delay
+1 point correct situation (retransmit all packets), if the situation is correct, will also get the point for unacceptable delay
+1 point correct answer

ii) Not segmented, delay at least 200%

If it is not segmented, then it takes 0.2 seconds to transmit the message one time.
If it has to be retransmitted, then it will take 200%.

So this is the chance that it does not succeed on the first try. P(fail at first transmission) = 0.02
+1 point correct unacceptable delay
+1 point correct situation (retransmit a packet), if the situation is correct, will also get the point for unacceptable delay
+1 point correct answer

9. [3 points] Based on your answer for expected delay and unacceptable delay, discuss how segmentation affects the performance of file transmission. Specifically, what explains your results? Explain in a few sentences.

At the same per-packet drop rate, segmented transmission achieves the same expected delay, while delay is more likely to fall into the “acceptable” margin, and less likely to fall into the “unacceptable” margin. (Its performance becomes even better if packet drop rate grows with packet size.) The reason for that is the use of packet-based retransmission mechanism. Segmentation brings faster recovery from loss, since we don’t retransmit the whole file when there’s a packet loss.
+2 points for how segmentation affects the performance: ‘more likely to fall into acceptable margin in segmentation case’ (or similar meanings)
+1 point for explanation (mention something about the retransmission mechanism, will have the point as long as the explanation makes sense)
No point given if ‘how segmentation affects the performance’ is not correct.

Let’s now zoom in to the switches between Alice and Bob. Suppose that a pair of switches X and Y are connected by a fiber optic link. The first switch begins transmitting the largest IP packet allowed by Ethernet, sending at the OC-24 rate (you will have to look up what these two values are). Just as switch X finishes sending the packet payload, the switch Y receives the first bit of the IP packet, meaning the "width" of the packet is the length of the fiber. The speed of light in fiber is 2/3 the speed of light in a vacuum.

10. [4 points] What is the distance between switch X and switch Y?

OC-24 R=1.244 Gbps
Ethernet MTU L = 1500 B
transmission delay D = 1500B / 1.244 Gbps
distance = D(s) * 3*10^8 (m/s) * 2/3 =
2/3 * 3*10^8 (m/s)*(1500 B / 1.244 Gbps) = 1.92926045 km
Note: For Ethernet MTU, 1518 Byte (1500 bytes + Ethernet headers=1518 bytes) can also be used, which will lead to the answer =1.95241158 km.
Answers with reasonable round up in the result/calculation steps, or used more exact value for OC-24 or speed of light will also be accepted.
+1 correct equation
+1 correct numbers for OC-24
+1 correct use of the transmission delay formula
+1 correct answer

11. [2 points] Suppose the link is swapped to OC-3. Does the "width" of a packet on the wire become shorter or longer? Why?

Longer. OC-3 is 8 times lower transmission than OC-24, so it takes more time to send a packet (much higher transmission delay). The propagation speed is the same, but the "front" of the packet has more time to propagate so it will travel farther before the full packet is on the wire.
Note: students can also use some equations to explain the relationship.
+1 Answer answer (longer)
+1 Correct explanation