CSEE 4119 Fall 2022

Homework 1

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

Corrections:

2022-09-25: fixed typo

2022-09-23: none yet

You are welcome to discuss your homework with other students, but each student is expected to write his or her final answer independently and in his or her own words. Students are expected to consult whatever resources are necessary to answer homework questions,

including things like class papers, textbooks, other papers or RFCs, and the web.

However, as above, all answers are expected to be formulated independently in your own words. If you find you must use materials from other sources, standard academic policies apply---you need to cite what material you use and clearly indicate (such as by quotes) what material was yours and what is theirs. You are welcome to use calculators to solve equations, but for full credit your answer must include the equation you entered in the calculator.

You will submit the homework as a PDF file to the class [CourseWorks under Assignments](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://courseworks2.columbia.edu/courses/61434/assignments%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578300640%26amp;usg%3DAOvVaw0t6dcRWcAC6cDU8C4MXh5Q&sa=D&source=docs&ust=1664399578344638&usg=AOvVaw2v1s8LHpd_akicCDd-fuRN). The homework can be either typeset and converted to pdf, or it can be handwritten and scanned to pdf. Simple text should be sufficient for most questions and is recommended as it is generally easier to mark up and give feedback about. Please identify yourself at the top of the first page. Write your full name and your UNI. This will make it easier for the TAs to grade the homeworks and ensure that everything gets recorded properly.

Homework is due on the specified due date and time. Any submission past the deadline (even 1 minute late) will be subject to the late policy described in the syllabus. To avoid last minute problems (with a scanner, network, etc), please submit your homework well in advance.

Please refer to the course information in the syllabus for further information and policies regarding late homeworks and collaboration, and the academic honesty policies. Make sure that you observe these policies.

Please make sure you label the parts of your answer 1, 2, 3a, 3b, etc. Answers that don't clearly identify what part of the question they are answering, or that are overly long, may not get full credit. If you choose to handwrite your answers, please make sure your writing is clear. Make sure any mathematical notation is clearly explained. Further, unless we derived a formula in class, for full credit you must derive it in your homework in order to use it, including explaining what each term represents and WHY your equations correctly capture the problem. For example, in class, we used an equation to get the probability of collisions under packet switching, but we only briefly discussed the derivation. If you use this equation in your homework answers, you must derive it to get full credit.

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](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://forms.gle/RMd5orjQbe689vYM7%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578303770%26amp;usg%3DAOvVaw3eeMf2JwUocN55V3N4vokO&sa=D&source=docs&ust=1664399578345647&usg=AOvVaw27NKcTsTiFUJRfuncVDYpE) and answer all the questions. During the survey, you will be asked to perform two Internet speed tests on the [Speakeasy Internet Speed Test](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://www.speakeasy.net/speedtest/%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578304153%26amp;usg%3DAOvVaw29DAfr5NRPMUl9gmH3iWwV&sa=D&source=docs&ust=1664399578345829&usg=AOvVaw26vX7RxKAgB-opFjwcNoST), 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.
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.
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](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://fast.com/%2523%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578305345%26amp;usg%3DAOvVaw1fModG5aepx2-A16atImVH&sa=D&source=docs&ust=1664399578346252&usg=AOvVaw21BPc5_W3gkkA-vU9Yfayg), 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?
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!)
5. [1 point] What are some possible negative performance effects for users of ISPs having deep buffers?

Question 3: Protocol Layers and Service Models [16 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 is a main disadvantage of layered architecture? Please provide one answer to this question, whichever you see as the biggest disadvantage.
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? When 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?

1. [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?
2. [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?
3. [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 link layer} - {Number of times processed by link layer} - {Number of times processed by physical layer}

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](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://help.netflix.com/en/node/306%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578309539%26amp;usg%3DAOvVaw1aiQGCY_ax7xNqYZGCGCGn&sa=D&source=docs&ust=1664399578347648&usg=AOvVaw24-ryV0bocNebE8E1ZOZ5K).

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.

Netflix videos are (or at least were) encoded in 4 second chunks (see this [paper](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttp://web.mit.edu/6.829/www/currentsemester/papers/buffer_rate_adap.pdf%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578310905%26amp;usg%3DAOvVaw0ldY3nv78YfFDvHNjJlYEz&sa=D&source=docs&ust=1664399578348058&usg=AOvVaw2T4MZC1KmzrxrZg2TSzNZc)). 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](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://medium.com/netflix-techblog/per-title-encode-optimization-7e99442b62a2%26amp;sa%3DD%26amp;source%3Deditors%26amp;ust%3D1664399578311418%26amp;usg%3DAOvVaw0CcexZSGzf5eS8Cuq2cfHq&sa=D&source=docs&ust=1664399578348249&usg=AOvVaw3RQ6EGM4NmqmbJVdMO33mJ). For example, BoJack Horseman is simply drawn, so it begins streaming 1080p video at 1.5 Mbps.

1. [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 Mbps (assume that includes all headers and overhead). What is the download time for a 4 second chunk of video? Show the math.
2. [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)?
3. [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.
4. [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.
5. [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.
6. [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.
7. [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.

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

1. [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?
2. [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?
3. [2 points] Which mechanism (or behavior) is the root cause of the difference we see in d)? Explain in a few sentences.

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).

1. [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.
2. [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).
3. [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%

ii) Not segmented, delay at least 200%

1. [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.

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, 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.

1. [4 points] What is the distance between switch X and switch Y?
2. [2 points] Suppose the link is swapped to OC-3. Does the "width" of a packet on the wire become shorter or longer? Why?