Engin 112
Module 11 Homework
Due: 5pm, 9th December 2022

Note: *for the calculations that you are doing in your homework, assume that terms of kilo, mega, giga, etc., all refer to base-10 versions of these metric prefixes. That is, do not use the base-2 interpretations.*

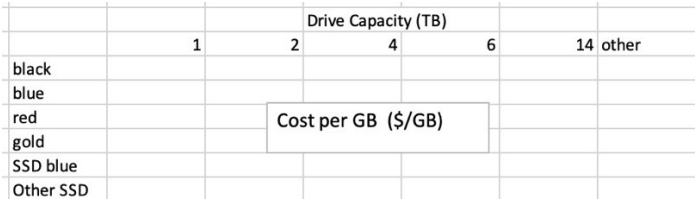
Question 1 (Electronic storage properties)
Examine an electronic device you have (such as laptop, smart phones, digital camera, etc.), and identify one memory or data storage component in these devices, and comment on them according to the following criteria:

a) Describe the particular memory or storage component.
b) Is it volatile or non-volatile?
c) Does it use sequential or random access approach?
d) Does it use floating gate transistors or regular transistors (or neither)?

Question 2 (Hard drive vs solid-state drive)
a) When you get a hard drive, you may see a lot of numbers on the specification sheets. Two of them may read, for example, 500 GB, 7200 rpm. What do these numbers mean?
b) Solid state drives (SSD) are a popular as a replacement for magnetic hard drives. Compare these two mediums in the following aspects: 1) mechanism; 2) volatility; 3) durability. Explain your answers briefly.

Question 3 (Cost of memory)
Hard drives these days come in many different sizes and flavors. For consistency, we will choose most of the drives to be from the same manufacturer: Western Digital. Western Digital has four classes of 3.5-inch drive: Red, Blue, Black and Gold. They also sell a 2.5-inch solid state drive (SSD). Most drive classes have capacities of 1, 2, 4, and 6 TB.

Use the internet to search out costs for the four sizes of the five different types of data storage. Add an additional class for finding the most inexpensive option for SSD, from any manufacturer. Combine the different sizes and costs into a table, as shown in the Figure below.



Once complete, using a spreadsheet program or something like Matlab/Python, make a plot of the cost per GB of memory, as a function of drive capacity, for the six different classes of data storage. Here, the horizontal axis should be the drive capacity, and the vertical axis should be $/GB. You should have a different line for each class of data storage (i.e. black, blue, etc.). You should label the horizontal and vertical axes as “drive capacity (TB)”, “cost per GB ($/GB)”, and

have a title, as "your name: Data Storage cost comparison (2020)". You should also use the labeling feature of your graphing program to specify which lines are associated with which drive class. When plotting, make sure that the horizontal axis is "linearly spaced", which is to say, that the spatial distance between 1 and 4 TB, should be four times larger than the spatial distance between 1 and 2 TB.

For your assignment, turn in a copy of the table (as above) and a copy of the plot. You should also provide a written comparison of the different classes of drives, on the advantages/disadvantages for each, and what you think is the reason for the differences in cost. Try to be as quantitative as possible (e.g. Cache-size, reliability, transfer rates, power consumption, etc.). In your table and/or description, it is okay to specify ranges (e.g. "The sustained transfer rate is 144-210 MB/s"). This doesn't have to be lengthy, just enough to provide the key details for each type of drive, and what the dominant reasons would be for the cost differences.
Based on the graphical analysis that you have done above, determine if there is an "inflection point" of hard drive capacity where for a given class of hard drive, the price per GB is the least expensive. This inflection point is typically the level where main-stream technology development is currently at. For the graphical analysis from above, determine the inflection point for the hard drives. Does it seem that there is a similar inflection point for solid state drives?

Question 4 (Cloud storage)
Storing data in the cloud is becoming increasingly common. Amazon Web Services (AWS) offers a number of ways of storing data of different volumes and different rates, usually dependent on how long it takes to retrieve data (Amazon Simple Storage Service, or Amazon S3). For this part of the homework, use the internet to look up the pricing for different classes of storage on S3, and compare these costs with those from the hard drive comparison in the previous problem.

For this part of the homework, you should also find information about the Amazon services known as "Snowball" and "Snowmobile". What are the purposes and costs for storage using these systems?

Question 5 (Data transfer rates)
Use an "internet speed test" program in your browser, or otherwise, to determine the download rate to your computer. Provide your answer in Megabytes per second (not megabits per second). Make a similar calculation for the upload rate.

Create and fill out a table with four columns and 9 rows. It should look like the following:

	Transfer rate	10 GB file transfer time	1 EB file transfer time
WD Black			
WD Blue			
WD Red			
WD Gold			
SSD			
Max SATA			

House download			
House upload			

In the table you will compare the amount of time derived from above, with the sustained data transfer rate to/from one of the disk drives that you analyzed in Problem 1, and other sources. As shown, you will have one column calculate amount of time it takes to transfer a 10 GB file (roughly 2 hours of a high-definition movie). Provide your answers in minutes. In another column, to transfer a 1 Exabyte of data ($1/10^{th}$ of the entire holdings of YouTube). In this case, use units of years instead of minutes.

Question 6 (Big data)

At one time or another, everyone has used GoogleEarth to look at images of the Earth from space, and to peer into the neighborhood that they grew up in. The resolution of GoogleEarth imagery can sometimes be as high as 1m x 1m.

Assuming that it takes 3 bytes of data for each pixel on GoogleEarth (one byte for each of the three colors: red, green and blue), and that we are only imaging over the Earth's land surfaces (1/3 of the total surface area), calculate the amount of memory (in PetaBytes, PB), that it will take to store one year's worth of color images of the Earth at 1m x 1m resolution, collected every day, for an entire year. Note, to do this problem, you will need to know the radius of the Earth (6371 km).