

Author: Andrea Fox

Title: Assignment 1

Subtitle: References

Data Sizes

Data Item	Size per Item	Reference
128 character message.	128 Bytes	https://www.unitconverters.net/data-storage/character-to-byte.htm Used calculator
1024x768 PNG image	1.1 MB	https://www.scantips.com/basics1d.html Used calculator
1024x768 RAW image	2.25 MB	https://www.scantips.com/basics1d.html Used calculator
HD (1080p) HEVC Video (15 minutes)	300MB	https://www.circlehd.com/blog/how-to-calculate-video-file-size Took the 5mbps bitrate * 20MB per 1 minute and got 300MB
HD (1080p) Uncompressed Video (15 minutes)	135,000MB	https://www.circlehd.com/blog/how-to-calculate-video-file-size 1080p(FHD) shows 5mbps with around 30FPS on common TV. $5 * 30 = 150$ mbps. Then found how many seconds are in 15 minutes which is 900seconds. I then took the $150\text{mbps} * 900\text{seconds} = 135,000\text{MB}$
4K UHD HEVC Video (15 minutes)	1,750MB	https://www.usatoday.com/story/tech/columnist/2017/12/10/youre-buying-4-k-tv-how-much-internet-bandwidth-do-you-need/933989001/ Basing off of Netflix estimates of 7GB per hour. Divide that by 4 for the quarter hour and I got 1.75GB. Then used a conversion calculator online to figure how many MB were in 1.75GB.
4k UHD Uncompressed Video (15 minutes)	156,430MB	https://www.digitalrebellion.com/webapps/videoalc Entered in the Uncompressed assuming 1080 8-bit for format with 30fps at 15 minutes.
Human Genome (Uncompressed)	200 GB	https://medium.com/precision-medicine/how-big-is-the-human-genome-e90caa3409b0 Mentioned in the article

Scaling

	Size	#HD	References
Daily Twitter Tweets (Uncompressed)	59.61GB	3 HD Calculated out to actually only use about 0.17TB of memory used. However, HDFS cannot have 1	https://www.internetlivestats.com/twitter-statistics/ Mentions 500 million tweets per day, github site said to assume each tweet is 128 characters. This brings us to about 64,000,000,000 bytes per day. I then found a calculator that took bytes to GB.

		hd must have at least 3.	
Daily Twitter Tweets (Snappy Compressed)	20.27GB	3HD	https://github.com/google/snappy Used 1.5 and then converted into a percentage which is roughly 66%. Then took the 59.61*66% and grabbed the difference.
Daily Instagram Photos	80,566.41GB	24HD	Took the 1.1MB per 1024x768 PNG photos and multiplied by 75,000,000, which is 75% of the estimated 100,000,000 per day. Then used calculator to take the MB to GB for storage. Then took that GB into TB which = 78.7TB * 3 for HDFS/10 to get HD
Daily YouTube Videos	843,776GB	248HD	Above I got 300MB for a 15 minute HD HEVC video. Took that * 4 = 1,200MB per hour * 500 hours = 600,000MB uploaded per minute. Then multiplied by 60 mins * 25 hours to get 864,000,000MB per day/1024/1024 = 824TB. Then converted TB to GB. For hard drive took the 824TB * 3/10 to get 247.2 HD so rounded up.
Yearly Twitter Tweets (Uncompressed)	21,757.65GB	7HD	Took the daily from above and multiplied by 365 days per year which gave me 21,757.65GB. I then used a convert calculator which gave me about 21.2477TB. I then took the TB and multiplied by 3 because it saves 3 copies which gives me 63.75TB. I then took that and divided by 10 which came to 6.375, which I rounded up to 7hd.
Yearly Twitter Tweets (Snappy Compressed)	7,398.55GB	3HD	Took the daily tweets compressed by snappy and multiplied by 365 to get yearly rate. Then took the yearly GB and used converter calculator to find TB, which comes to 7.2251 then multiplied by 3 for HDFS. This came to 21.67 then divided by 10 per HD came to 2.16.
Yearly Instagram Photos	28,717.52TB	8,616HD	Took daily number * 365 days, then converted the GB to TB because of the size of number. For hard drive I took the TB *3 for HDFS/10 to get my number of HD rounded up from 8,615.256
Yearly YouTube Videos	300,760TB	90,228HD	Took daily number * 365 days, then converted the GB to TB because of the size of number. For hard drive I took the TB *3 for HDFS/10 to get my number of HD.

Reliability

	#HD	# Failures	References
Twitter Tweets (Uncompressed)	7HD	0.07 if rounded a complete failure it would be 1	Multiplied my number of HD * 0.93% then rounded up
Twitter Tweets (Snappy Compressed)	3HD	0.03 if rounded to a complete failure number it would be 1	Multiplied my number of HD * 0.93% then rounded up
Instagram Photos	8,616HD	81	Multiplied my number of HD * 0.93% then rounded up
YouTube Videos	90,228HD	840	Multiplied my number of HD * 0.93% then rounded up

Latency

	One way Latency	References
Los Angeles to Amsterdam	70.32ms	https://wondernetwork.com/pings/Los%20Angeles/Amsterdam Found ping times from LA to Amsterdam and got an average from the min and max which gave me 140.63ms. Since a ping is roundtrip I then divided in half to arrive at 70.32ms of latency for my average based on the data from March 2021 pings.
Low Earth Orbit Satellite	20ms	https://www.omniaccess.com/leo/#:~:text=The%20GEO%20latency%20is%20of,an%20essential%20part%20if Under the section "What is LEO?" they talk about round trip latency based on the location of the orbit satellite. With a roundtrip latency of 40ms for LEO, I cut that in half for my one way measurement.
Geostationary Satellite	300ms	https://www.omniaccess.com/leo/#:~:text=The%20GEO%20latency%20is%20of,an%20essential%20part%20if Used the same section area as above and cut the latency in half to represent the one way. 600ms round trip becomes 300ms one way.
Earth to the Moon	1300ms	https://www.spaceacademy.net.au/spacelink/commldly.htm This listed out 1.3 seconds and then I found a calculator online to convert seconds to milliseconds.
Earth to Mars	180,000 to 1,260,000ms	https://www.spaceacademy.net.au/spacelink/commldly.htm This referenced 3-21 minutes from earth to Mars. I then converted minutes to milliseconds for the range.