---

title: Computer performance, reliability, and scalability calculation

author: Manish Kalkar

---

a. Data Sizes

Data Item	**Size per Item**	**Calculations / Assumptions**
128 character message	128 Bytes	UTF-8 format - 'a' character will use 1 byte
1024x768 PNG image	3.15 MB	Uncompressed 4x8bit RGBA
1024x768 RAW image	1.57 MB	Uncompressed 16bit monochrome
HD (1080p) HEVC Video (15 minutes)	1080 MB	Pixel:WxH=1920x1080=2.07Mpixels | 19.29KB/Mpixel x 2.07Mpixels=40KB/frame = 1200KB/30 Frames = 1.2MB/sec| File size: 900 seconds x 1.2MB/sec = 1080 MB
HD (1080p) Uncompressed Video (15 minutes)	3240 MB	Pixel:WxH=1920x1080x24bit(RGB)=6.22MB/frame | 19.29KB/Mpixel x 6.22 = 120KB/frame=3600KB/30 Frames=3.6MB/sec | File size:900sec x 3.6MB/sec=3240 MB
4K UHD HEVC Video (15 minutes)	1512 MB	Pixel:WxH=1920x1080=2.07Mpixels | 27.13KB/Mpixel x 2.07Mpixels = 56KB/frame = 1200KB/30 Frames=1.68MB/sec | File size: 900 seconds x 1.68MB/sec = 1512 MB
4k UHD Uncompressed Video (15 minutes)	4563 MB	Pixel:WxH=1920x1080x24bit(RGB)=6.22MB/frame | 27.13KB/Mpixel x 6.22 = 169KB/frame=5070KB/30 Frames=5.07MB/sec | File size:900secx5.07MB/sec=4563 MB
Human Genome (Uncompressed)	0.715 GB	No. of base pairs in the human genome = 3 Billion | each base pair takes 2 bits | 2 * 3 billion = 6,000,000,000 bits / 8 = 750,000,000 bytes / 1024 = 732,422 kilobytes / 1024 = 715 megabytes / 1000 = 0.715 GB

b. Scaling

Data Item	# Hard Disks	Size per Item	Calculations / Assumptions
Daily Twitter Tweets (Uncompressed)	1	192000 MB	Assuming - 128 characters/tweet and UTF-8 format - 'a' character = 1 byte. 500000000×128=64,000,000,000 Bytes | (x 3) 192,000,000,000 to store in HDFS - 192,000,000,000 B = 192000 MB = 192 GB = 0.192 TB
Daily Twitter Tweets (Snappy Compressed)	1	112941 MB	Snappy having a compression ratio of 1.5-1.7x for plain text. Assuming compression ratio of 1.7x - Snappy Compressed Size = 112941 MB
Daily Instagram Photos	83	826.5 TB	1 1024x768 PNG image = 3.15 MB. 75% of 100M = 75M PNG photos/Day. 75M × 3.15 = 236,250,000 MB PNG Photos. Assume rest of the photos are RAW images | 1 1024x768 RAW image = 1.57 MB. 25M RAW photos are equivalent to 25M×1.57=39,250,000. Total photo size=236,250,000+39,250,000=275,500,000 (x 3) 826,500,000 to store in HDFS = 826500 GB = 826.5 TB.
Daily YouTube Videos	2799	27,993.6TB	60 x 24 = 1440 min/day. So, 1440 x 500 = 720,000 hours of videos per day. Assume HD (1080p) Uncompressed YouTube Video with 30 frames per seconds. Pixel:WxH=1920x1080x24bit(RGB)=6.22MB/frame | 19.29KB/Mpixel x 6.22 = 120KB/frame=3600KB/30 Frames=3.6MB/sec. 3.6 x 3600 = 12,960 MB per hour. Total file size per day = 12,960 MB x 720,000 = 9,331,200,000 MB 9,331,200,000 MB = 9331200 GB = 9331.2 TB (x 3) for HDFS = 27,993.6 TB
Yearly Twitter Tweets (Uncompressed)	7	70.08 TB	Daily Tweeter Tweets (Uncompressed) = 192000 MB Yearly Tweeter Tweets (uncompressed) = 192000 x 365 = 70,080,000 MB 70,080,000 MB = 70080 GB = 70.08 TB
Yearly Twitter Tweets (Snappy Compressed)	4	41.2 TB	Daily Tweeter Tweets (Snappy Compressed) = 112941 MB Yearly Tweeter Tweets (Snappy Compressed) = 112941 x 365 = 41,223,465 MB 41,223,465 MB = 41223.5 GB = 41.20 TB
Yearly Instagram Photos	30,167	301,672 TB	Daily Instagram Photos = 826.5 TB. Yearly Instagram Photos = 826.5 x 365 = 301,672.5 TB
Yearly YouTube Videos	1,021,766	10,217,664TB	Daily YouTube Videos = 27,993.6 TB. Yearly YouTube Videos = 27,993.6 x 365 = 10,217,664 TB

c. Reliability

Data Item	# Hard Disks	# Failures	Calculations / Assumptions
Twitter Tweets (Uncompressed)	7	0.056	Assuming 0.85% Annual Failure Rate - 7 × 0.0085 = 0.0595
Twitter Tweets (Snappy Compressed)	4	0.034	Assuming 0.85% Annual Failure Rate - 4 × 0.0085 = 0.034
Instagram Photos	30,167	256	Assuming 0.85% Annual Failure Rate - 30,167 × 0.0085 = 256.4195
YouTube Videos	1,021,766	8,685	Assuming 0.85% Annual Failure Rate - 1,021,766 × 0.0085 = 8,685.011

d. Latency

Distance	One Way Latency	Reference
Los Angeles to Amsterdam	69.748 ms	https://wondernetwork.com/pings/Los%20Angeles/Amsterdam (RTT 139.496/2=69.748 ms
Low Earth Orbit Satellite	20 ms	https://www.omniaccess.com/leo/ (RTT 40/2 = 20 ms)
Geostationary Satellite	300 ms	https://www.omniaccess.com/leo/ (RTT 600/2 = 300 ms)
Earth to the Moon	1280 ms	https://en.wikipedia.org/wiki/Earth%E2%80%93Moon%E2%80%93Earth_communication(2.56/2=1.28s)
Earth to Mars	19.40 minutes	The latency is due to radio waves travelling at the speed of light. It depends on the orbit positions of the Mars and Earth. Depending upon their relative positions, it can take about 5 to 20 minutes for a signal to travel the distance between Mars and Earth. https://mars.nasa.gov/mars2020/spacecraft/rover/communications/ Based on current positions of the planets, it takes 19.40 minutes for the signal to reach Mars from Earth. Reference: https://interimm.org/comms-latency/en/