Must-know Estimations for a System Design Interview

Dinesh Varyani

Dinesh Varyani

Must know Back-of-the-envelope estimations

In a system design interview, the chances are high that you will be asked to estimate QPS (Queries per second) and Storage of the system using a back-of-the-envelope estimation.
There are various scales and numbers anyone appearing for System Design Interview should know. Some of the scales are as follows -

Number Scale

Name	Number	Number of Zeroes
1 hundred	100	2 Zeroes
1 thousand (K)	1000	3 Zeroes
1 million (M)	1000000	6 Zeroes
1 billion (B)	1000000000	9 Zeroes
1 trillion (T)	1000000000000	12 Zeroes
1 quadrillion	1000000000000000	15 Zeroes

Power of Two's Scale

In a system design interview, the volume of huge data is measured on the power of 2. It gets as low as to bits and bytes. A byte is measured as 8 bits. Estimations becomes easier if we co-relate below table with number table and make rough approximation. The interviewer will expect you to know these scales.

Name	Power	Value
1 KB (Kilobyte)	2^{10}	1024 ~ 1K
1 MB (Megabyte)	2^{20}	1048576 ~ 1M
1 GB (Gigabyte)	2^{30}	1073741824 ~ 1B
1 TB (Terabyte)	2^{40}	1099511627776 ~ 1T
1 PB (Petabyte)	2^{50}	1125899906842624 ~ 1 Quadrillion

Latency Numbers

In a system design interview, the latency numbers play a vital role in estimations and in having the knowledge, like how much time certain components take to perform certain operations. Below are some of the latency numbers of various operations -

Operation	Time taken
L1 cache reference	0.5 ns
Branch mispredict	5 ns
L2 cache reference	7 ns
Mutex lock/unlock	100 ns
Main memory reference	100 ns
Compress 1K bytes with Zippy	10,000 ns = 10 µs
Send 2K bytes over 1 Gbps network	20,000 ns = 20 µs

Latency Numbers

Operation	Time taken
Read 1 MB sequentially from memory	250,000 ns = 250 µs
Round trip within the same datacenter	500,000 ns = 500 µs
Disk seek	10,000,000 ns = 10 ms
Read 1 MB sequentially from the network	10,000,000 ns = 10 ms
Read 1 MB sequentially from disk	30,000,000 ns = 30 ms
Send packet CA (California) → Netherlands → CA	150,000,000 ns = 150 ms

***ns = nanosecond, µs = microsecond, ms = millisecond

1 ns = 10^-9 seconds
1 µs= 10^-6 seconds = 1,000 ns
1 ms = 10^-3 seconds = 1,000 µs = 1,000,000 ns

Availability Numbers

In a system design interview, the **High Availability** discussion will happen for sure. It is define as the ability of the system to be operational for a longer period of time. Below are some of the availability numbers you should know -

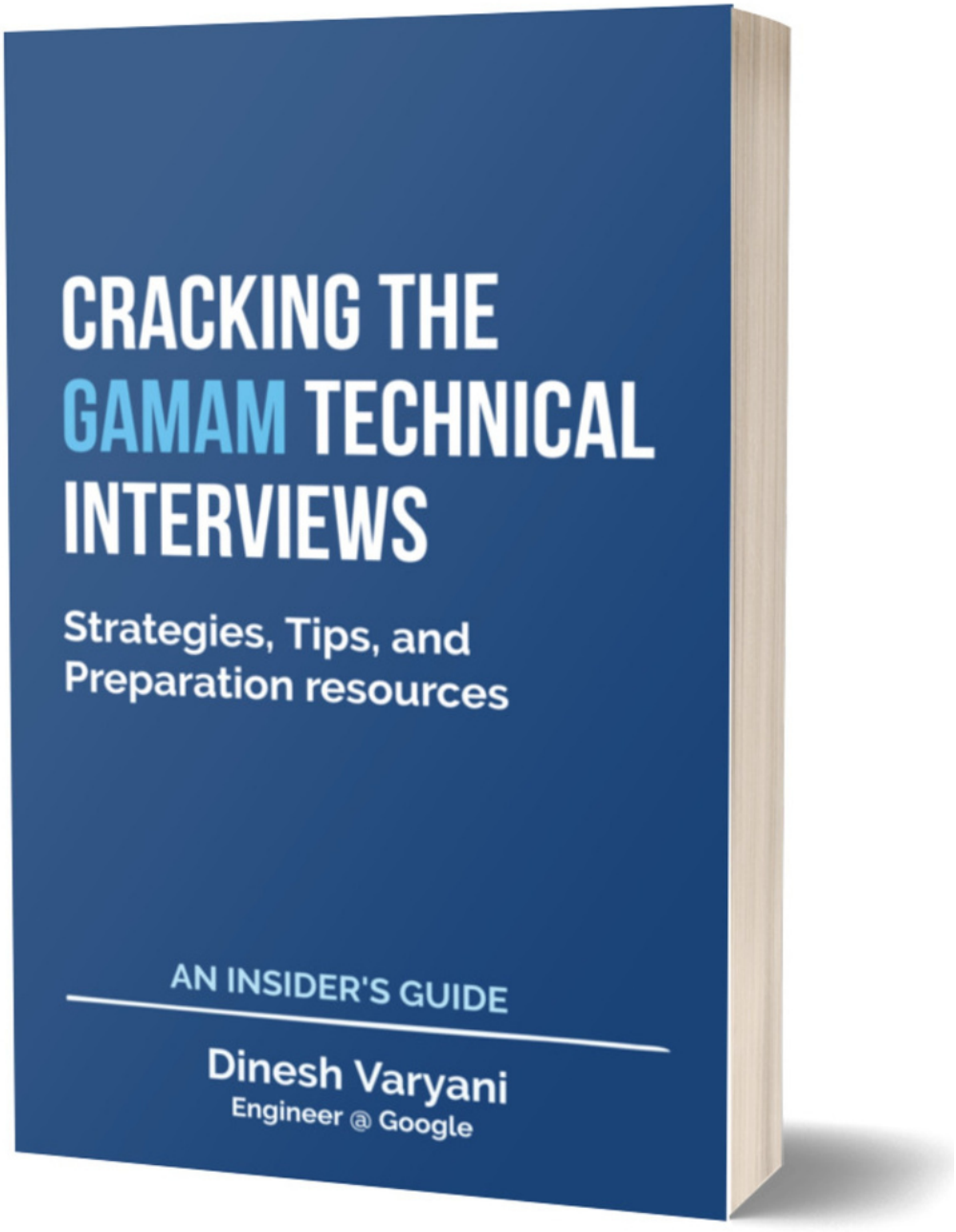
Availability %	Downtime per year	Downtime per month	Downtime per day
90% (one nine)	36.53 days	73.05 hours	2.40 hours
99% (two nines)	3.65 days	7.31 hours	14.40 mins
99.9% (three nines)	8.77 hours	43.83 mins	1.44 mins
99.99% (four nines)	52.60 mins	4.38 mins	8.64 secs
99.999% (five nines)	5.26 mins	26.30 secs	864.00 millisecs
99.9999% (six nines)	31.56 secs	2.63 secs	86.40 millisecs
99.99999% (seven nines)	3.16 secs	262.98 millisecs	8.64 millisecs
99.999999% (eight nines)	315.58 millisecs	26.30 millisecs	864.00 microsecs
99.9999999% (nine nines)	31.56 millisecs	2.63 millisecs	86.40 microsecs

Blob/Object Storage Sizes

In a system design, there are various big systems that involve types of blobs/objects like images, videos, audios etc. Below are some of the approximate storage sizes of various blobs/objects -

Object Type	Size
char	1 B
char (Unicode)	2 B
UUID	16 B
Thumbnails	20-30 KB
Website image	200-300 KB
Mobile image	2-3 MB
Documents like books, reports, govt Ids etc	1-3 MB
Audio files like songs, recordings etc	4-5 MB
1 min 720px video	60 MB
1 min 1080px video	130 MB
1 min 4K video	350 MB

Cracking the GAMAM Technical Interviews



Buy Now

THANK YOU !!!
For more such content follow
@Dinesh Varyani

Dinesh Varyani

Dinesh Varyani