

Q1: Assume that a 50000 server cluster is built from reliable machines with a mean=-time between failure (MTBF) of 2500 days. What is the failure rate?

Failure rate calculation:

$50,000 \times (1/2,500) = 20$ failures per day.

If one failure a day means 100% failure, what is the percentage?

$(1/20) \times 100 = 5\%$

Q2: What are the four Vs in Big Data?

Volume – The large amount of data.

Velocity – The speed at which data is generated and processed.

Variety – The different types and formats of data (structured, unstructured, etc.).

Veracity – The quality and reliability of the data.