

① Volume — Data Size.

② Variety — Structured, Semi-Structured, un-structured.

③ Velocity. — Batch, Near-Time, Real-time, Stream

④ Variety —

⑤ Value

Day 1



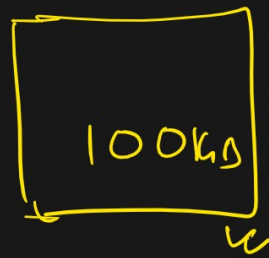
X

Approach 1

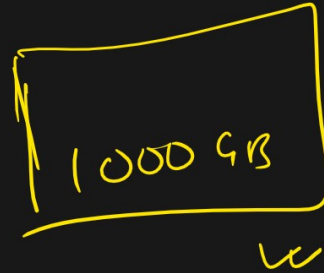


Scale-up

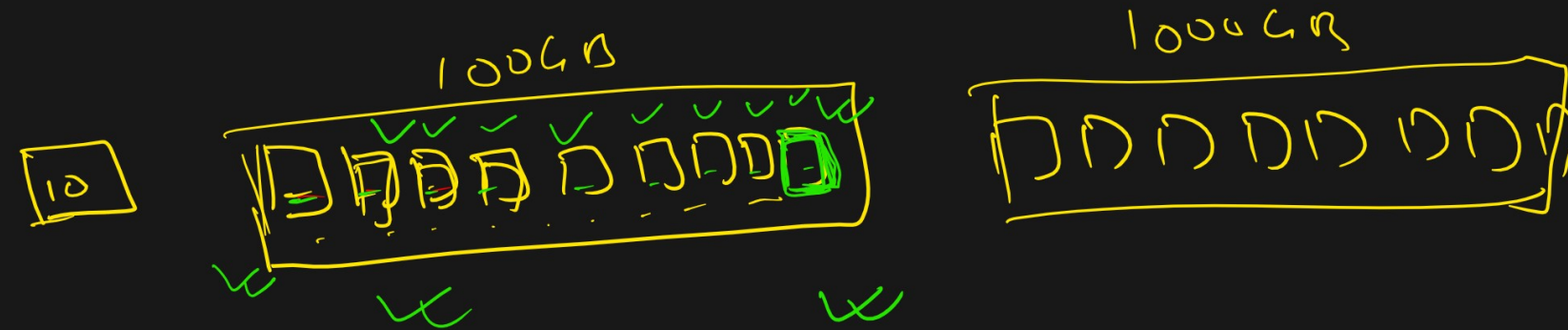
Day 2 →



Day 3 →



Approach - 2 → Scale out



HDFS → Hadoop distributed File System.

Master - Slave

