

1/6  
Volume  
Variety  
Velocity

Flipkart

IBM

Veracity  
Value

Log Files - Unstructured / 2TB

Transactions - Structured / ~GB

Catalog - Semi-Structured / 200GB

Social Media - Semi, Un / 3-4TB

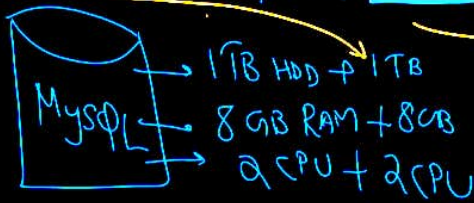
Logistics - / ~GBs

Internal - / TB

Google - 1999 - Distributed Computing

Scale-up / Vertical scaling

Scale-out / Horizontal Scaling



Dist. Comp | Framework

Google File System (GFS)  
Distributed

MapReduce  
Cluster



Doug  
Mike } Hadoop

