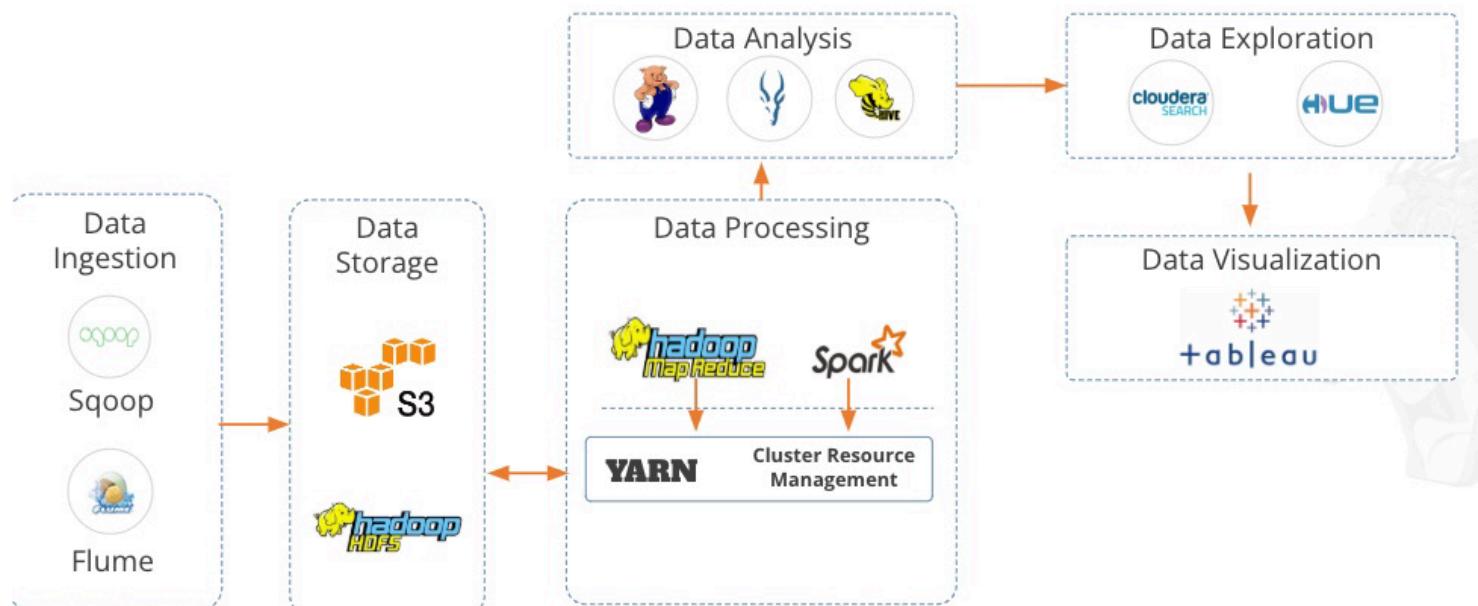
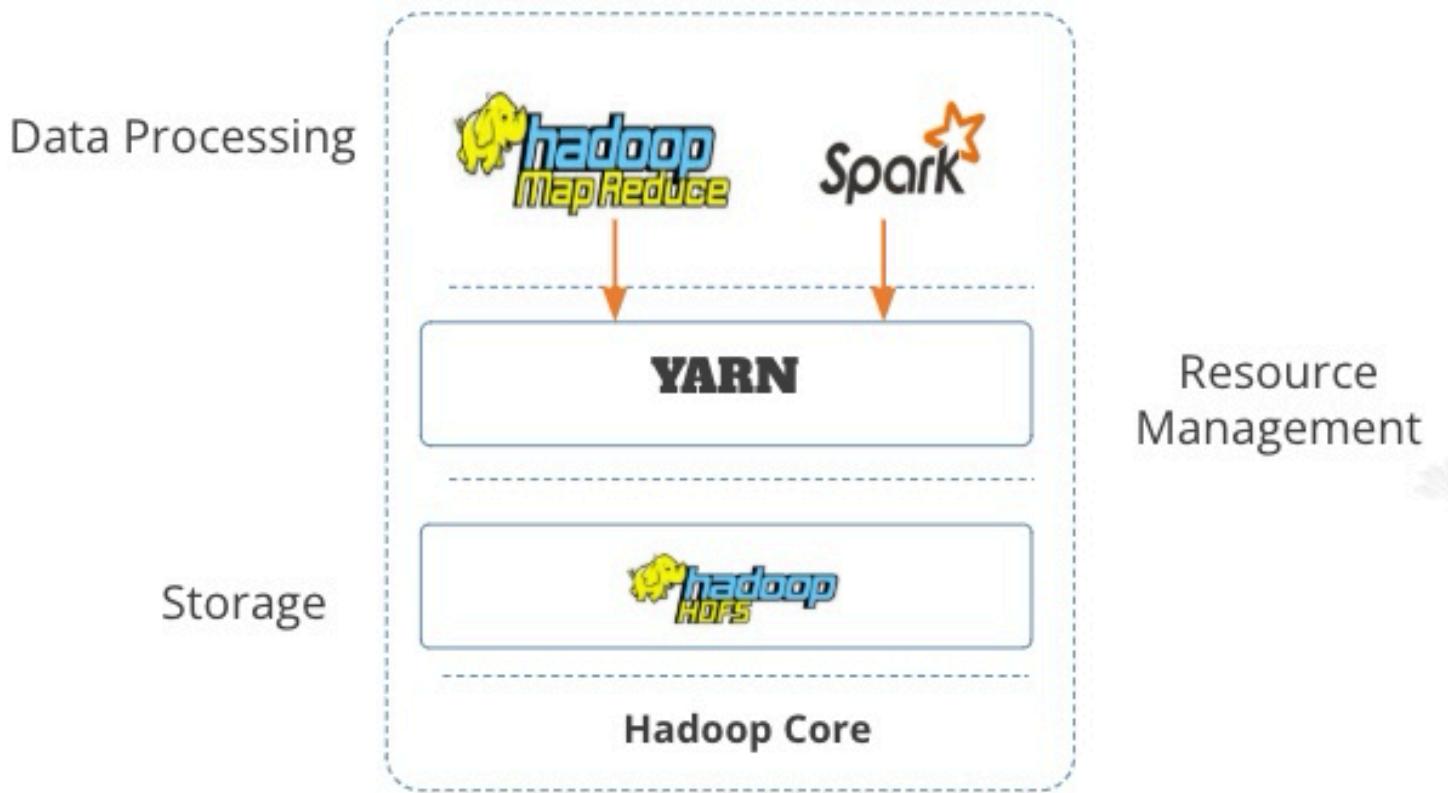
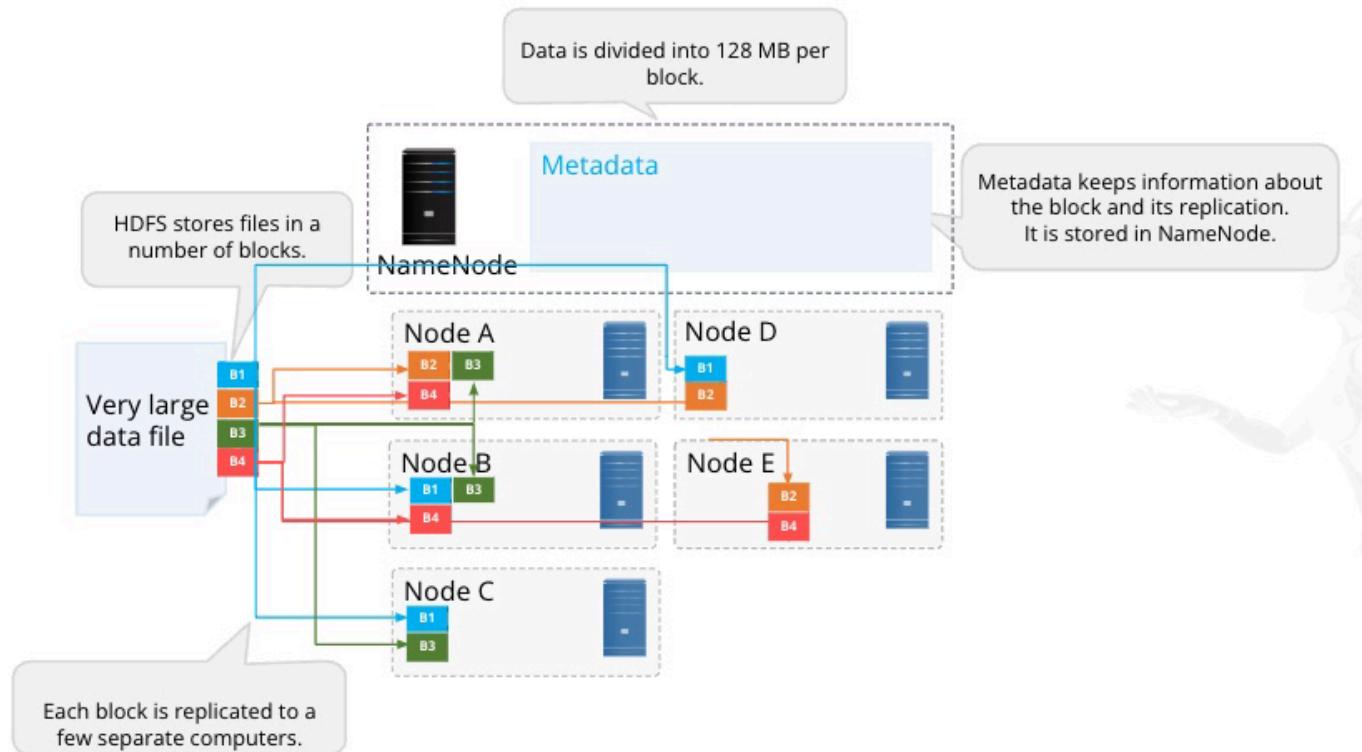
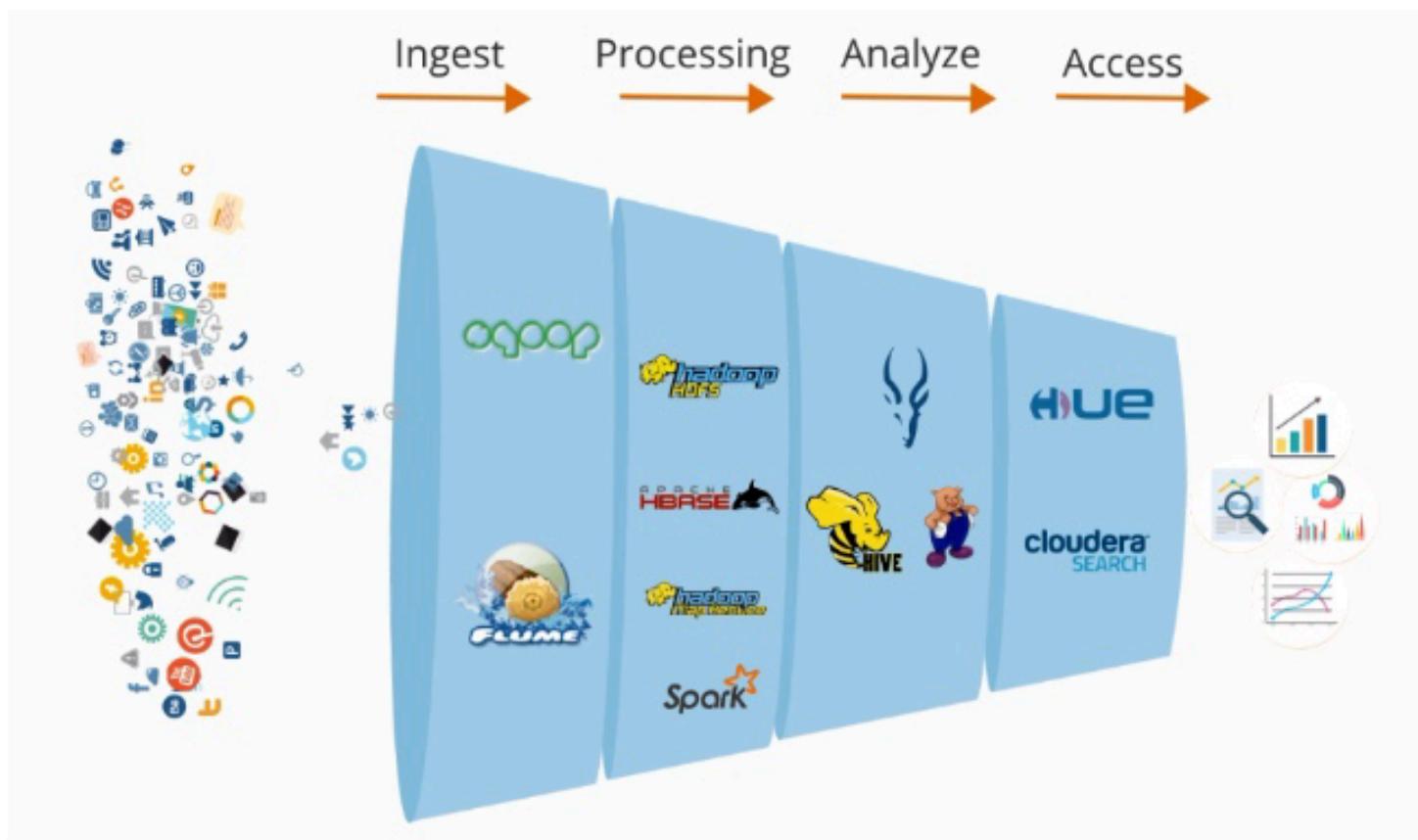


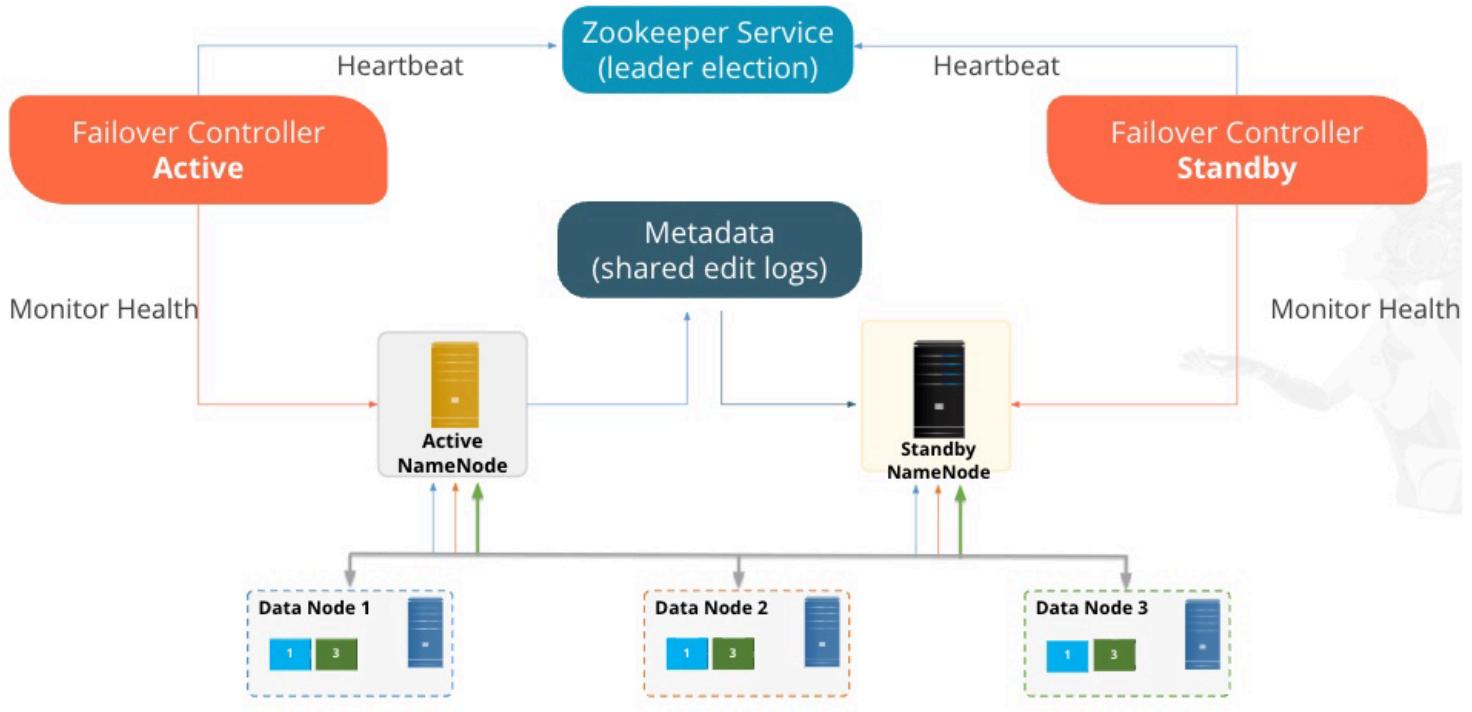
Hadoop Ecosystem Fundamentals Of Distributed Systems

- Intro to Hadoop
- Traditional system vs Hadoop
- Hadoop Core System
- Components of Hadoop system
- Big data processing
- HDFS
- YARN

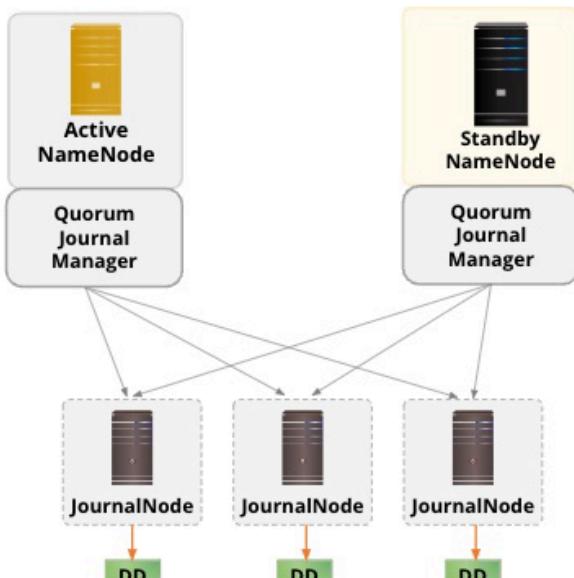




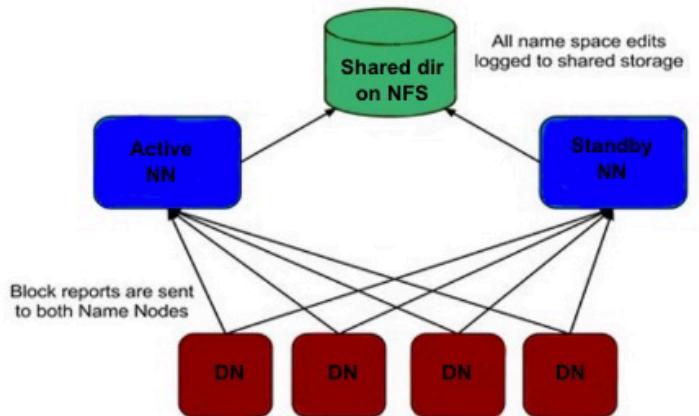




Quorum-based Storage

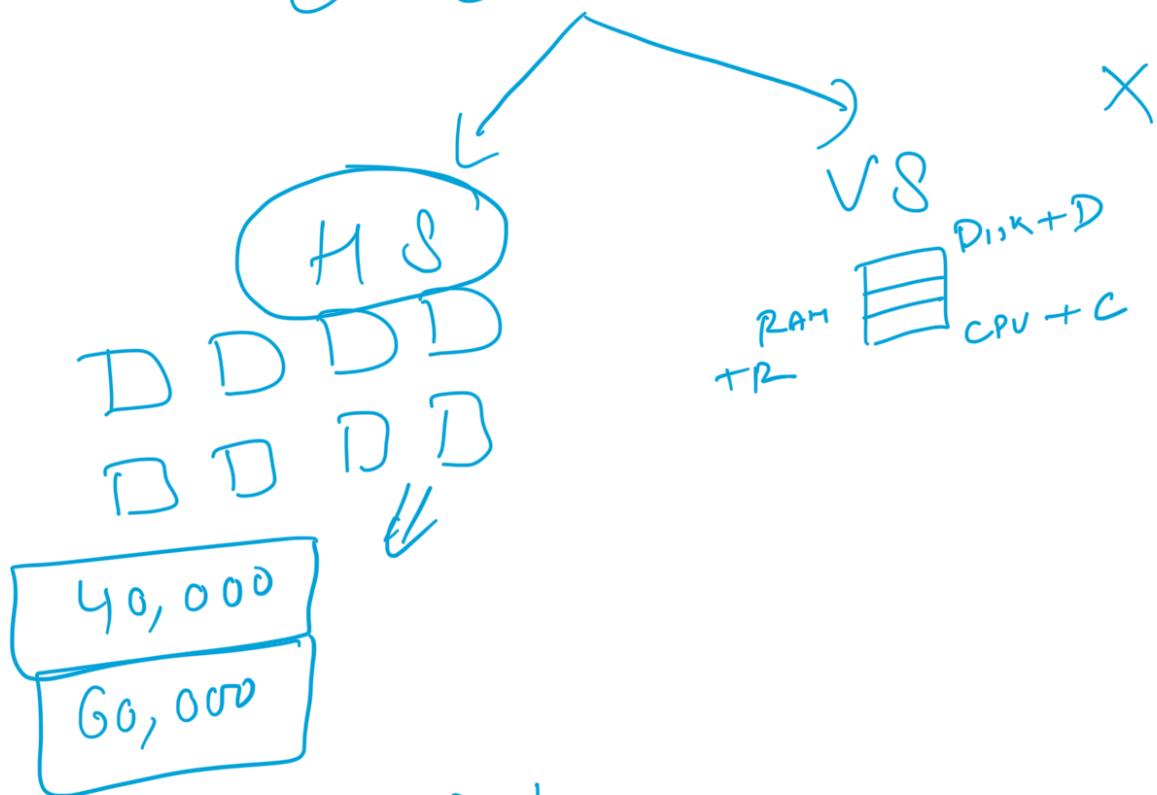


Shared storage using NFS



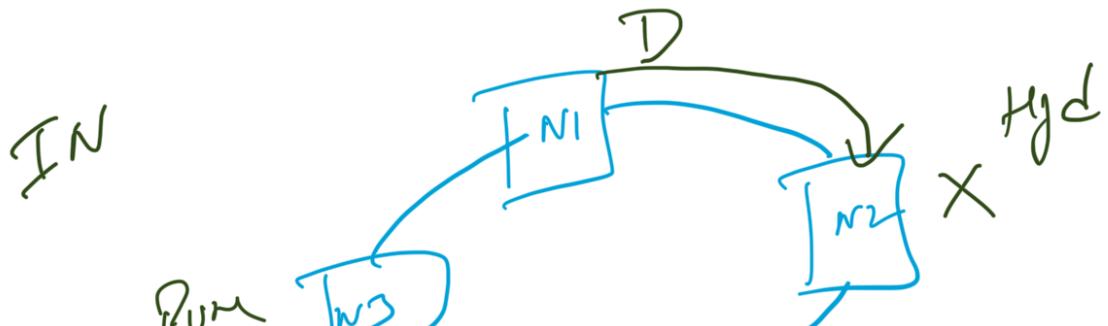
Challenges of traditional System

- ① Data Size :- EB / PB / HB
- ② Unstructured Data :- Video / Image / PDF
- ③ Scalability



Distributed System

↳ it is a model in which components located on networked systems communicate and coordinate their action via passing msgs.



IT
you
Ch

Challenges :-

- ① System Failure
- ② Limited Bandwidth
- ③ High Programming Complexity

\Rightarrow Hadoop

\Rightarrow Doug Cutting

Hadoop is a framework that allows distributed processing of large data sets across clusters of commodity computers using simple programming models.

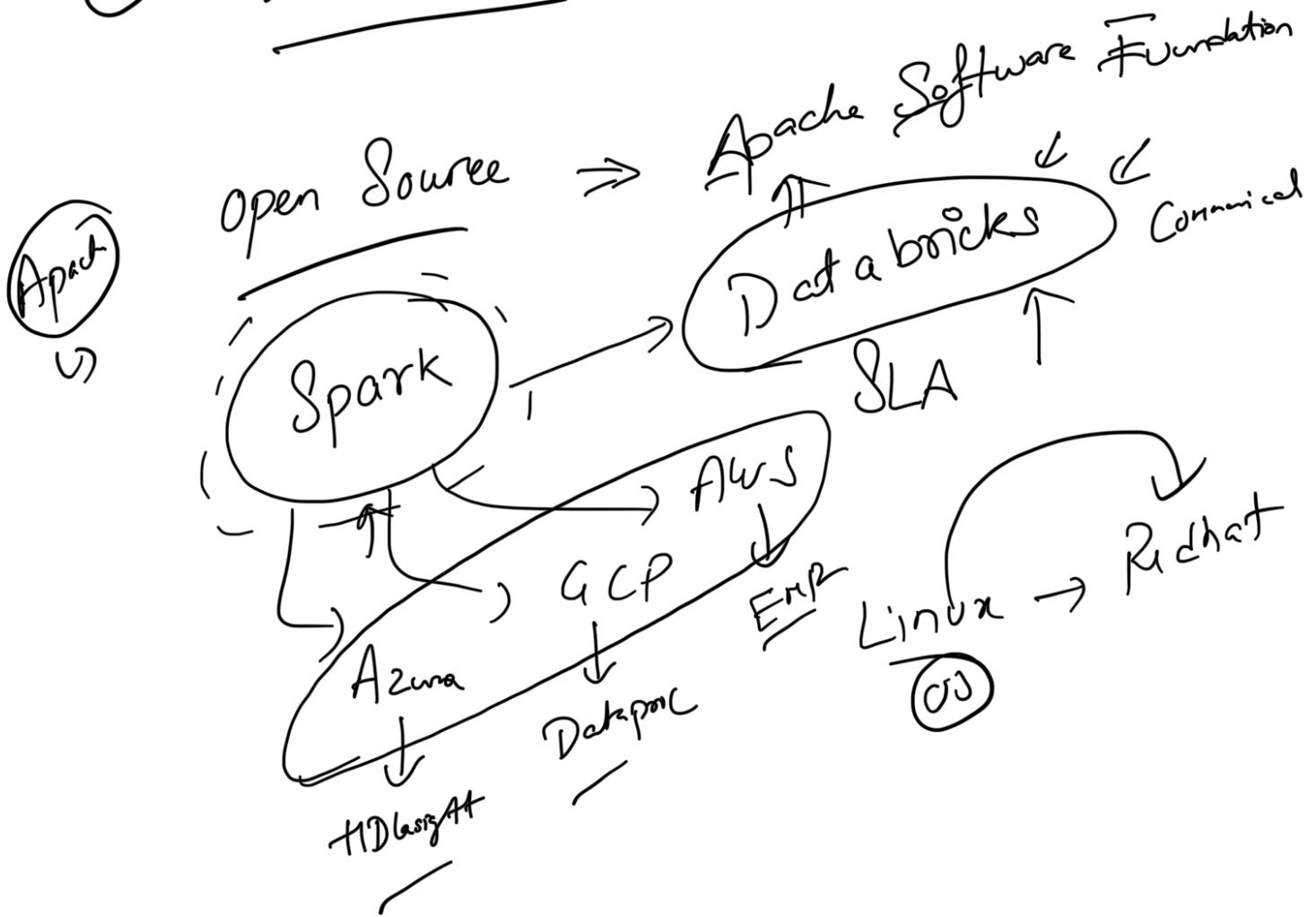
- ① System Failure :- Data on multiple servers
- ② Scalable :- $HIS + VS$

$$1. x = 4,000$$

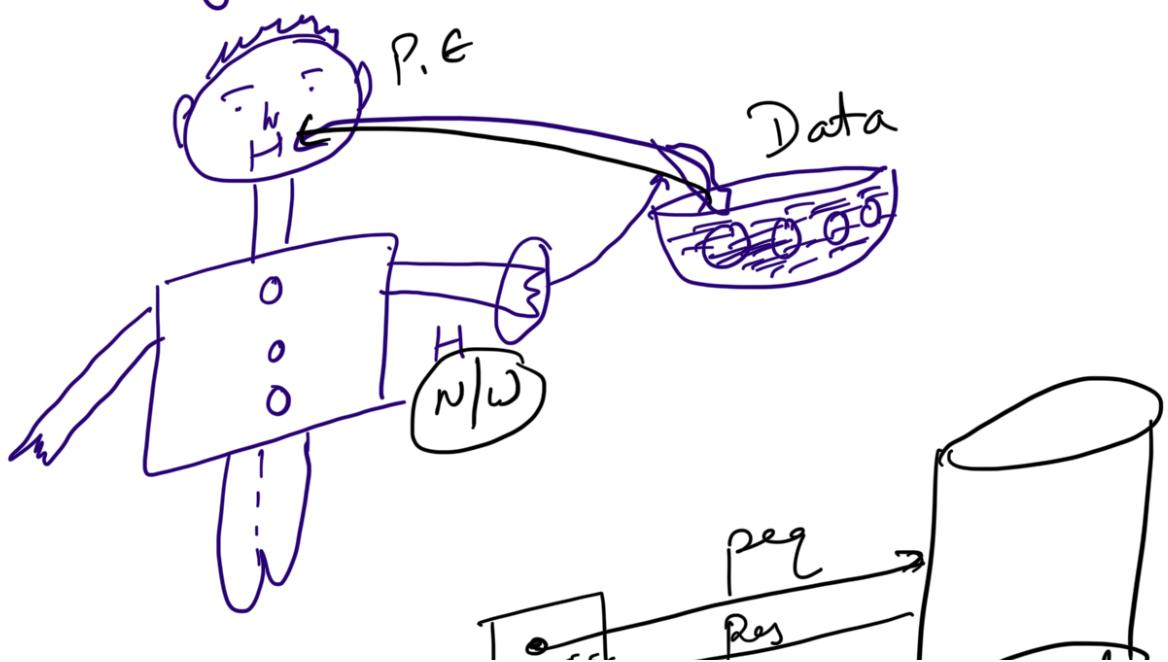
$$\begin{aligned} 2 \cdot n &\rightarrow 12020 \\ 3 \cdot n &\rightarrow \text{No Unit} \end{aligned}$$

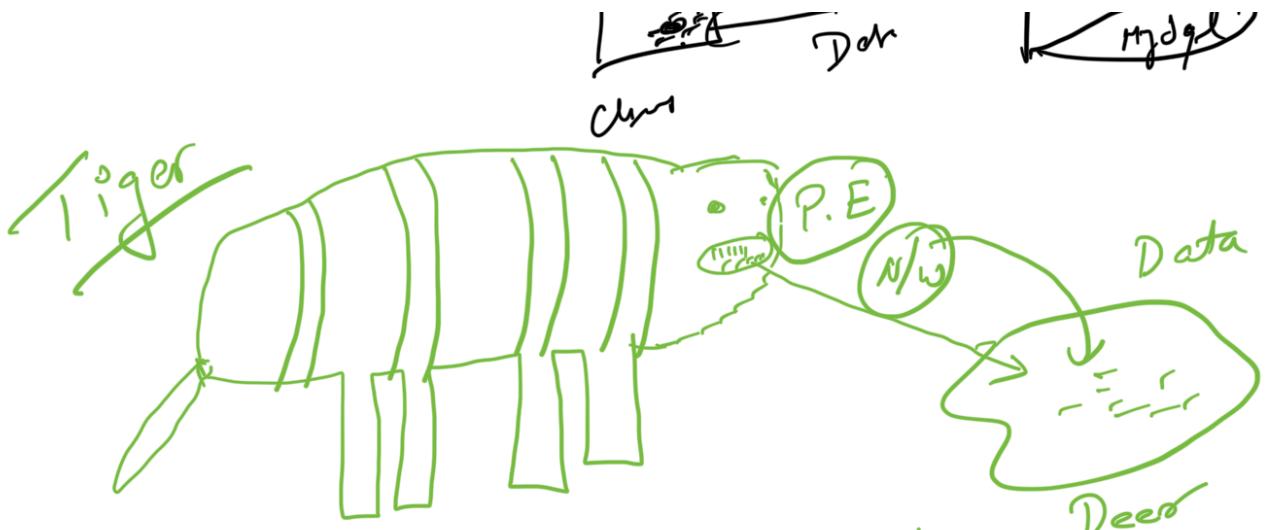
③

Economical :-



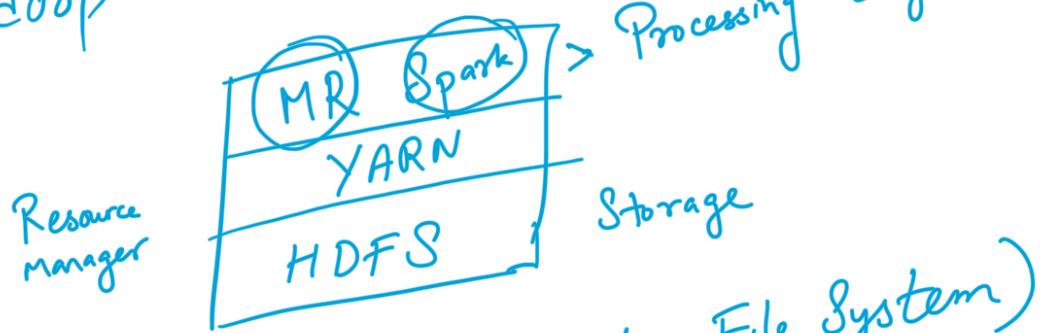
Sujeet



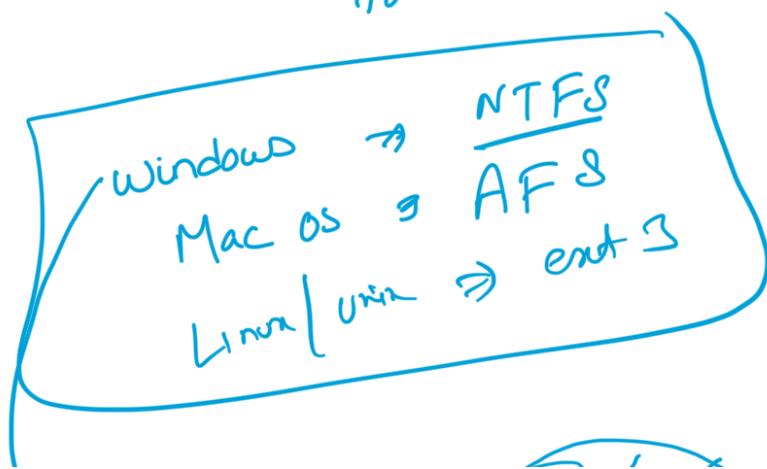


P.E. = Jar | query file
Data

Hadoop Core Components



① HDFS (Hadoop Distributed File System)
↳ is a DFS that provides the access to data across hadoop clusters.



C:
D: [abc/movies]
abc.txt

Resume.docx
61MB

