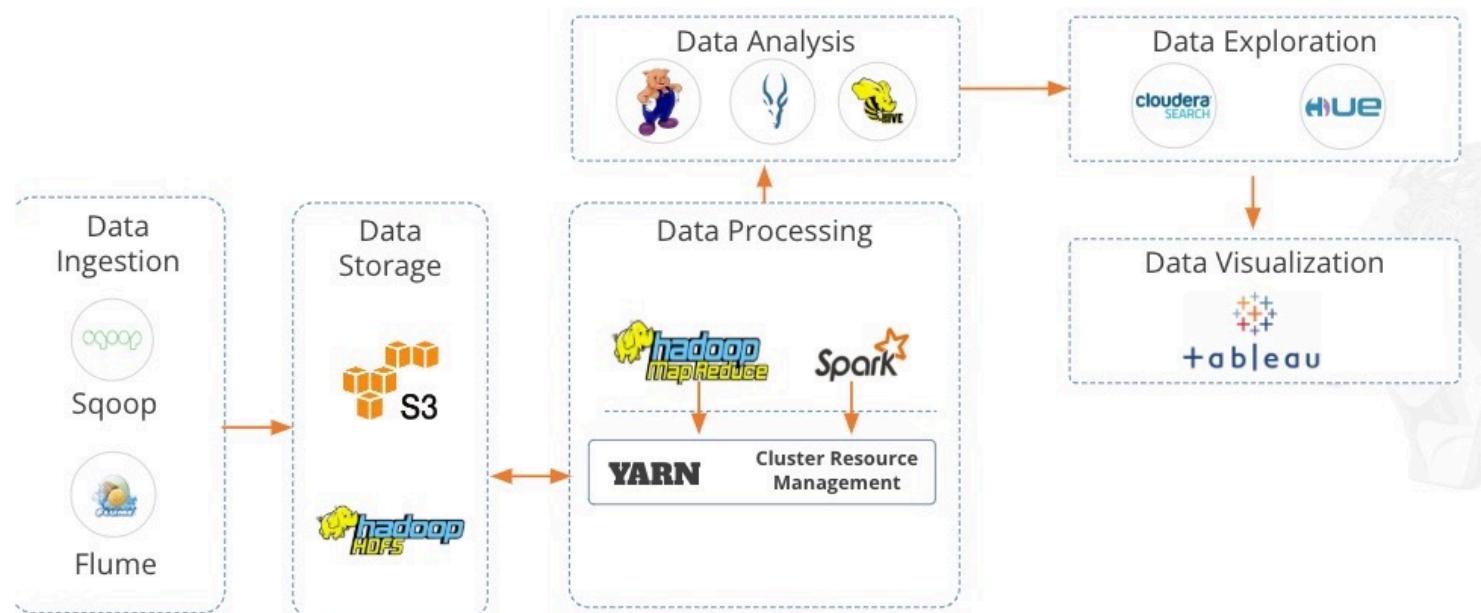
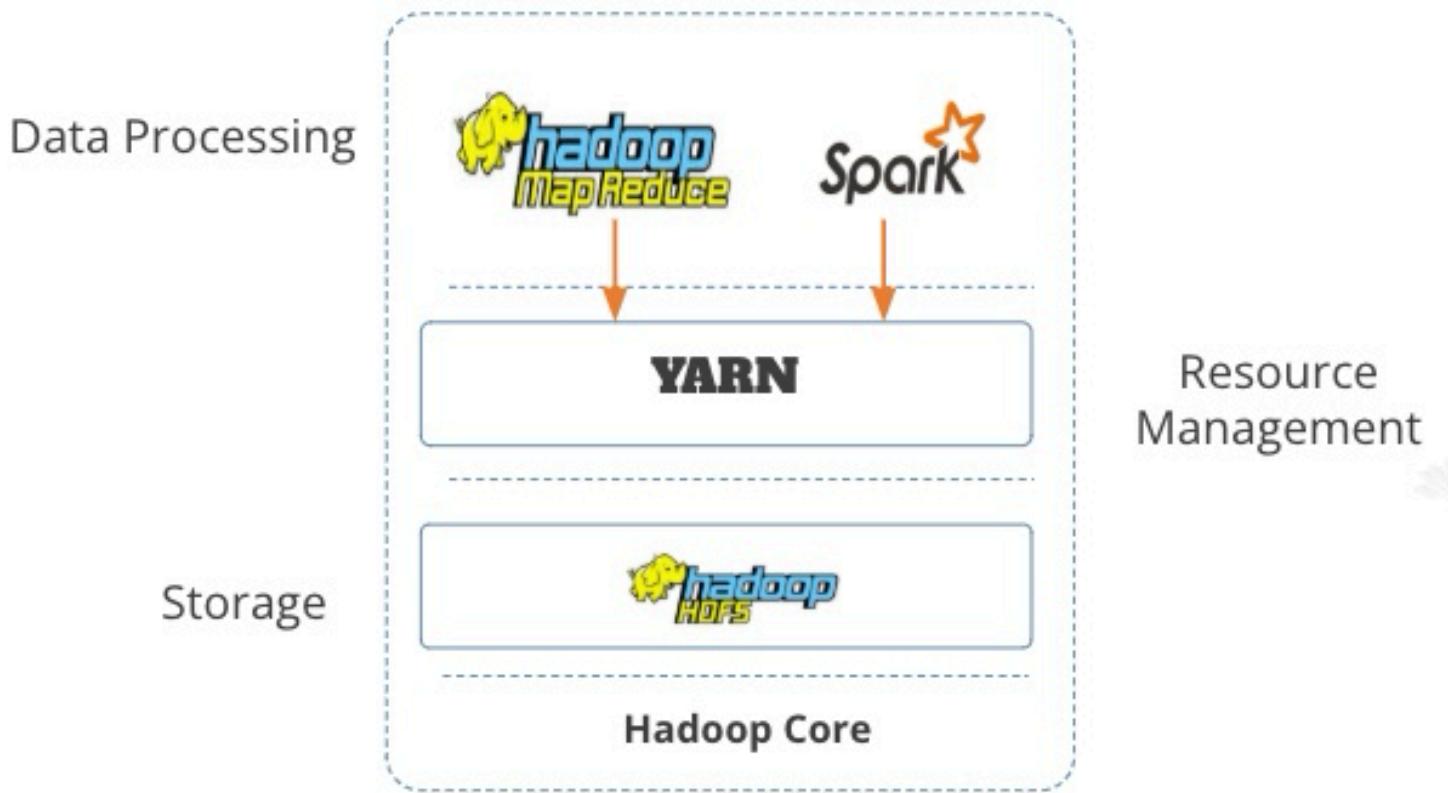
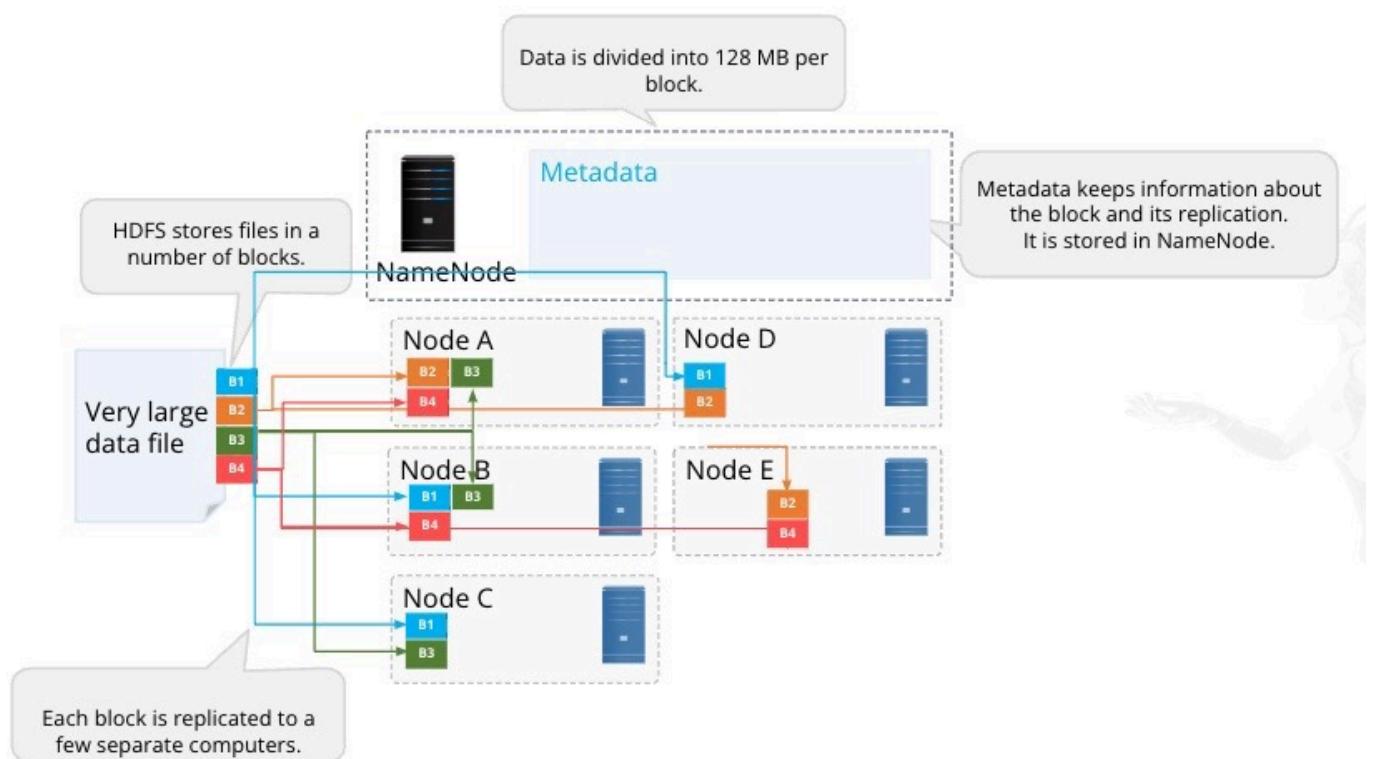
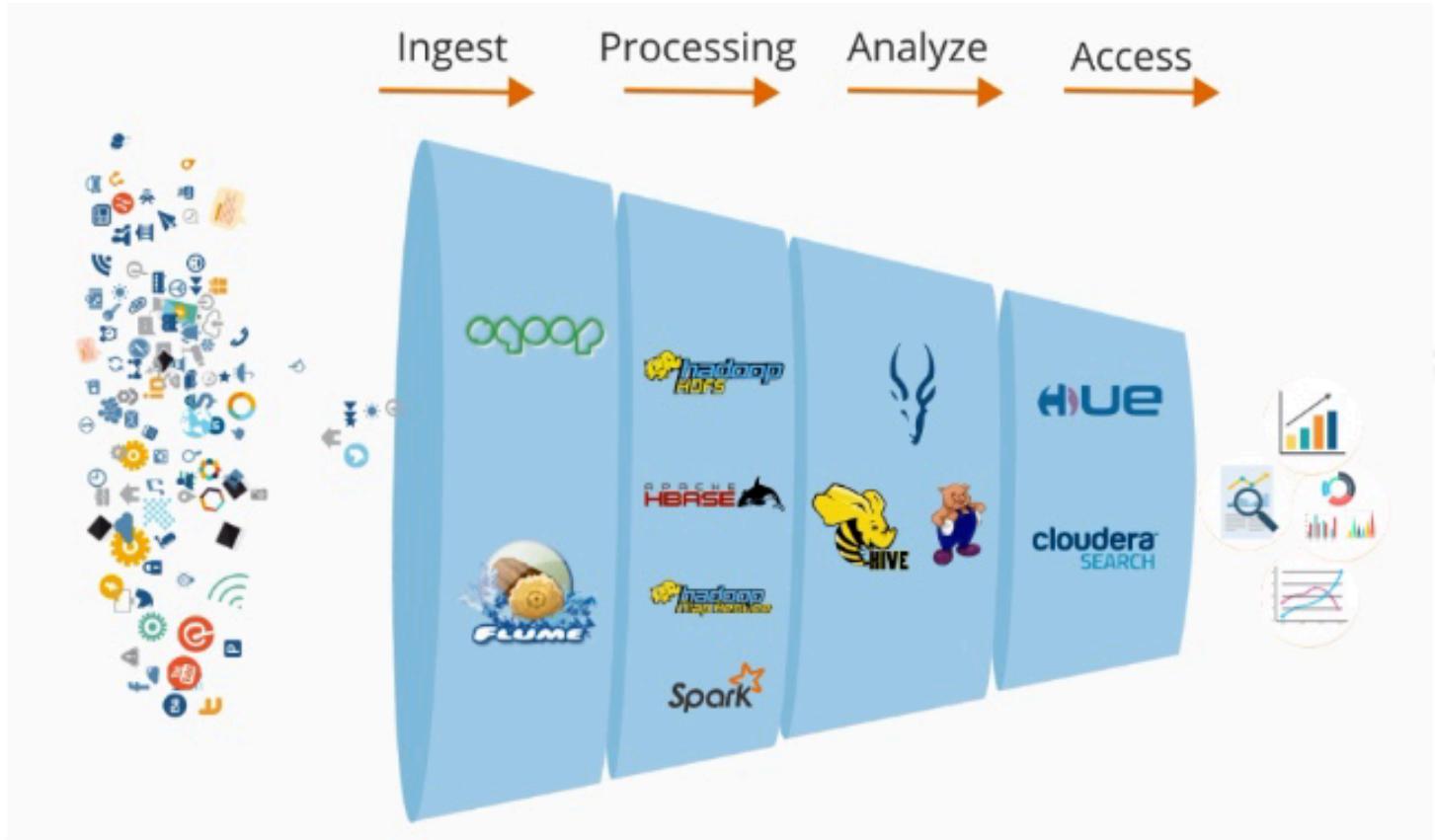


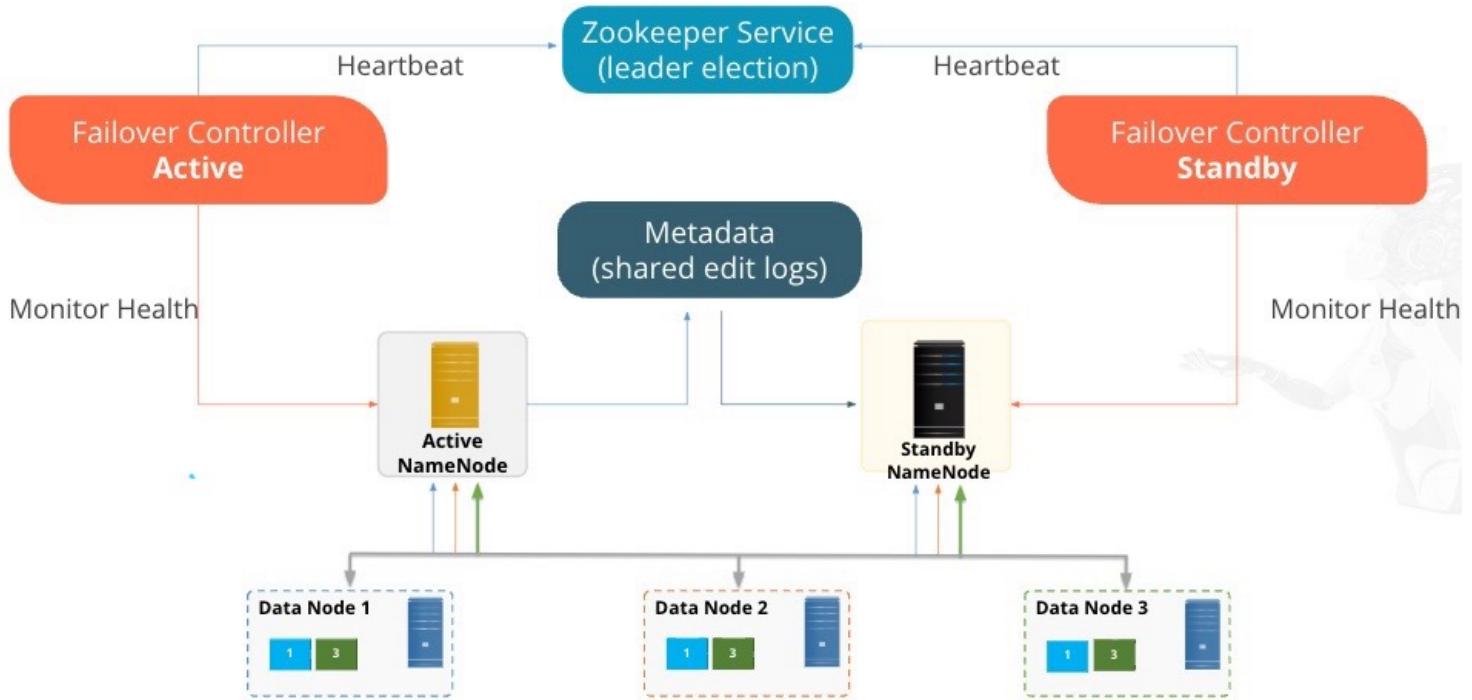
Hadoop Ecosystem Fundamentals Of Distributed Systems

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

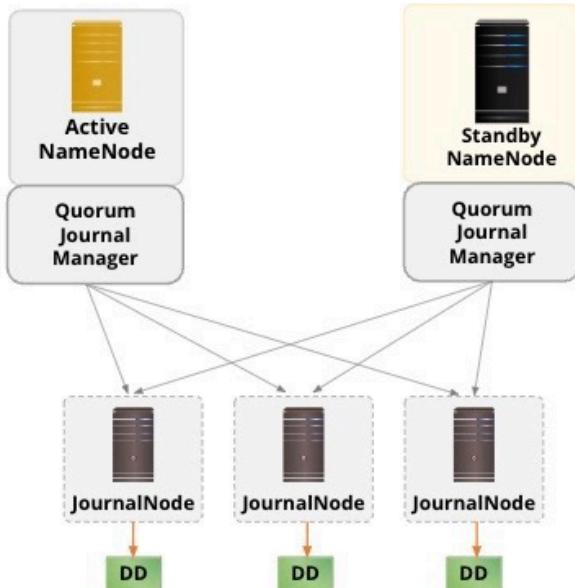




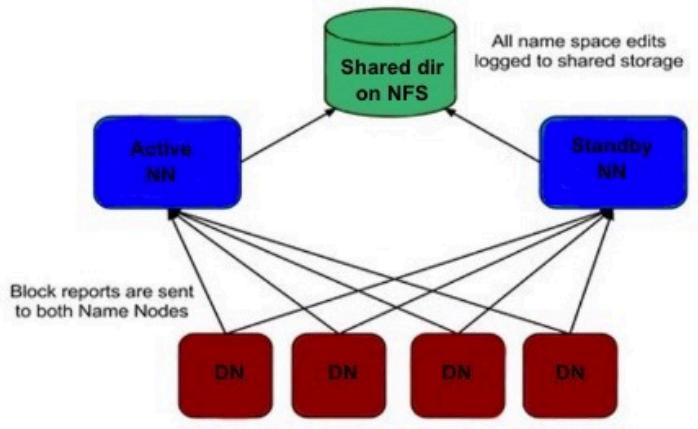




Quorum-based Storage



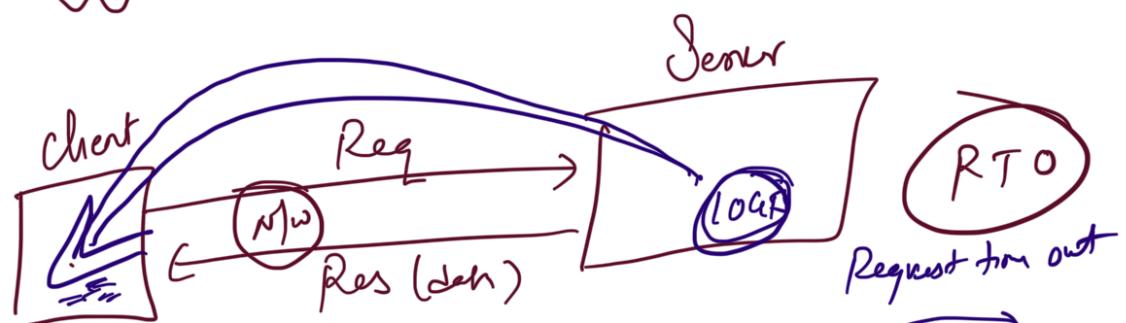
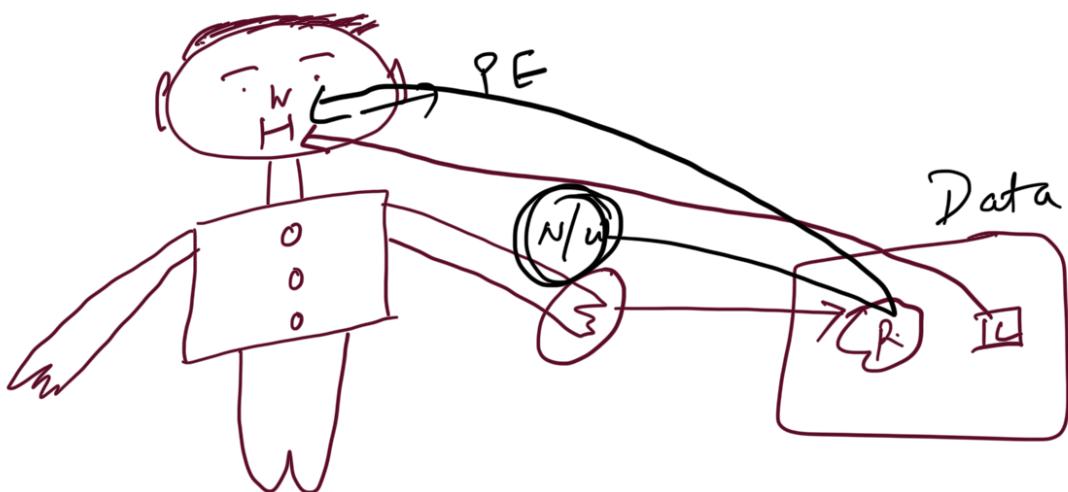
Shared storage using NFS



Hadoop

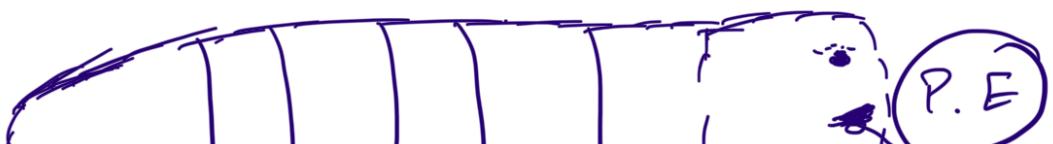
→ It is a framework that allows distributed processing of large datasets across commodity computers using simple programming models.

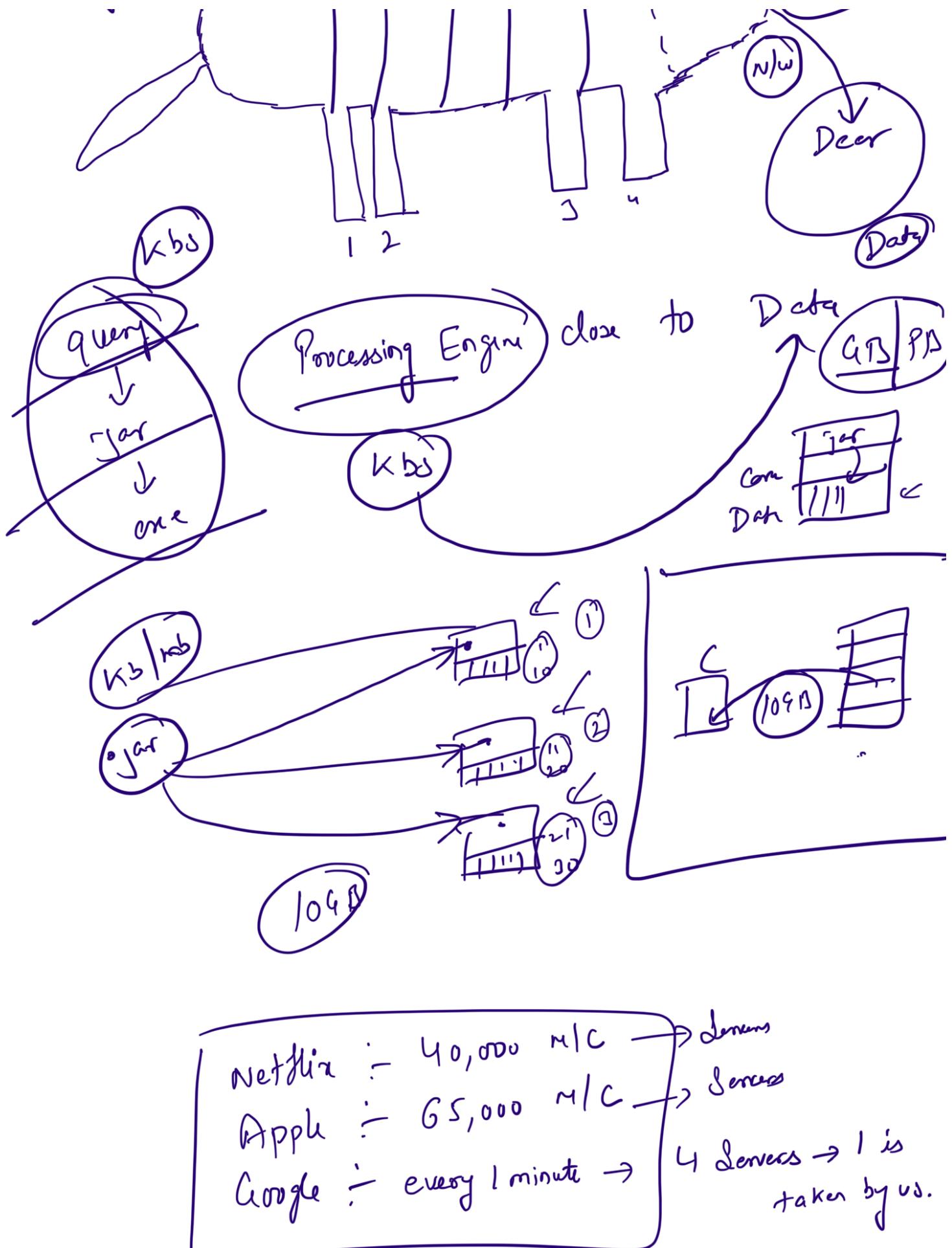
Doug Cutting



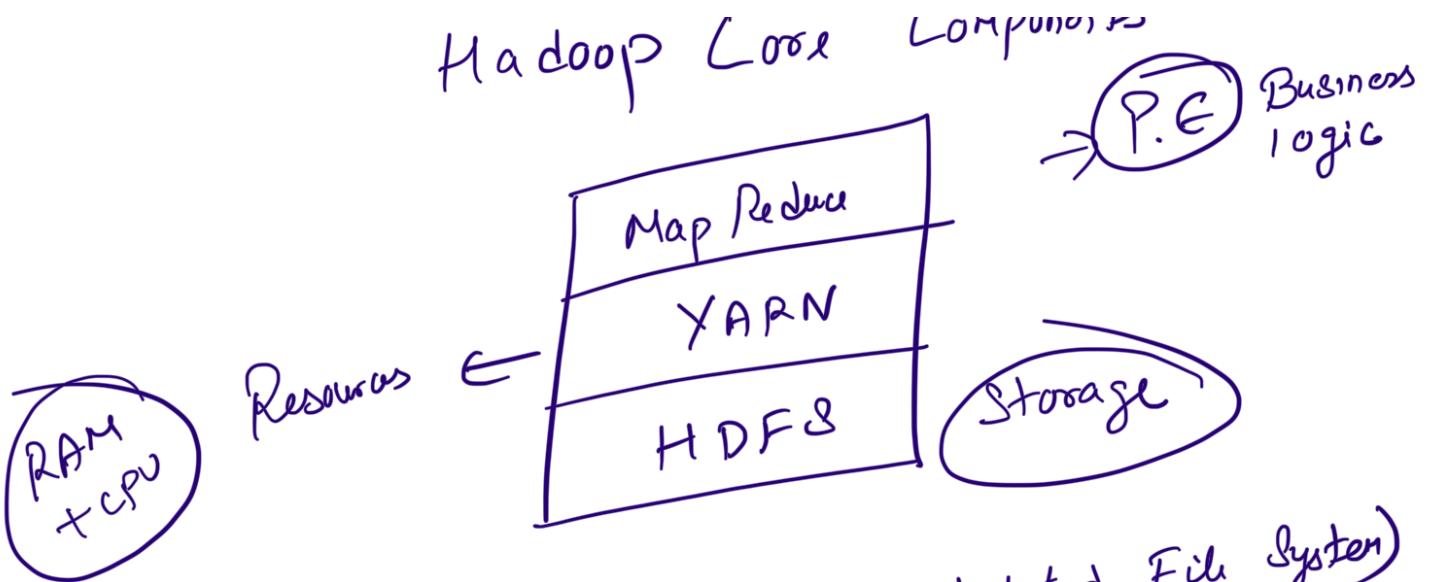
Data is coming over Network.

Tiger



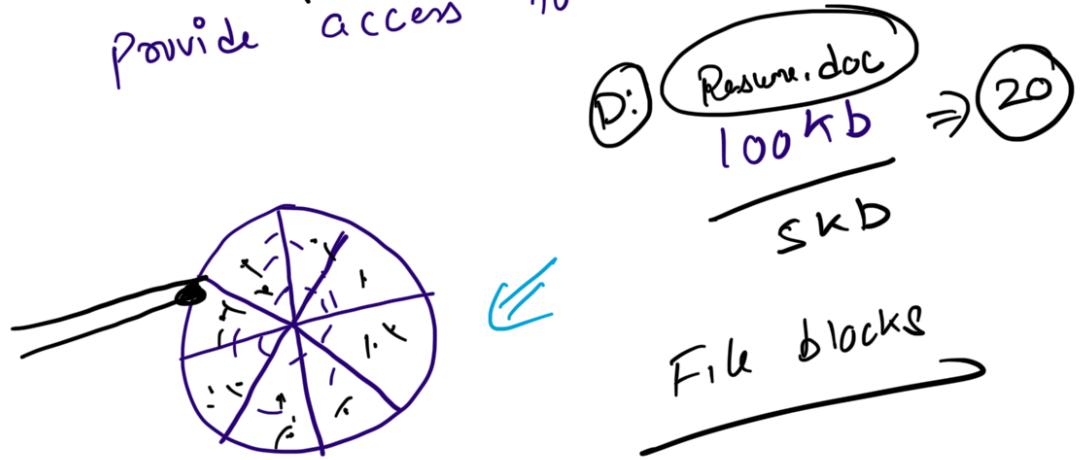


n ... ments



HDFS (Hadoop Distributed File System)

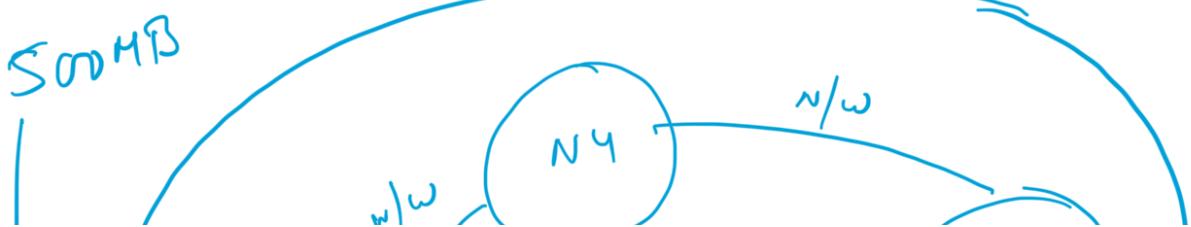
↳ is a distributed file system that provide access to data across hadoop.

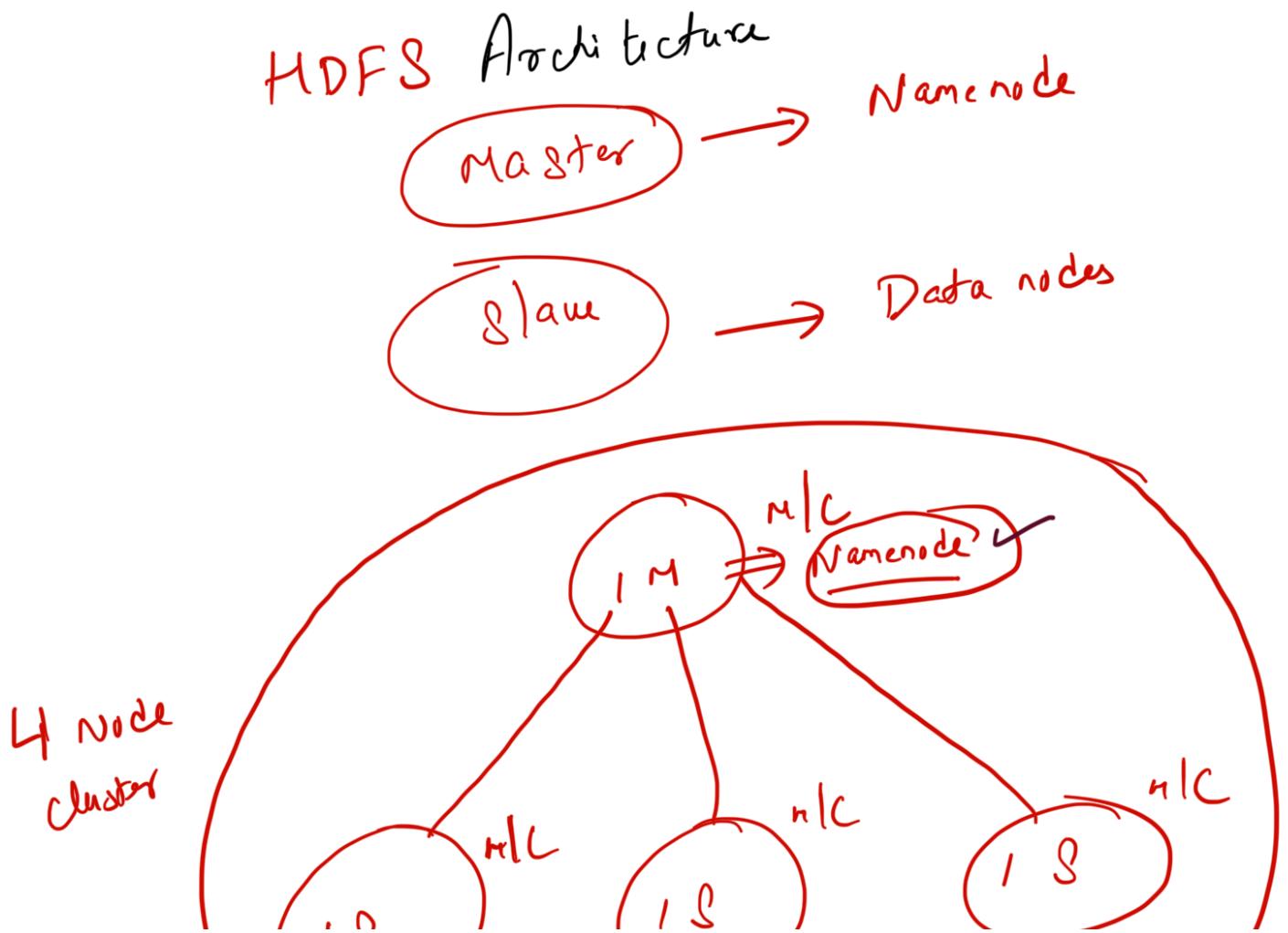
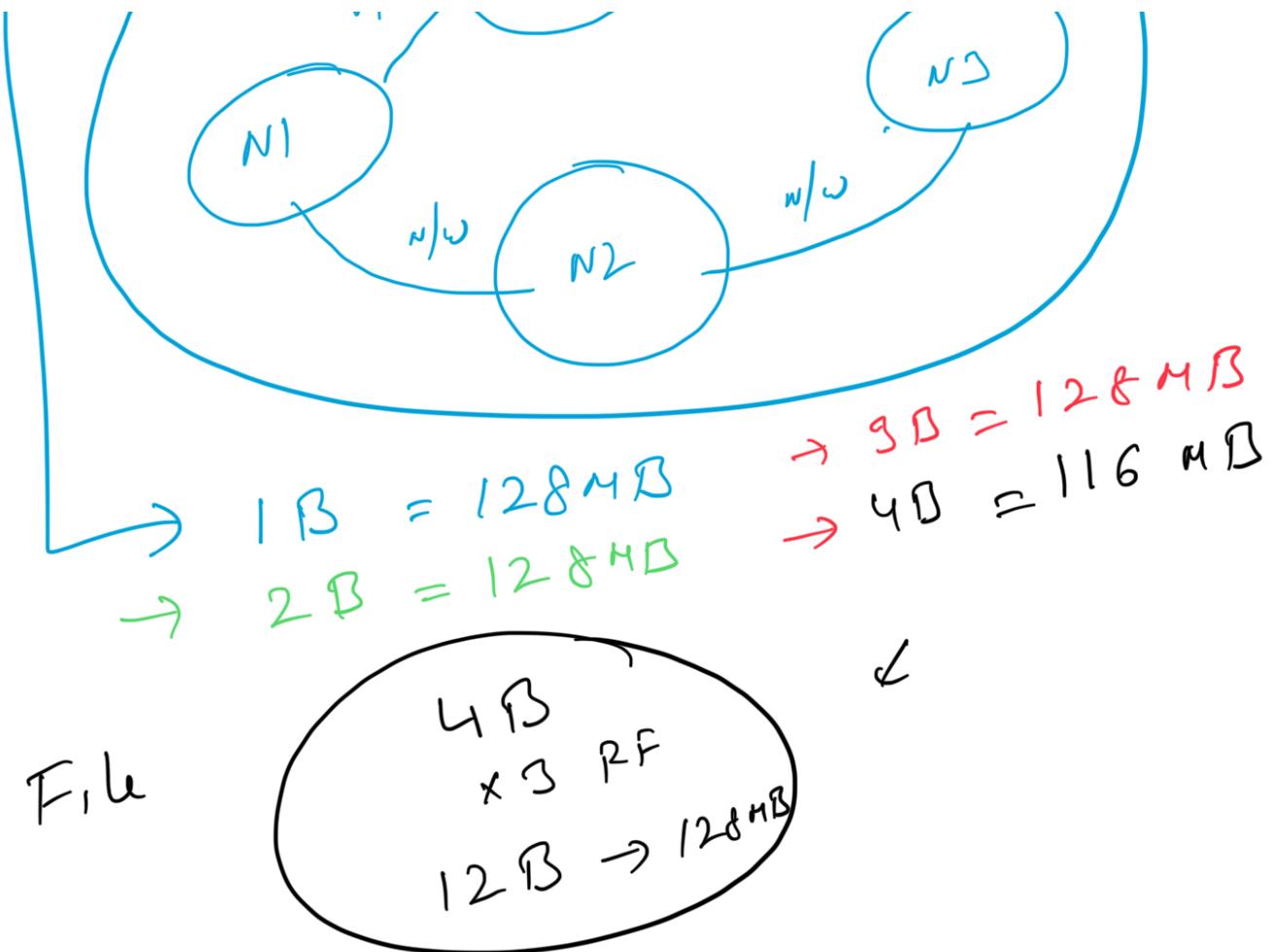


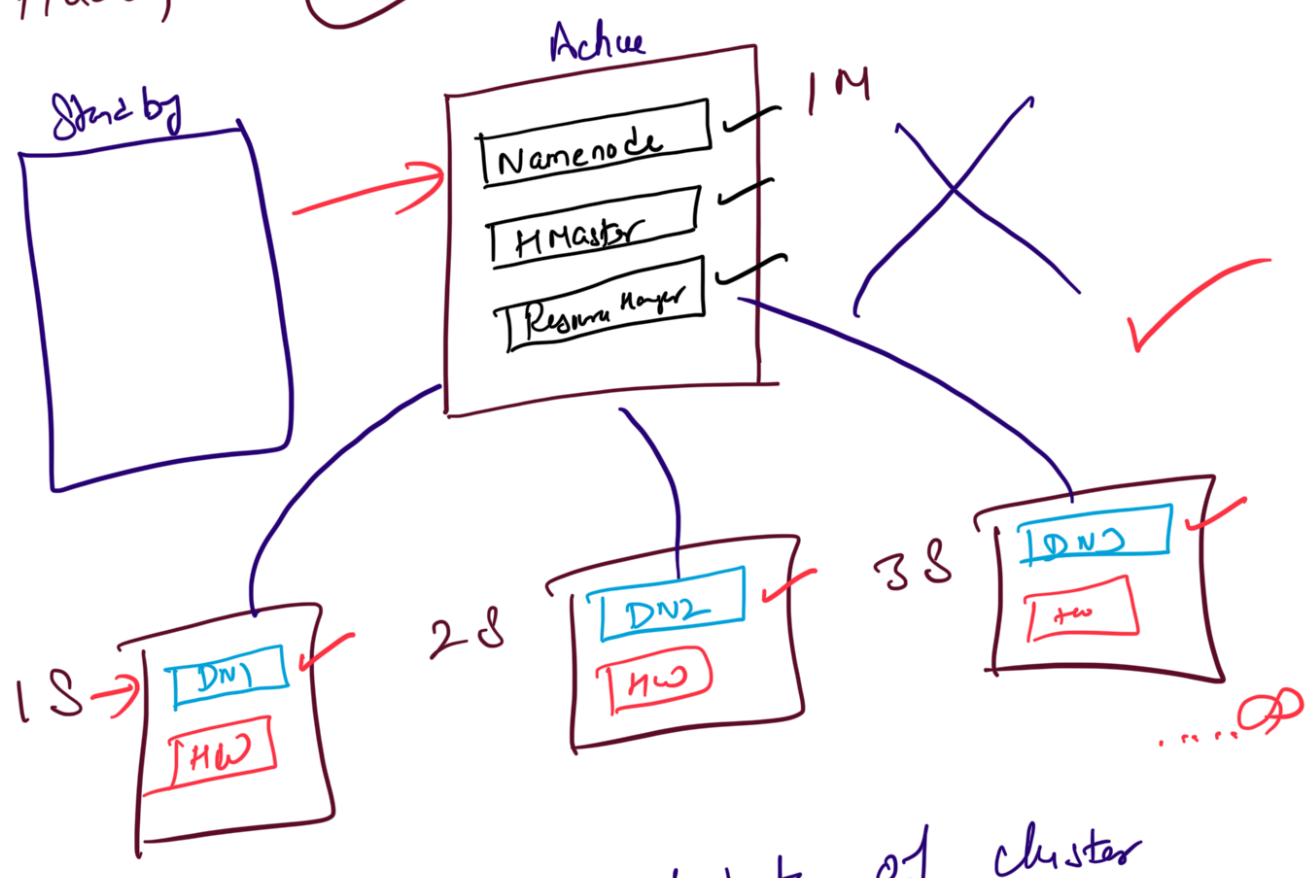
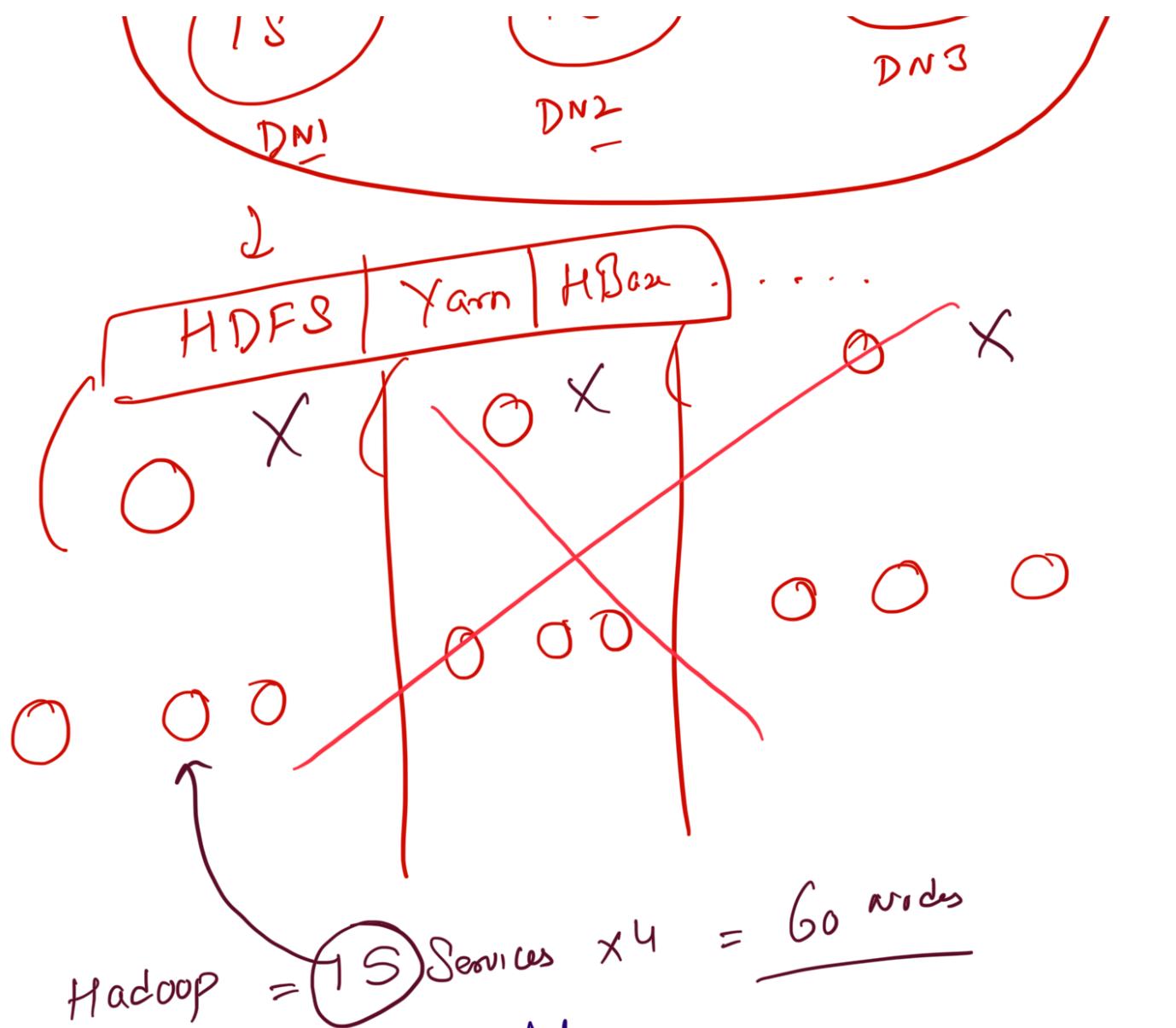
1GB → 3GB

HDFS

- ↳ Do Create File blocks
- ↳ Replicate FB by 3 times (Replication Factor)
- ↳ every FB will be of 128 MB.







Namenode :- Storing metadata "O"
↓
→ DNI : B1, B2
→ Permission
→ filename → blocks → file location
who create / edit
Big RAM

Dat node :- store the actual data