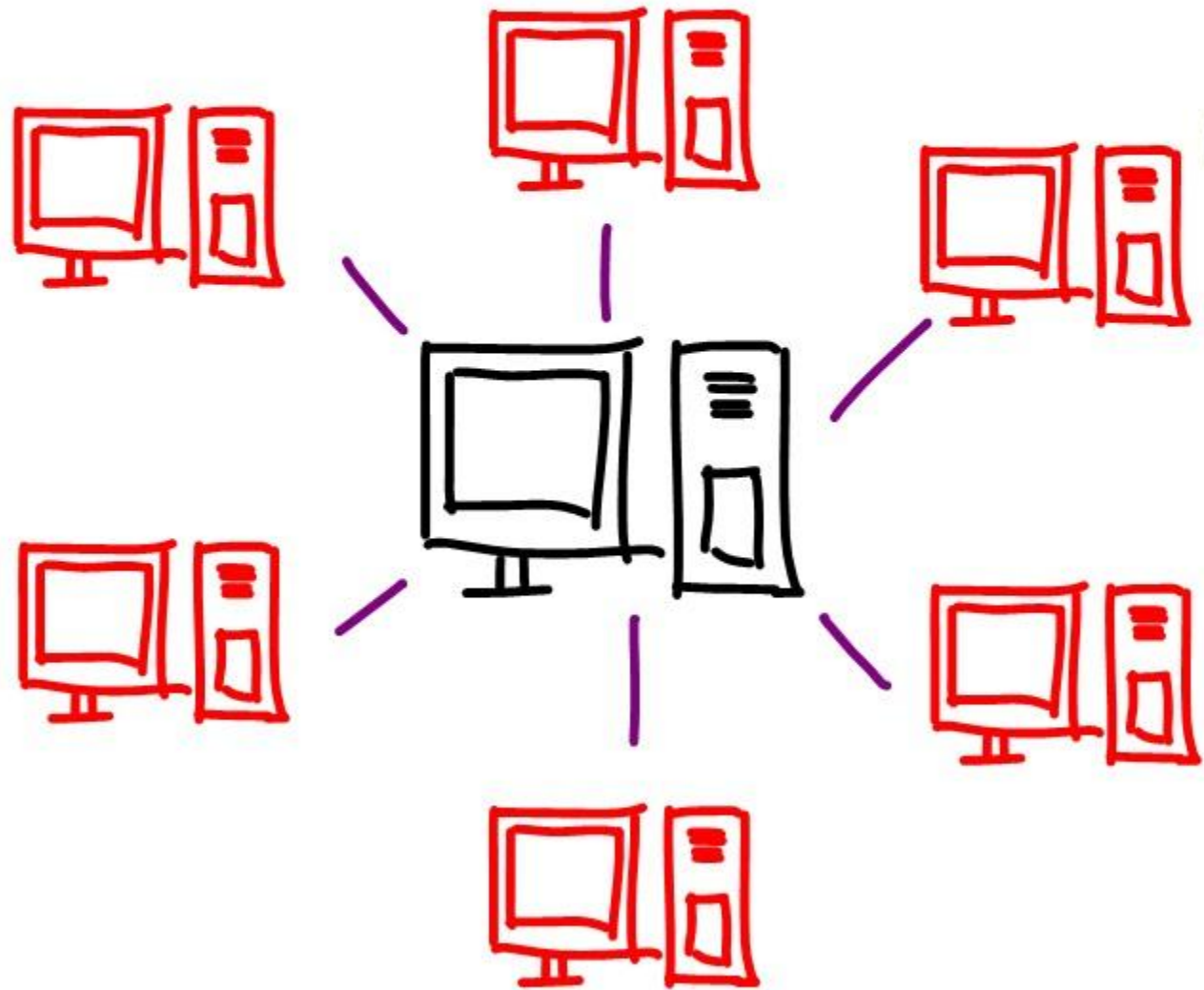
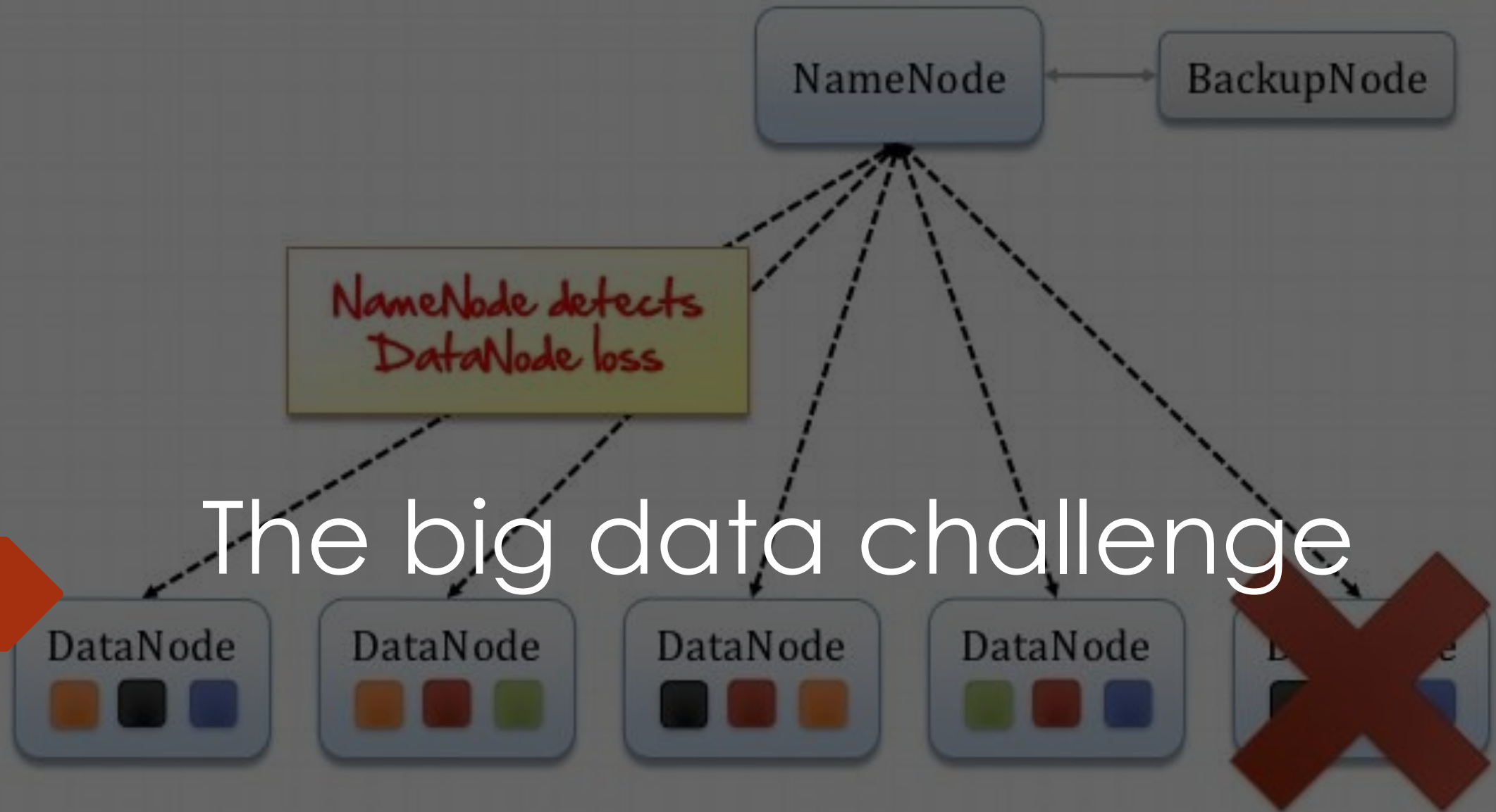




Distributed computing and data processing frameworks

Distributed Computing







The major hardware restrictions

- • Processor time
- • Memory
- • Hard drive space
- • Network bandwidth

Memory (RAM)

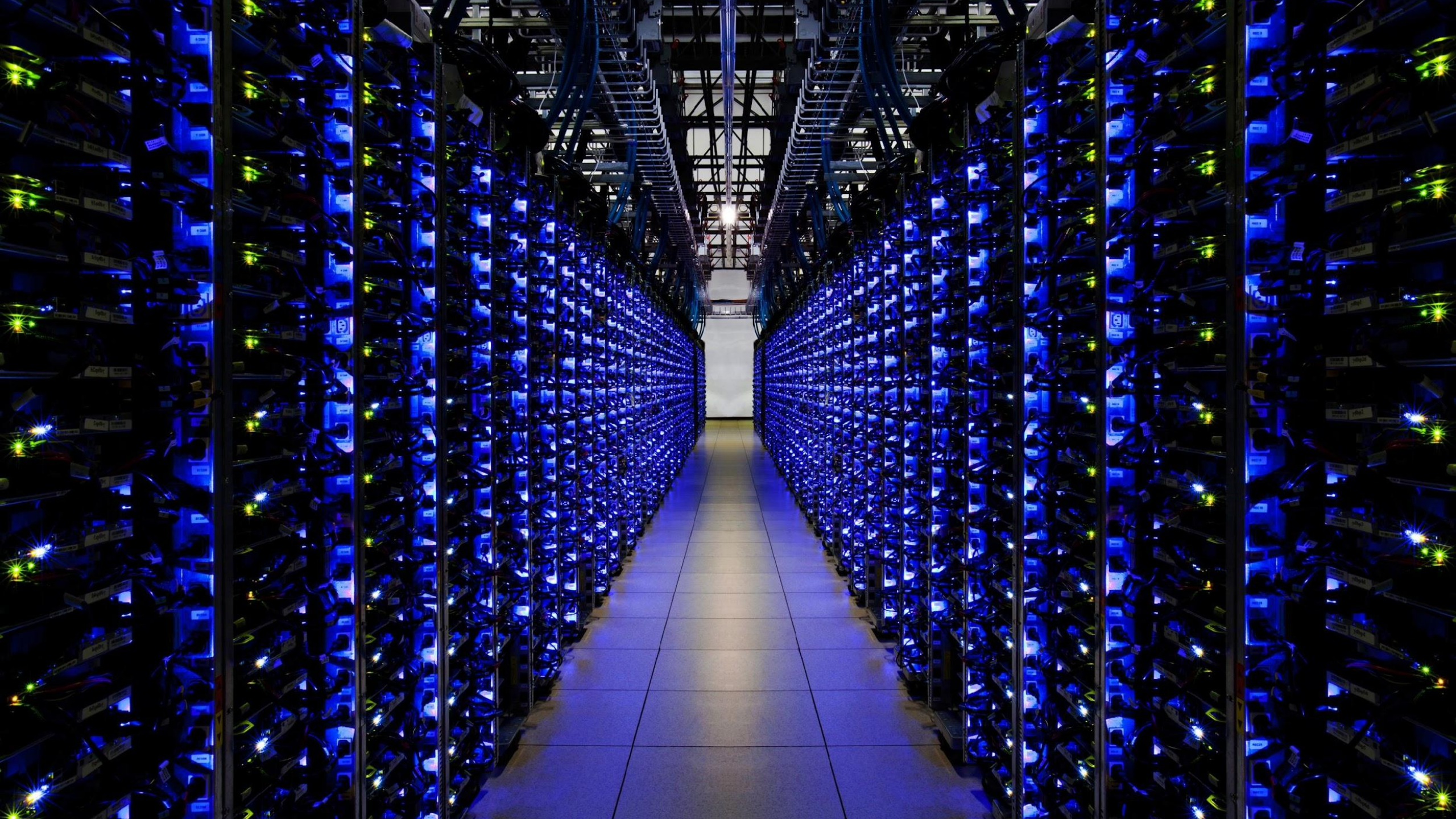


Hard drive space



BANDWIDTH

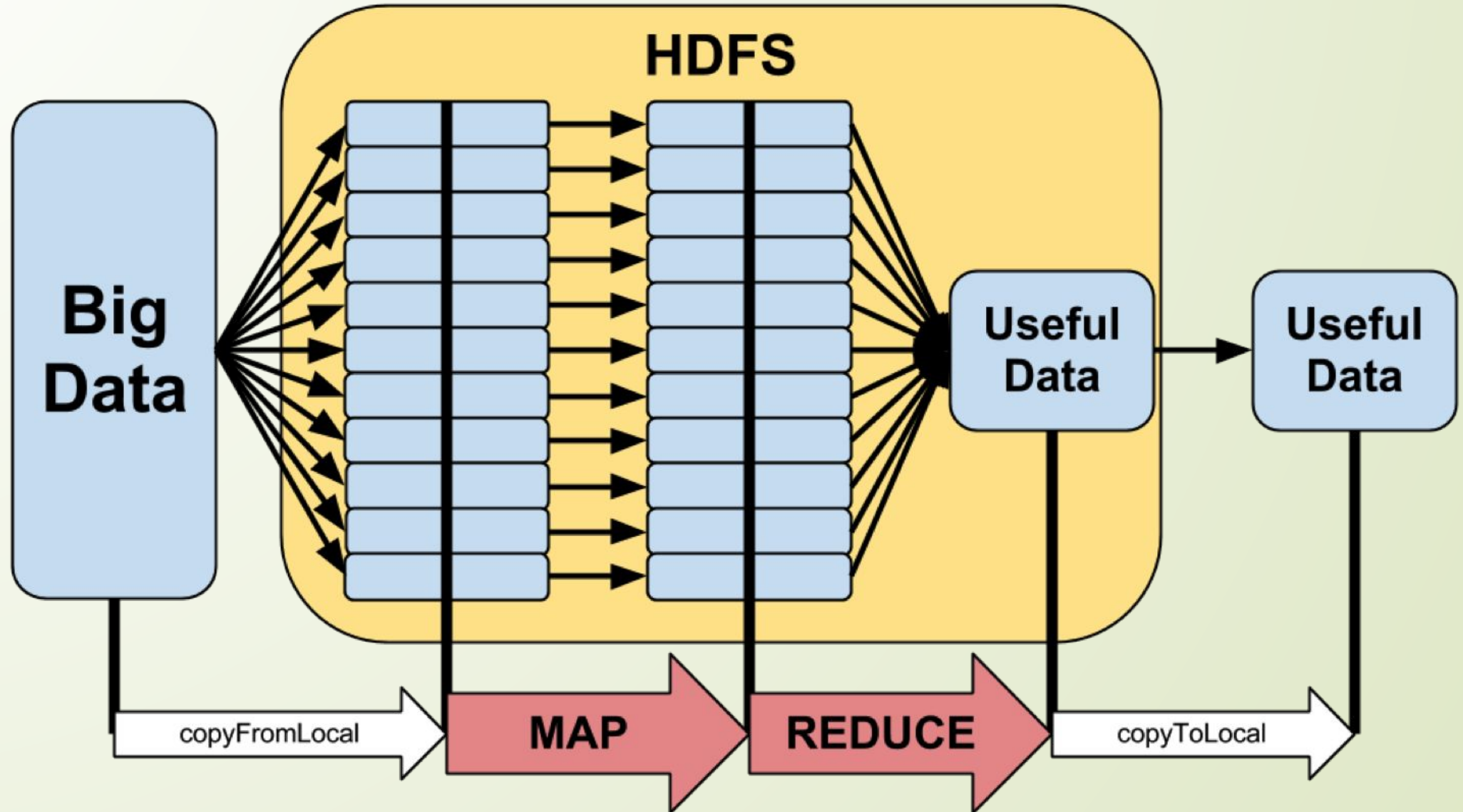


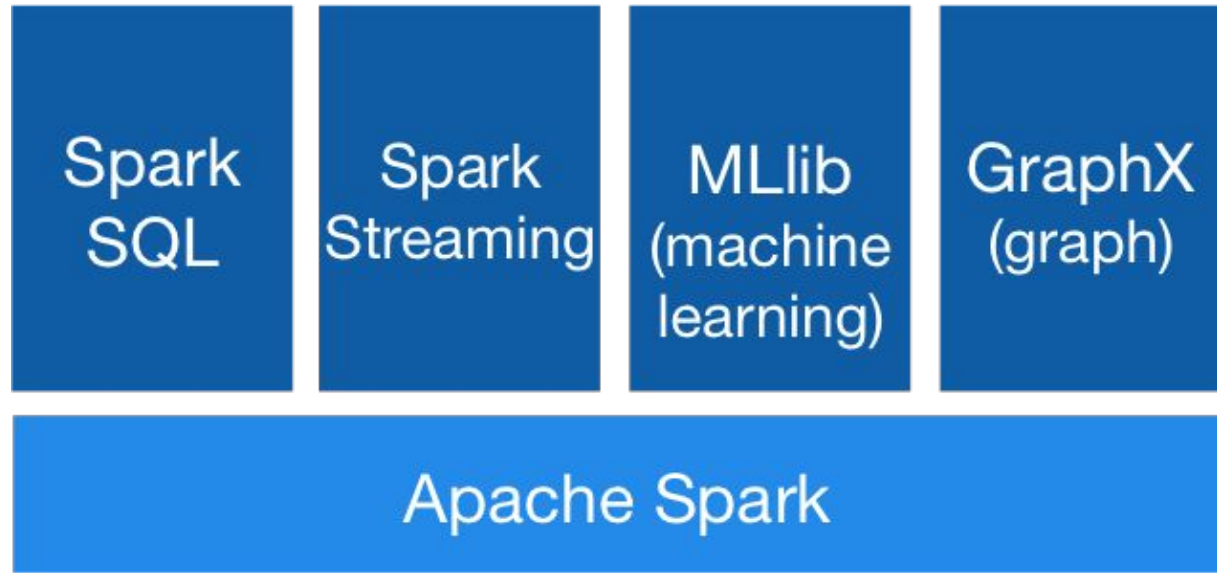




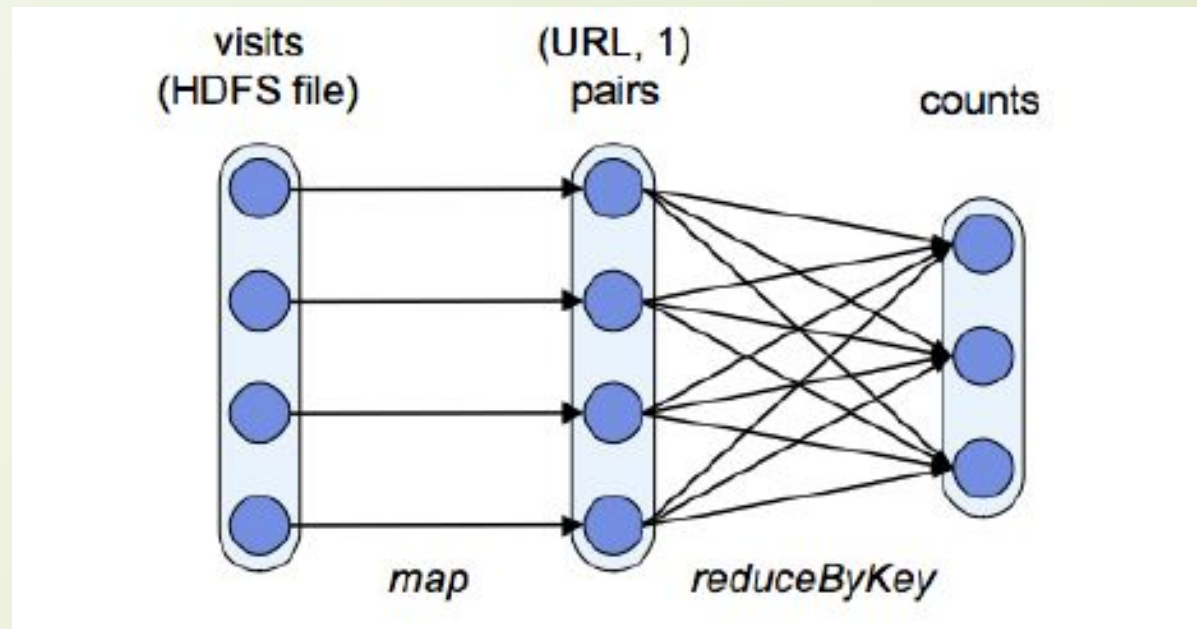
Hadoop

MapReduce



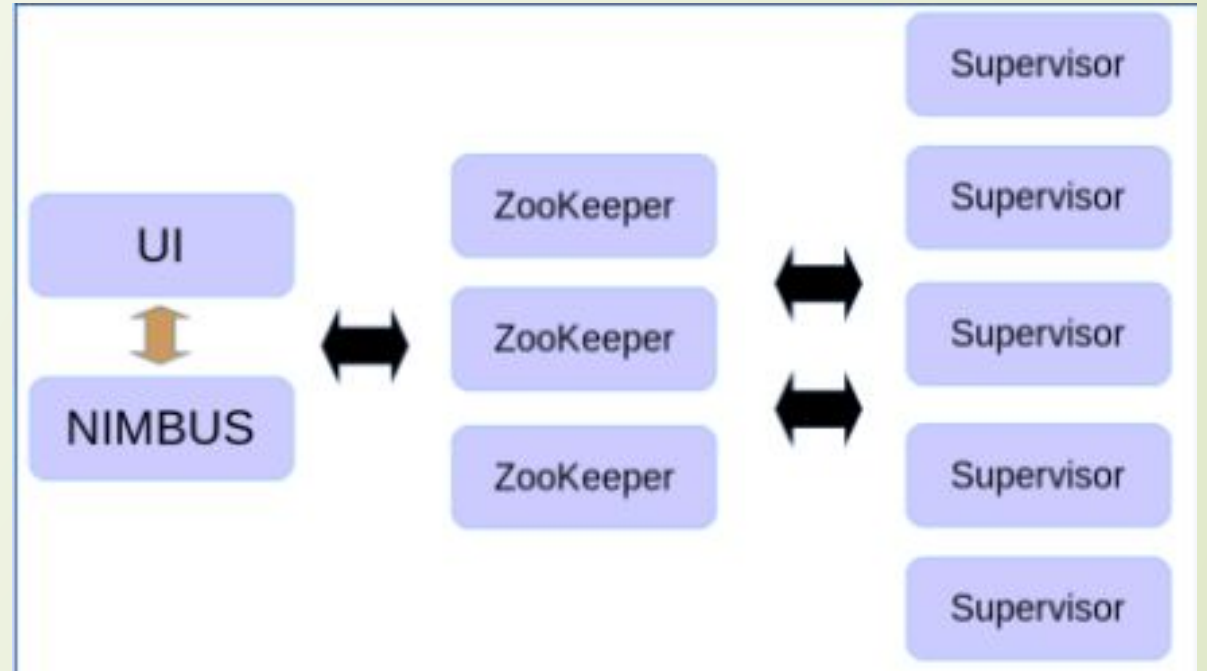


Resilient Distributed Datasets (RDDs)

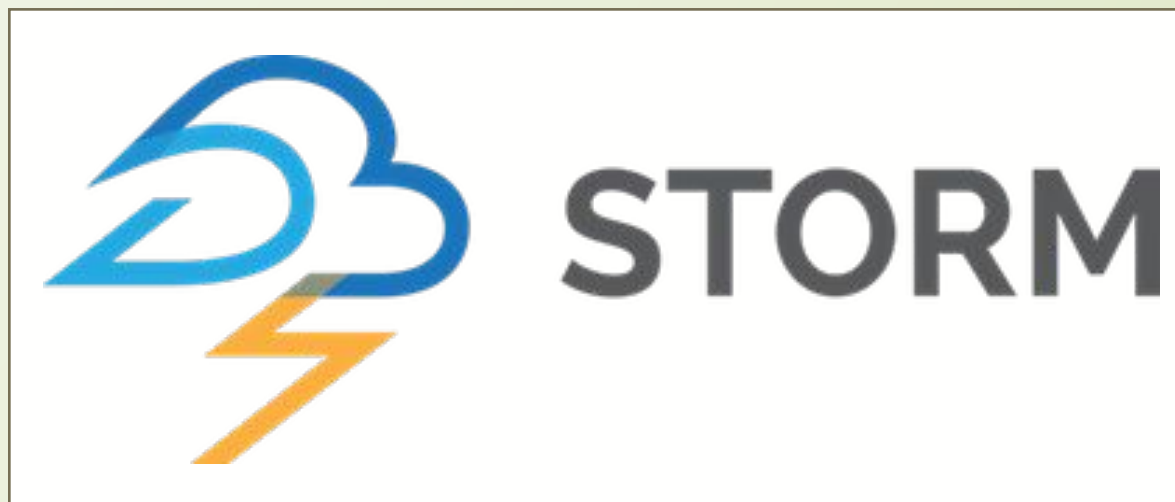
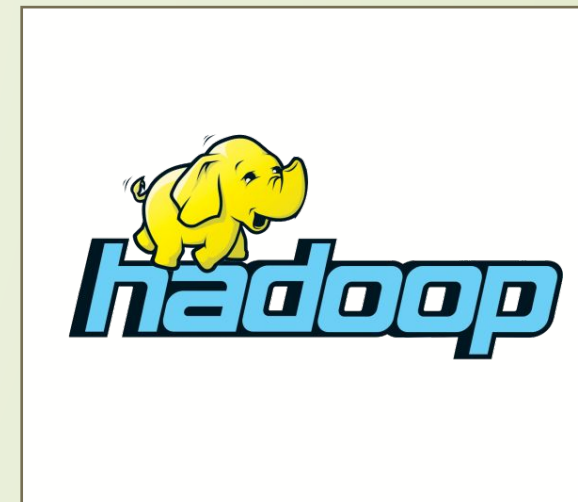


Spark  Streaming

Storm framework



Discussion



**THANKS FOR
WATCHING**