

Hadoop 101: Запусти слона

Yaroslav Yermilov
Senior Software Engineer, EPAM Systems



we have lots of data

it will be great to use it somehow

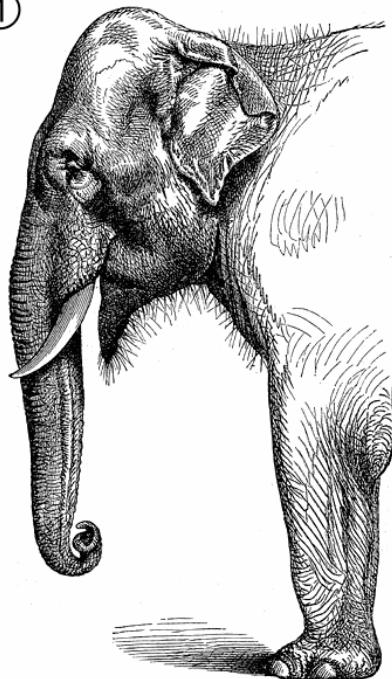




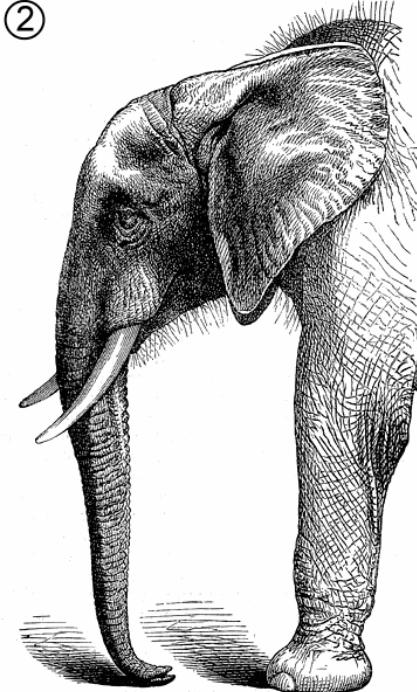
but it's really hard

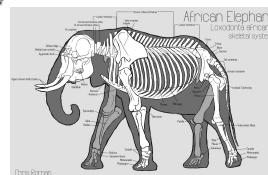
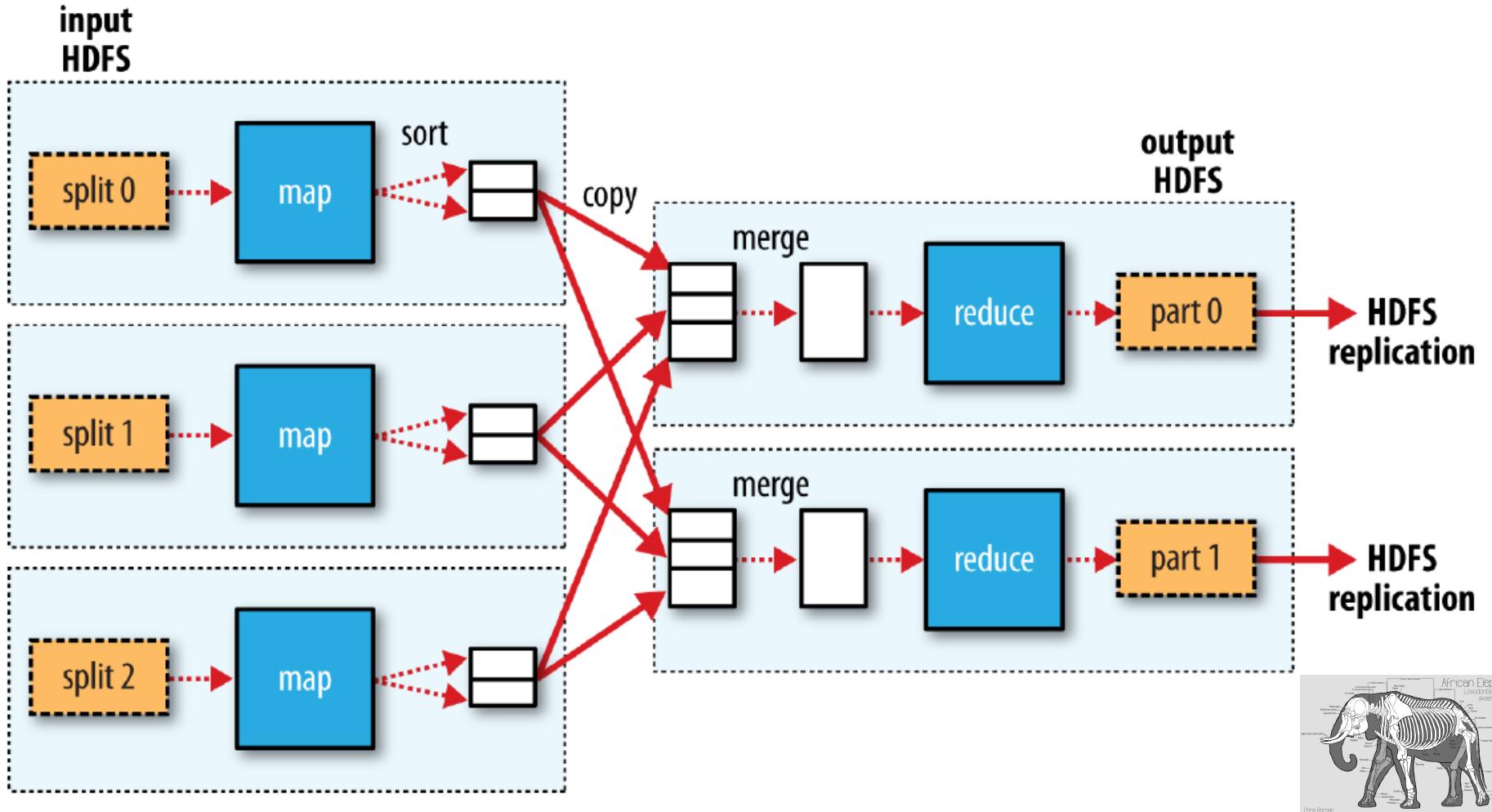
- RDBMS
- Grid computing
- Volunteer computing
- Apache Hadoop

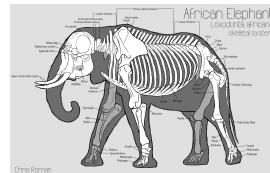
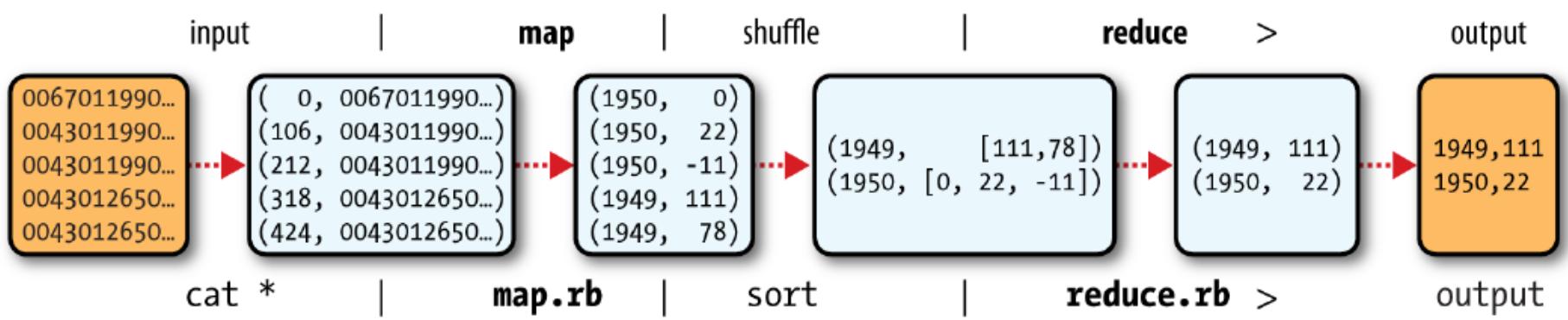
①

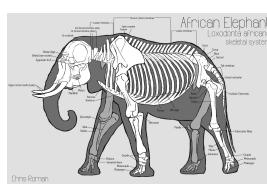
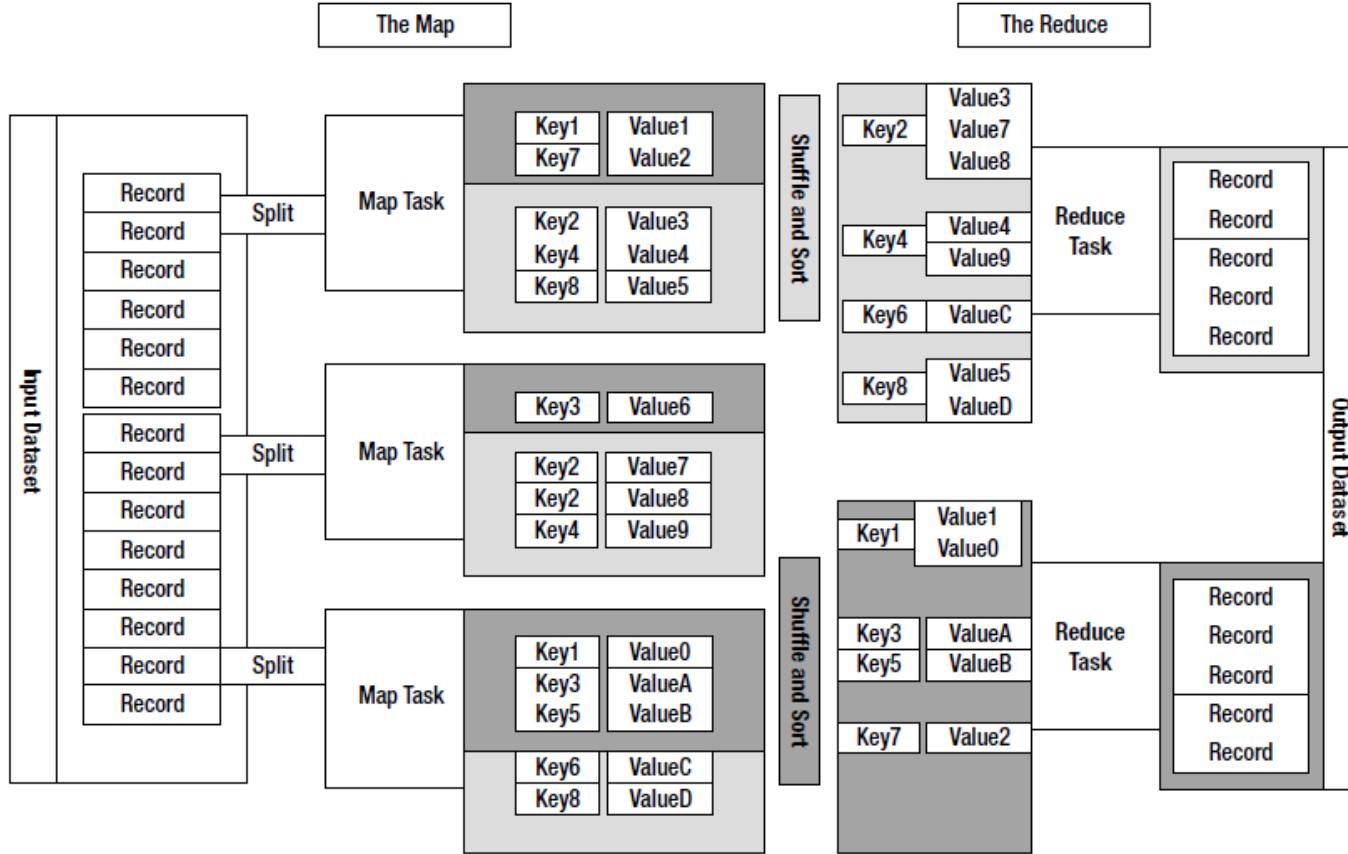


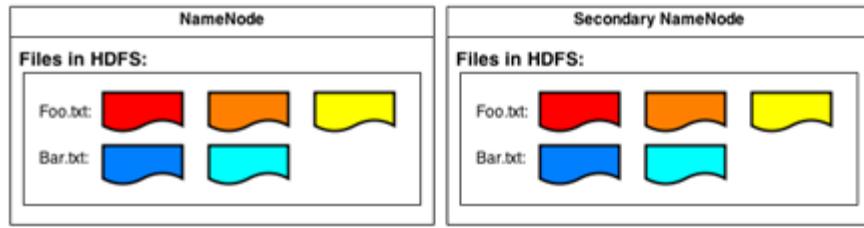
②



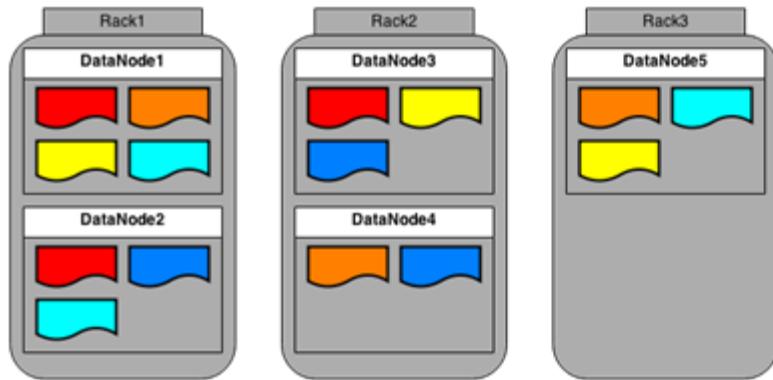




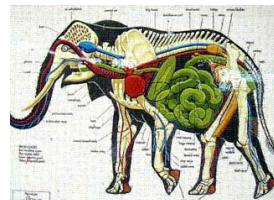


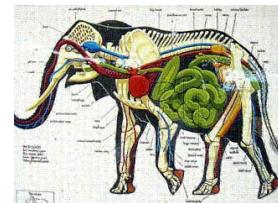
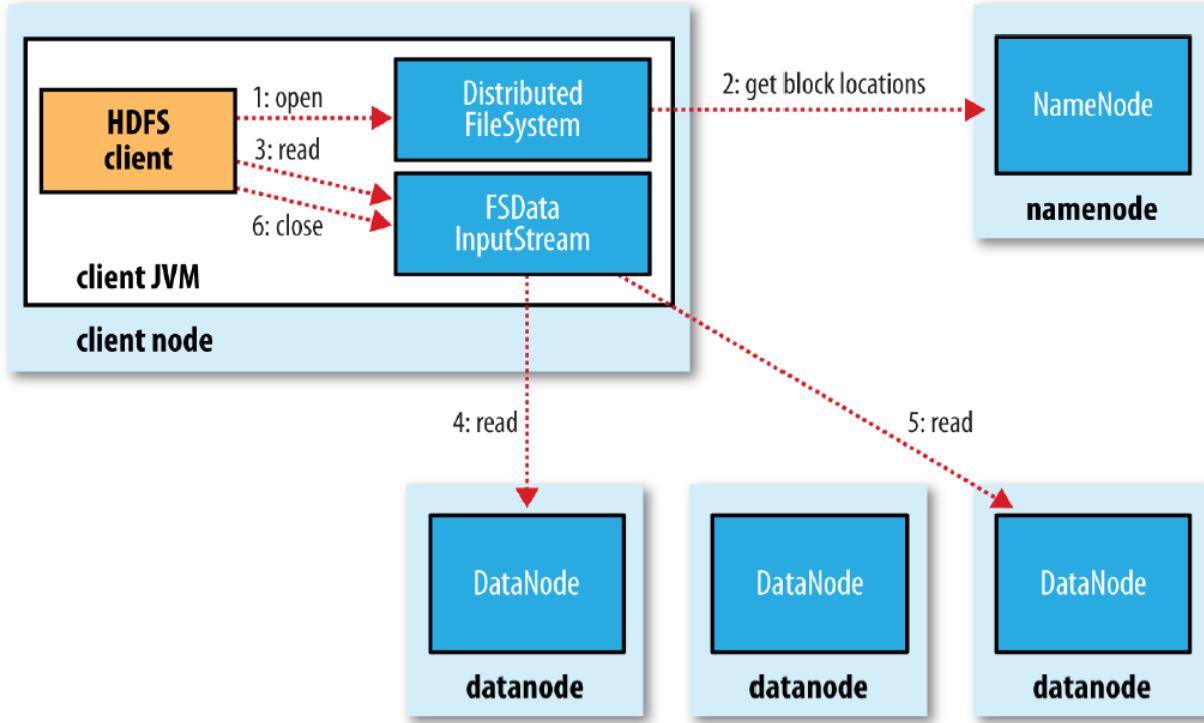


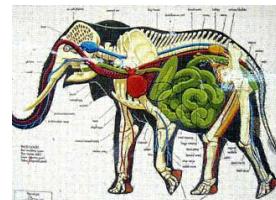
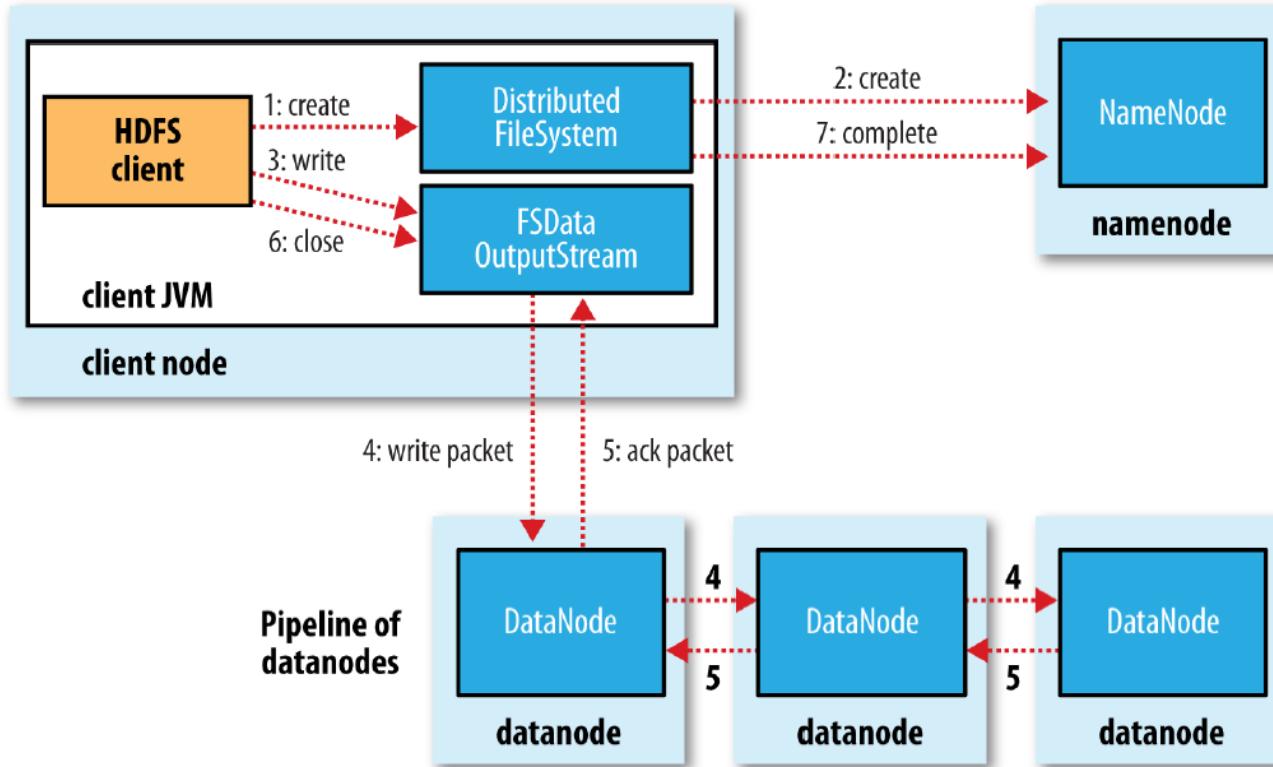
Master
Node(s)



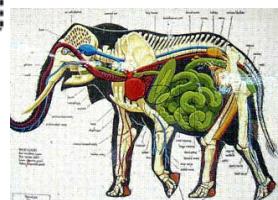
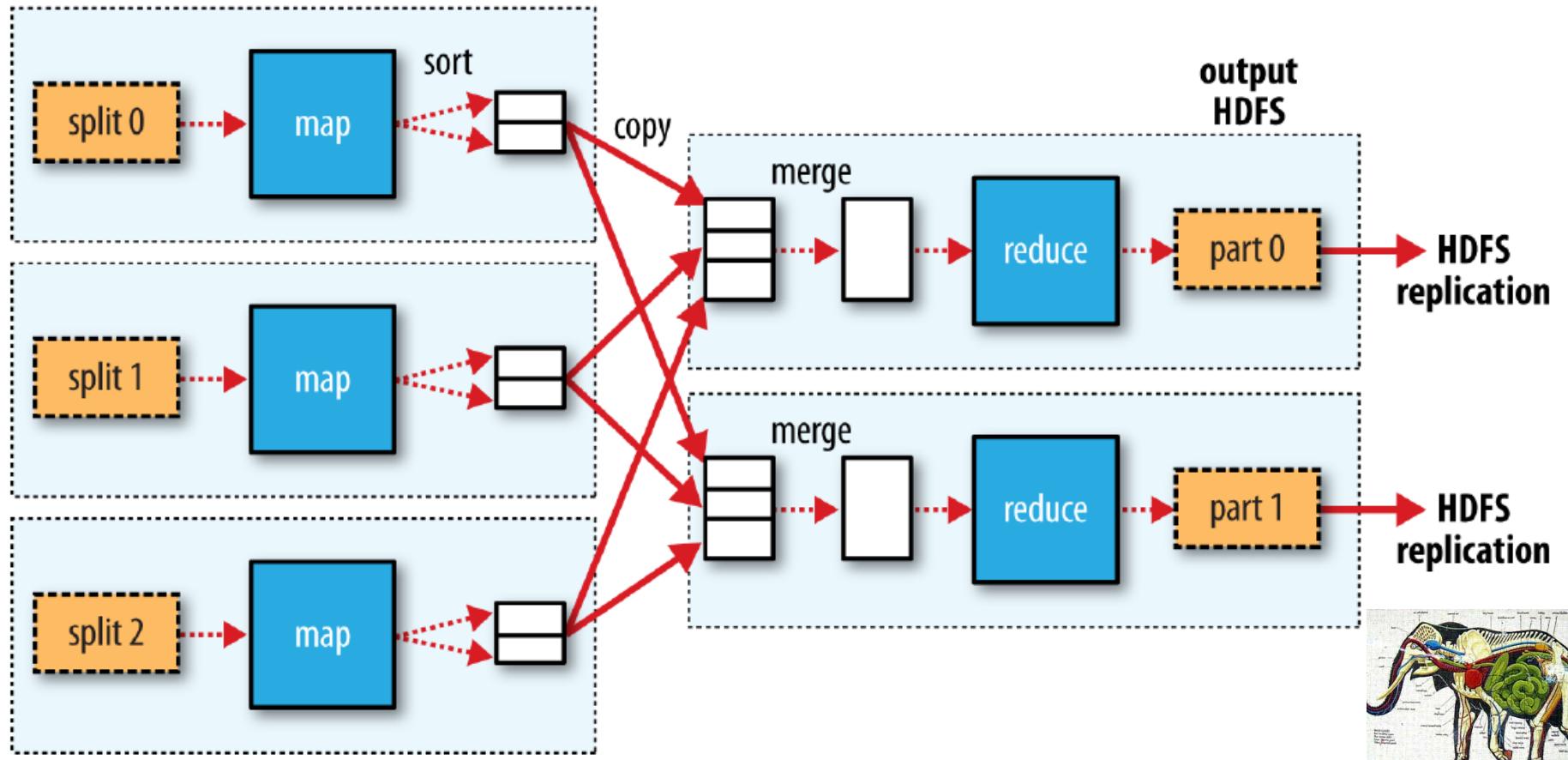
Slave
Nodes

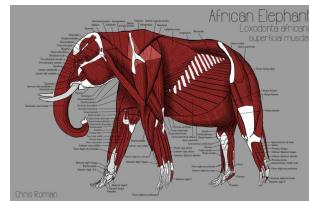
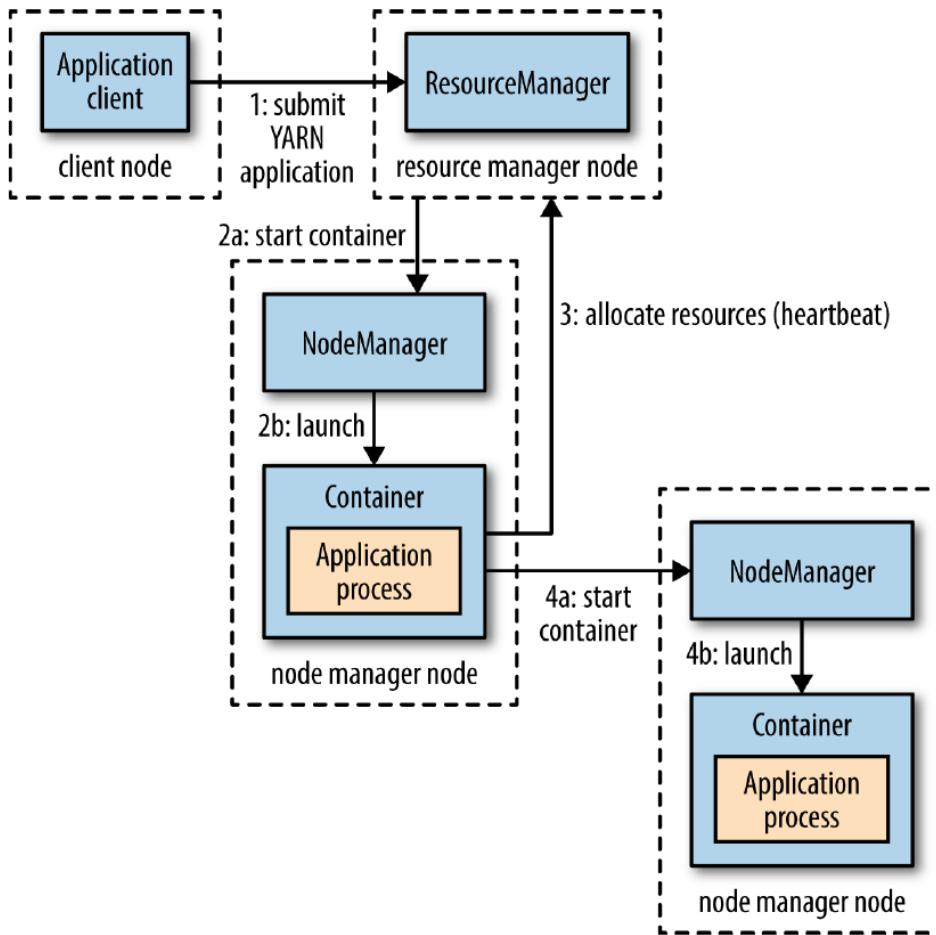


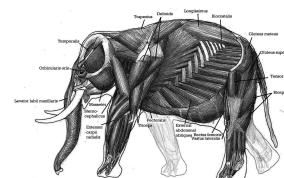
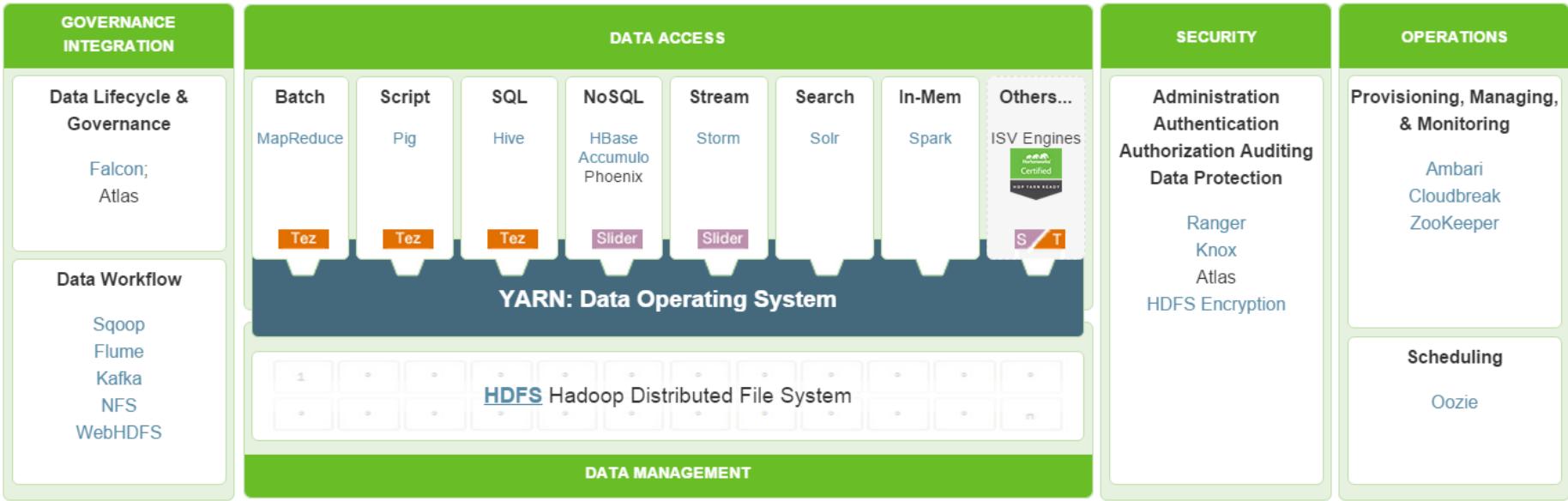


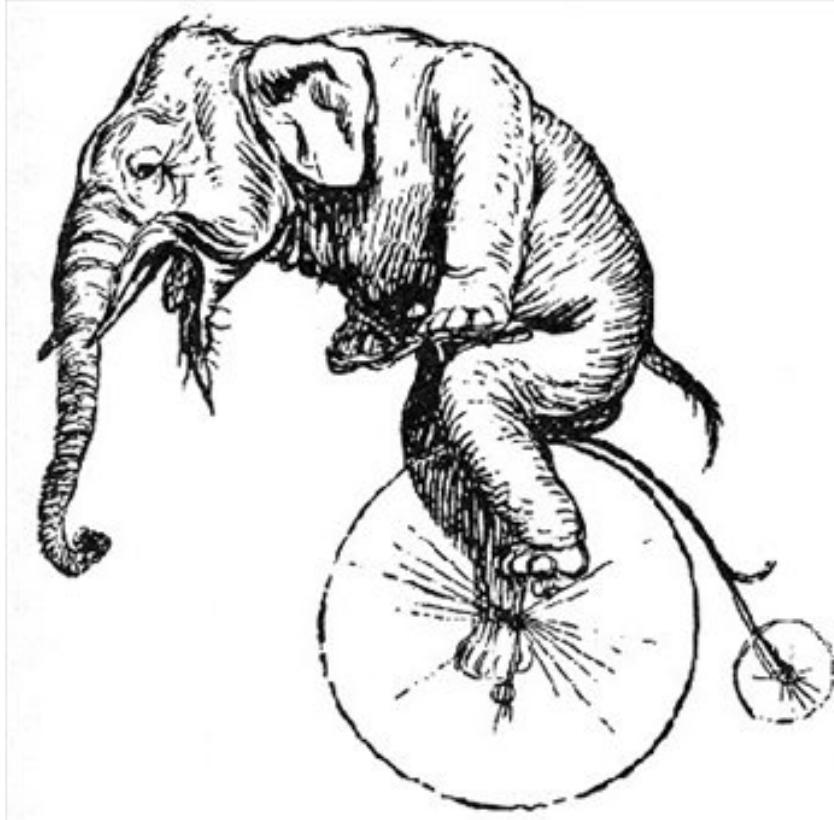


**input
HDFS**







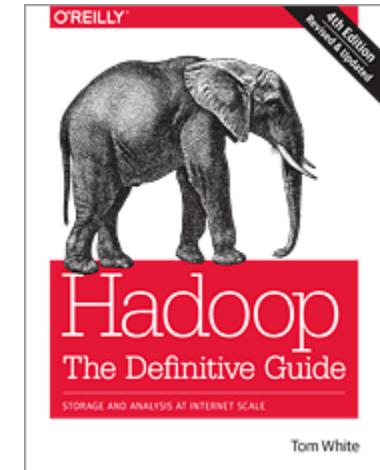






What's next?

1. Get HDP Sandbox from <http://hortonworks.com/hdp/downloads/>
2. Install and run
3. Try to run sample cases from <http://hortonworks.com/use-cases/>
4. Read the book



Thank you!

