

Map-reduce

Data science
Master 2 ISIDIS

SÉBASTIEN VEREL

verel@lisic.univ-littoral.fr

<http://www-lisic.univ-littoral.fr/~verel>

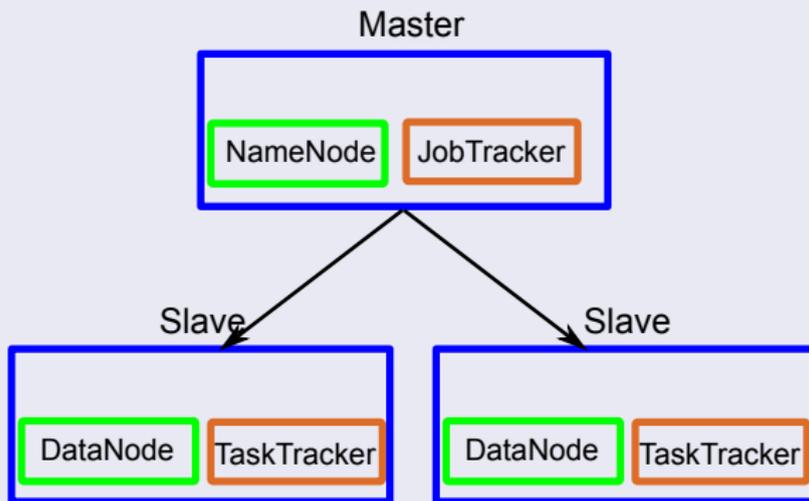
Université du Littoral Côte d'Opale

Laboratoire LISIC

Equipe OSMOSE

Le cours de big data de
Benjamin Renault
MBDS de l'université de Nice :
<http://cours.tokidev.fr/bigdata/>

Schéma Hadoop 1



Pour HDFS : NameNode, DataNode

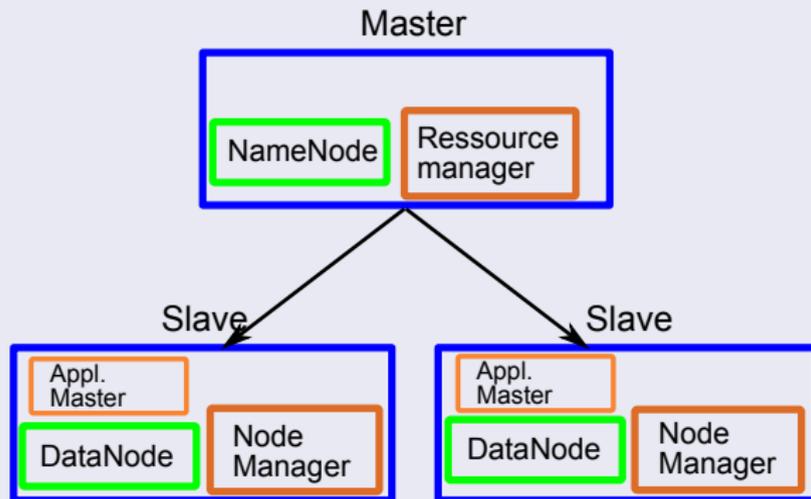
Pour MapReduce :

JobTracker : reçoit demande client et communique avec le NameNode, gère l'ensemble de déroulement de l'exécution en communiquant avec les TaskTracker

TaskTracker : exécute un map/réduce.

Schéma Hadoop 2 - Yarn

Une architecture générique (map/reduce, ou autres à définir)



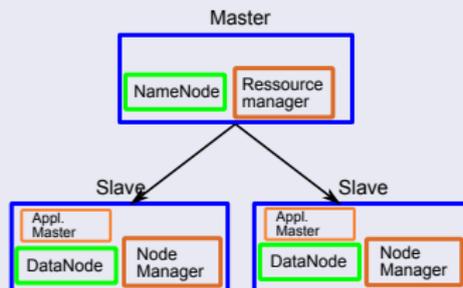
Ressource Manager : Scheduler (planifie les taches) et Application Manager (ordonnancer des taches, s'occupe de l'exécution, négocie avec les Application Masters)

Node Manager : exécution local du travail,

"responsible for launching the applications, containers, monitoring their

Schéma Hadoop 2 - Yarn

Une architecture générique (map/reduce, ou autres à définir)



Ressource Manager : Scheduler (planifie les taches) et Application Manager (ordonnancer des taches, s'occupe de l'exécution, négocie avec les Application Masters)

Node Manager : exécution local du travail,

"responsible for launching the applications, containers, monitoring their resource usage (cpu, memory, disk, network) and reporting the same to the ResourceManager." from cloudera.