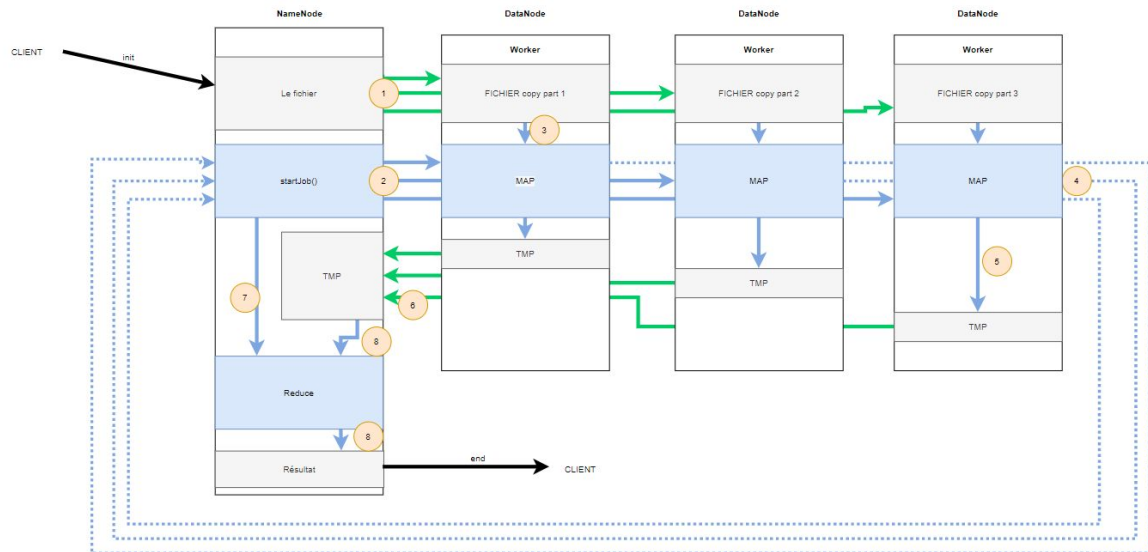


Architecture de principe

Projet HIDOOP - Cyril Figuin - Christopher James - Salma Mejbar - Cédric Martin - Axelle Pagnier



Légende :

Vert : HDFS

Bleu : MapReduce

Etapes :

init : Démarrage de l'application : le client fournit le fichier

1 : HDFS : copie sur les DataNodes du fichier

2 : startJob : lancement des MAP (appel de runMap sur les Workers)

3 : Les MAP calculent (lecture de fragments)

4 : Les MAP appellent le Callback pour dire qu'ils ont fini

5 : Génération d'un fragment de résultat (fragment de tmp)

6 : Le Reduce copie dans son FS local le fichier tmp résultat des Map (composé de fragments), grâce à HdfsClient qui contacte les HdfsServer

7 : startJob lance le Reduce

8 : Le Reduce calcule en lisant localement dans tmp

9 : Le Reduce génère le fichier résultat final