

Projet Hagidooop

Réalisé par :
Louiza Chekraoui
Safae Belahrach

Encadré par:
Daniel Hagimont



Notre Contribution I

- Nous avons implémenté le service HDFS qui connecte le Hdfs client au Hdfs Serveur via une connexion TCP en utilisant la classe NetworkReaderWriter.
- Pour le Client, on a créé des sockets pour chaque nodes en donnant leurs numéros de ports et leur adresse IP ("localhost" si sur la même machine). Pour le connecter aux différents serveurs.
- Pour chaque serveur, on lui créé un ServerSocket qu'on mettera par la suite sur accept() pour accepter les connexions provenant du client.

Notre Contribution II

- Nous avons implémenté le service Hagidooop, on a lancé comme demandé des workers sur chaque node, et cela, en utilisant RMI qui va créer des registres liés au port de la ligne de commande.
- Chaque chemin RMI est utilisé pour rechercher et obtenir une référence à un objet Worker à distance, et ces références sont stockées dans un tableau de travailleurs.
- Ensuite, du côté JobLauncher, nous allons obtenir des références vers les objets workers à l'aide de la méthode lookup de la classe Naming afin de lancer leurs runmap par le startJob du JobLancher, qui va récupérer les KV résultant des map dans un PipeReaderWriter (Queue) pour ainsi appliquer le reduce .

Limitations

- Script de déploiement : lancer toutes les commandes sur les différents démons à la fois et les nettoient. On lance à la main les commandes sur différents terminaux.
- Les fichiers traités sont que ce de la forme .txt on n'a pas traité les fichiers sous forme de KV.




Evaluation des performances

Nombre de noeuds = 3 Noeuds

Taille du fichier initial 	2 Ko	1305 Ko	41 Ko	1,3 Go	...
Temps de Performance MyMapReduce 	261 ms	1286 ms	265 ms	72363 ms	
Temps de Performance Count 	35 ms	229 ms	44 ms	109612 ms	




Evaluation des performances en local

Taille de fichier = 1.3 Go

Nombre des noeuds 	1 node	2 nodes	3 nodes	4 nodes	...
Temps de Performance MyMapReduce 	97810 ms	48956 ms	72363 ms	40099 ms	
Temps de Performance Count 	89261 ms	82951 ms	109612 ms	101459 ms	

Evaluation des performances en distribué

Taille de fichier = 1.3 Go

Nombre des noeuds 	2 nodes	3 nodes
Temps de Performance MyMapReduce 	12277 ms	10530 ms
Temps de Performance Count 	24812 ms	24587 ms