

SUPDAT - MINI PROJET5DAT

Ce projet à été généré en utilisant Vagrant. Afin de fournir un environnement Hadoop disponible au correcteur, lui permettant d'exécuter nos scripts PIG.

Version Logiciel : - Devian 8 Jessie 64 bits - Hadoop 1.2.1

Contexte

Mini projet permettant de manipuler nos gros fichiers volumineux dans un environnement Hadoop, avec PIG.

IMPORTANT : - Le projet se compose de 2 parties distincts. - Blabla partie I. - Blabla partie II.

EQUIPES

NOM	PRENOM	IDBooster
TRAORE	Oumar	165366
BOYER	Clément	159913
BOMO	NGABA Xavier	216113

Rôle : - **TRAORE Oumar** : Installation de l'infrastructure Hadoop, Gestion du projet. - **BOYER Clément** : Partie I. - **BOMO NGABA Xavier** : Partie II.

PREREQUIS

Vous devez avoir ces prérequis

- vagrant <https://www.vagrantup.com/>
- virtualbox <https://www.virtualbox.org/>
- Memoire RAM 1024 MB
- Memoire Physique 20 GB

INSTALLATION

Exécuter ces commandes depuis votre terminal

- `vagrant up` Démarrer la VM
 - `vagrant ssh` Connexion à la VM --> Machine Guest (Vous êtes sur la machine Guest à ce niveau)
 - `start-all.sh` Lancement des daemons HDFS & Mapreduce en mode Pseudo Distribué, sur un nœud (Suffisant pour PIG, dans ce POC)
-

DEPLOIEMENT (TEST || DEMO || ...)

Depuis la machine **Hôte**, sur votre Navigateur

- `http://localhost:50070` Interface du NameNode
- `http://localhost:50030` Interface du JobTracker

Depuis la machine **Guest** sur votre terminal, exécuter

- `cd /vagrant` Pour accéder au répertoire partager, contenant les sources à réaliser
- `cd PartieI/scriptX.sh` X: [1..5] Répéter cette action, pour chaque script à tester de la Partie I
- `cd PartieII/scriptX.sh` X: [1..5] Répéter cette action, pour chaque script à tester de la Partie II

IMPORTANT :

- A l'attention du **Correcteur**, il n'est pas nécessaire d'exécuter l'environnement, vous pouvez immédiatement consulter le résultat attendu dans le dossier source "src" présent à la racine de l'archive.

Fin du Prof of concept du projet 5DAT.