## TP: Hadoop HDFS

 Tout d'abord, vous devez cloner le dépôt Git suivant dans le répertoire de votre choix sur votre système. Je préfère le cloner dans le home pour la démonstration et l'accès facile.

"git clone https://github.com/big-data-europe/docker-hadoop"

- 2. puis entrez dans le référentiel du projet à l'aide de la commande change directory : cd docker-hadoop
- 3. tapez : **Is -I**

```
(base) paras@IUDX:~/docker-hadoop$ ls -l
total 52
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 base
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 datanode
-rw-rw-r-- 1 paras paras 2522 Mar 17 14:44 docker-compose-v3.yml
-rw-rw-r-- 1 paras paras 1559 Mar 17 14:44 docker-compose.yml
-rw-rw-r-- 1 paras paras 2507 Mar 17 14:44 hadoop.env
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 historyserver
-rw-rw-r-- 1 paras paras 1437 Mar 17 14:44 Makefile
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 namenode
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 nginx
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 nodemanager
-rw-rw-r-- 1 paras paras 2171 Mar 17 14:44 README.md
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 resourcemanager
drwxrwxr-x 2 paras paras 4096 Mar 17 14:44 submit
```

Output of the command "Is -I"

## 4. Obtenez les images Docker et créez des conteneurs

```
docker-compose up -d
```

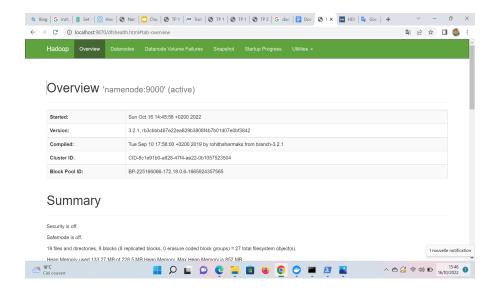
La commande ci-dessus téléchargera toutes les images Docker nécessaires à partir du hub Docker pour configurer les conteneurs HDFS. Le téléchargement des images peut prendre un peu de temps en fonction de votre vitesse Internet.

Maintenant, pour voir vos conteneurs Docker en cours d'exécution, utilisez la commande pour répertorier tous les conteneurs actifs.

```
docker container ls
or
docker ps
```



Et l'état actuel peut également être vérifié à l'aide de la page Web http://localhost:9870 :



## Premier pas avec Hadoop

Toutes les commandes interagissant avec le système Hadoop commencent par **hadoop fs** (aussi **hdfs dfs**). Ensuite, les options rajoutées sont très largement inspirées des commandes Unix standard.

1- Créer un répertoire dans HDFS, appelé input (Tapez l'une de ces commandes)





Ajouter des fichiers textes (.txt) dans le namenode

docker cp C:\Users\amina\OneDrive\Bureau\TP\words.txt
namenode:/input/

Travail à faire (en utilisant les commandes HDFS):

- 1. Vérifier si le fichier words.txt est bien placé dans votre répertoire locale,
- 2. Placer le fichier words.txt dans le répertoire que vous avez créer
- 3. afficher le contenu de répertoire /input
- 4. télécharger le contenu de répertoire /input
- 5. lire les derniers lignes de fichier words.txt