

## Travaux pratiques (TP – 02)

### Manipulation de HDFS

#### Objectifs

- Comprendre les bases de la manipulation de fichiers dans HDFS, incluant le stockage, la lecture et la gestion de fichiers. Ce TP permettra aux étudiants d'apprendre les commandes essentielles d'HDFS et de se familiariser avec la gestion de données dans un système de stockage distribué.

#### Prérequis :

- Avoir un accès à un environnement Hadoop avec HDFS configuré.

#### Étape 1 : Stocker un fichier dans HDFS

Création d'un fichier local :

Créez un fichier texte sur votre machine locale. Par exemple, vous pouvez utiliser la commande suivante pour créer un fichier nommé exemple.txt :

```
echo "Bienvenue dans le TP sur HDFS !" > exemple.txt
```

Copier le fichier dans HDFS :

Utilisez la commande put pour copier le fichier exemple.txt dans HDFS.

Commande :

```
hdfs dfs -put exemple.txt /user/<votre_nom_utilisateur>/
```

Objectif : Vérifier que le fichier est bien stocké dans HDFS.

Vérifier la présence du fichier dans HDFS :

Listez les fichiers dans le répertoire /user/<votre\_nom\_utilisateur>/ pour confirmer que exemple.txt a bien été copié.

Commande :

```
hdfs dfs -ls /user/<votre_nom_utilisateur>/
```

#### Étape 2 : Lire un fichier depuis HDFS

Lire le contenu du fichier :

Utilisez la commande cat pour lire le contenu du fichier exemple.txt depuis HDFS.

Commande :

```
hdfs dfs -cat /user/<votre_nom_utilisateur>/exemple.txt
```

Objectif : Visualiser le contenu du fichier pour s'assurer qu'il a été correctement chargé dans HDFS.

Lire les premières lignes d'un fichier :

Utilisez la commande head pour afficher les premières lignes du fichier exemple.txt dans HDFS (surtout utile pour les fichiers volumineux).

Commande :

```
hdfs dfs -head /user/<votre_nom_utilisateur>/exemple.txt
```

### Étape 3 : Gérer les fichiers dans HDFS

Renommer un fichier dans HDFS :

Utilisez la commande mv pour renommer exemple.txt en exemple\_modifié.txt.

Commande :

```
hdfs dfs -mv /user/<votre_nom_utilisateur>/exemple.txt  
/user/<votre_nom_utilisateur>/exemple_modifié.txt
```

Objectif : Comprendre comment renommer un fichier dans HDFS.

Copier un fichier dans HDFS :

Créez une copie de exemple\_modifié.txt dans HDFS sous un autre nom, par exemple exemple\_copie.txt.

Commande :

```
hdfs dfs -cp /user/<votre_nom_utilisateur>/exemple_modifié.txt  
/user/<votre_nom_utilisateur>/exemple_copie.txt
```

Supprimer un fichier dans HDFS :

Utilisez la commande rm pour supprimer exemple\_copie.txt de HDFS.

Commande :

```
hdfs dfs -rm /user/<votre_nom_utilisateur>/exemple_copie.txt
```

Objectif : Apprendre à gérer la suppression de fichiers dans un environnement distribué.

Créer un répertoire dans HDFS :

Créez un répertoire nommé mon\_dossier dans HDFS pour organiser vos fichiers.

Commande :

```
hdfs dfs -mkdir /user/<votre_nom_utilisateur>/mon_dossier
```

Déplacer un fichier dans un répertoire HDFS :

Déplacez exemple\_modifié.txt dans le répertoire mon\_dossier.

Commande :

```
hdfs dfs -mv /user/<votre_nom_utilisateur>/exemple_modifié.txt  
/user/<votre_nom_utilisateur>/mon_dossier/
```

#### Étape 4 : Vérification et Nettoyage

Vérifier la structure des fichiers et des dossiers :

Listez le contenu de /user/<votre\_nom\_utilisateur>/ pour vérifier que le fichier a bien été déplacé dans mon\_dossier.

Commande :

```
hdfs dfs -ls /user/<votre_nom_utilisateur>/mon_dossier/
```

Supprimer les fichiers et répertoires créés (si nécessaire) :

Si vous n'avez plus besoin des fichiers et dossiers créés, vous pouvez les supprimer pour nettoyer votre espace HDFS.

Commande pour supprimer le répertoire et son contenu :

```
hdfs dfs -rm -r /user/<votre_nom_utilisateur>/mon_dossier
```

À la fin de ce TP, vous devriez être à l'aise avec les commandes de base de HDFS pour :

- Stocker des fichiers,
- Lire et vérifier le contenu des fichiers,
- Gérer les fichiers (renommer, copier, déplacer, et supprimer),
- Organiser les données dans des répertoires.

Ces compétences sont essentielles pour travailler efficacement avec HDFS dans un environnement Big Data distribué.