

Question 2.8 – Lire un fichier compact

Nous lançons notre instance HDFS.

Tout d’abord, nous créons un dossier `/user/thibaut/question2.8` à l’aide de la commande :

```
hdfs dfs -mkdir /user/thibaut/question2.8
```

Ensuite, nous plaçons le fichier `noaa.txt` comme demandé :

```
hdfs dfs -put noaa.txt /user/thibaut/question2.8
```

Enfin, nous affichons les dernières lignes à l’aide de la commande `tail` :

```
hdfs dfs -tail /user/thibaut/question2.8 /noaa.txt
```

Résultat :

```
227.1 19680101 19721231
999999 94908 DUBUQUE MUNICIPAL AP      US  IA  KDBQ  +42.398 -090.704 +0329.2 19510201 19721231
999999 94910 WATERLOO MUNICIPAL AP     US  IA  KALO  +42.554 -092.401 +0267.6 19490101 19721231
999999 94925 GRAND FORKS AF            US  ND  KRDR  +47.967 -097.400 +0277.7 19710101 19710101
999999 94931 HIBBING CHISHOLM-HIBBING AP US  MN  KHIB  +47.386 -092.839 +0413.6 19720101 19721231
999999 94995 LINCOLN 8 ENE              US  NE           +40.848 -096.565 +0362.4 20020115 20171006
999999 94996 LINCOLN 11 SW             US  NE           +40.695 -096.854 +0418.2 20020114 20171006
999999 96404 TOK 70 SE                  US  AK           +62.737 -141.208 +0609.6 20110924 20171006
999999 96406 RUBY 44 ESE                US  AK           +64.502 -154.130 +0078.9 20140828 20171006
999999 96407 SELAWIK 28 E              US  AK           +66.562 -159.004 +0006.7 20150813 20171006
999999 96408 DENALI 27 N               US  AK           +63.452 -150.875 +0678.2 20150819 20171006
MacBook-Pro-de-Thibaut:Data td$ █
```

Codage et tests

Bien que le lancement en **Standalone mode** soit suffisant (simplicité de débogage, d’utilisation, ...) ; nous allons coder cette partie en **pseudo-distribuée** pour utiliser HDFS et non pas local file system, et ainsi prendre un main l’API java associé à HDFS.

Pour cela, on exporte en JAR, avec les paramètres `input file - output file`.

Pour retracer l’historique des commandes, un fichier `bash` est mis à disposition. Le résultat de sortie est stocké dans le dossier *result*.