



Lancer Hadoop

En vous connectant à la machine virtuelle distante, vous allez avoir accès à l'environnement Hadoop.

Commençons par un `ls` afin de voir les différents scripts à notre disposition.

```
root@node175820-env-1839015-etudiant1:~  
login as: root  
root@node175820-env-1839015-etudiant1.sh1.hidora.com's password:  
Last login: Mon Jan 29 11:54:59 2024 from 10.101.0.9  
[root@node175820-env-1839015-etudiant1 ~]# ls  
bash_hadoop_master.sh  installation  stop_docker_digi.sh  
bash_hadoop_slave1.sh  lance_srv_slaves.sh  
bash_hadoop_slave2.sh  start_docker_digi.sh  
[root@node175820-env-1839015-etudiant1 ~]#
```

Avec la commande `cat` vous allez pouvoir analyser le contenu des scripts.

```
[root@node175820-env-1839015-etudiant1 ~]# cat start_docker_digi.sh  
docker start hadoop-master  
docker start hadoop-slave1  
docker start hadoop-slave2  
  
[root@node175820-env-1839015-etudiant1 ~]# cat stop_docker_digi.sh  
docker stop hadoop-master  
docker stop hadoop-slave1  
docker stop hadoop-slave2  
  
[root@node175820-env-1839015-etudiant1 ~]# cat lance_srv_slaves.sh  
docker exec hadoop-slave1 /bin/bash -c './services_hbase_thrift.sh'  
docker exec hadoop-slave2 /bin/bash -c './services_hbase_thrift.sh'  
  
[root@node175820-env-1839015-etudiant1 ~]# cat bash_hadoop_master.sh  
docker exec -it hadoop-master bash  
  
[root@node175820-env-1839015-etudiant1 ~]# cat bash_hadoop_slave1.sh  
docker exec -it hadoop-slave1 bash  
[root@node175820-env-1839015-etudiant1 ~]# cat bash_hadoop_slave2.sh  
docker exec -it hadoop-slave2 bash
```

Les noms explicites des scripts nous expliquaient déjà leur utilité, mais au moins maintenant nous en sommes certains, et on va pouvoir les utiliser.

On commence par exécuter le `./start_docker_digi.sh` afin de lancer le docker sur la machine.

```
[root@node175820-env-1839015-etudiant1 ~]# ./start_docker_digi.sh  
hadoop-master  
hadoop-slave1  
hadoop-slave2
```

Ensuite, on lance les services thrift des slaves avec `./lance_srv_slaves.sh`

```
[root@node175820-env-1839015-etudiant1 ~]# ./lance_srv_slaves.sh
hadoop-master: Warning: Permanently added 'hadoop-master,172.18.0.2' (ECDSA) to the list of known hosts.
hadoop-master: running zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-root-zookeeper-hadoop-master.out
running master, logging to /usr/local/hbase/logs/hbase--master-hadoop-slave1.out
OpenJDK 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
: running regionserver, logging to /usr/local/hbase/logs/hbase--regionserver-hadoop-slave1.out
: OpenJDK 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
: OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
running thrift, logging to /usr/local/hbase/logs/hbase--thrift-hadoop-slave1.out
hadoop-master: Warning: Permanently added 'hadoop-master,172.18.0.2' (ECDSA) to the list of known hosts.
hadoop-master: zookeeper running as process 83. Stop it first.
running master, logging to /usr/local/hbase/logs/hbase--master-hadoop-slave2.out
OpenJDK 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
: running regionserver, logging to /usr/local/hbase/logs/hbase--regionserver-hadoop-slave2.out
: OpenJDK 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
: OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
running thrift, logging to /usr/local/hbase/logs/hbase--thrift-hadoop-slave2.out
```

Puis pour pouvoir utiliser le master, on lancera `./bash_hadoop_master.sh` pour arriver sur le terminal du container hadoop-master. Réflexe à avoir, toujours un `ls` pour voir où l'on est et à quoi on a accès.

```
root@hadoop-master:~# ls
happybase.sh      hdfs          services_hbase_thrift.sh
hbase_create.sh   purchases.txt  setup.sh
hbase_drop.sh     purchases2.txt start-hadoop.sh
hbase_odbc_rest.sh run-wordcount.sh start-kafka-zookeeper.sh
```

Il y a un exemple d'utilisation avec le script `run-wordcount.sh`. Pour l'exécuter, lancez bien les services grâce à `./start-hadoop.sh`, puis vous pourrez le lancer avec `./run-wordcount.sh`

```

root@hadoop-master:~# ./start-hadoop.sh

Starting namenodes on [hadoop-master]
hadoop-master: Warning: Permanently added 'hadoop-master,172.18.0.2' (ECDSA) to the list of known hosts.
hadoop-master: starting namenode, logging to /usr/local/hadoop/logs/hadoop-root-namenode-hadoop-master.out
hadoop-slave1: Warning: Permanently added 'hadoop-slave1,172.18.0.3' (ECDSA) to the list of known hosts.
hadoop-slave2: Warning: Permanently added 'hadoop-slave2,172.18.0.4' (ECDSA) to the list of known hosts.
hadoop-slave1: starting datanode, logging to /usr/local/hadoop/logs/hadoop-root-datanode-hadoop-slave1.out
hadoop-slave2: starting datanode, logging to /usr/local/hadoop/logs/hadoop-root-datanode-hadoop-slave2.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: Warning: Permanently added '0.0.0.0' (ECDSA) to the list of known hosts.
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-root-secondarynamenode-hadoop-master.out

starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn--resourcemanager-hadoop-master.out
hadoop-slave2: Warning: Permanently added 'hadoop-slave2,172.18.0.4' (ECDSA) to the list of known hosts.
hadoop-slave1: Warning: Permanently added 'hadoop-slave1,172.18.0.3' (ECDSA) to the list of known hosts.
hadoop-slave2: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-root-nodemanager-hadoop-slave2.out
hadoop-slave1: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-root-nodemanager-hadoop-slave1.out

root@hadoop-master:~# ./run-wordcount.sh
mkdir: cannot create directory 'input': File exists
24/01/31 11:19:23 INFO client.RMPProxy: Connecting to ResourceManager at hadoop-master/172.18.0.2:8032
24/01/31 11:19:23 INFO input.FileInputFormat: Total input paths to process : 2
24/01/31 11:19:24 INFO mapreduce.JobSubmitter: number of splits:2
24/01/31 11:19:24 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1706699953857_0001
24/01/31 11:19:24 INFO impl.YarnClientImpl: Submitted application application_1706699953857_0001
24/01/31 11:19:24 INFO mapreduce.Job: The url to track the job: http://hadoop-master:8088/proxy/application_1706699953857_0001/
24/01/31 11:19:24 INFO mapreduce.Job: Running job: job_1706699953857_0001
24/01/31 11:19:30 INFO mapreduce.Job: Job job_1706699953857_0001 running in uber mode : false
24/01/31 11:19:30 INFO mapreduce.Job: map 0% reduce 0%
24/01/31 11:19:35 INFO mapreduce.Job: map 100% reduce 0%
24/01/31 11:19:40 INFO mapreduce.Job: map 100% reduce 100%
24/01/31 11:19:40 INFO mapreduce.Job: Job job_1706699953857_0001 completed successfully
24/01/31 11:19:40 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=56
    FILE: Number of bytes written=352386
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=258
    HDFS: Number of bytes written=26
    HDFS: Number of read operations=9
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters

```

À vous de jouer !

Soucis machines

Si vous avez un problème de configuration ou que vous avez cassé votre machine, allez dans le dossier "installation" grâce à `cd installation` et relancer le script `./install_cont_hadoop`.

```

[root@node175820-env-1839015-etudiant1 ~]# cd installation/
[root@node175820-env-1839015-etudiant1 installation]# ls
happybase.sh      hbase_drop.sh      install_cont_hadoop.sh  setup.sh
hbase_create.sh  hbase_odbc_rest.sh  services_hbase_thrift.sh
[root@node175820-env-1839015-etudiant1 installation]# █

```

Le processus prendra un peu de temps, donc soyez patient. Ne paniquez pas si le terminal bloque à l'installation de pandas. Vous pouvez partir en pause et prendre un café ☕ en attendant.

Si le problème ne se résout pas, prévenez nous ! ([Christophe Germain](#) ou [Robin Hotton](#))