

RESULTATS D'INSTALLATION DE SPARK

Parmi les différentes façons présentée pour installer spark nous avons choisi l'installation localement

Tout d'abord nous tenons à préciser que nous avons eu a également installer spark dans le livre sur scala. Mais nous allons recommencer le processus ici pour mieux expliquer.

Nous utilisons ici le système d'exploitation linux (ubuntu 20.04)

Les étapes sont les suivantes :

1. Télécharger Apache Spark
2. Extraire les fichiers du dossier compressé
3. déplacer le fichier extrait spark répertoire vers /opt et Configurer les variables d'environnement pour Spark

```
brice@wambdevps:~/spark$ sudo mv spark-3.3.1-bin-hadoop3 /opt/spark
brice@wambdevps:~/spark$ echo "export SPARK_HOME=/opt/spark" >> ~/.profile
brice@wambdevps:~/spark$ echo "export PATH=$PATH:/opt/spark/bin:/opt/spark/sbin" >> ~/.profile
brice@wambdevps:~/spark$ echo "export PYSPARK_PYTHON=/usr/bin/python3" >> ~/.profile
brice@wambdevps:~/spark$ source ~/.profile
brice@wambdevps:~/spark$ brice@wambdevps:~/spark$
bash: brice@wambdevps:~/spark$: No such file or directory
brice@wambdevps:~/spark$ ls -l /opt/spark
total 156
drwxr-xr-x 2 brice brice 4096 oct. 15 11:32 bin
drwxr-xr-x 2 brice brice 4096 oct. 15 11:32 conf
drwxr-xr-x 5 brice brice 4096 oct. 15 11:32 data
drwxr-xr-x 4 brice brice 4096 oct. 15 11:32 examples
drwxr-xr-x 2 brice brice 20480 oct. 15 11:32 jars
drwxr-xr-x 4 brice brice 4096 oct. 15 11:32 kubernetes
-rw-r--r-- 1 brice brice 22940 oct. 15 11:32 LICENSE
drwxr-xr-x 2 brice brice 4096 oct. 15 11:32 licenses
-rw-r--r-- 1 brice brice 57842 oct. 15 11:32 NOTICE
drwxr-xr-x 9 brice brice 4096 oct. 15 11:32 python
drwxr-xr-x 3 brice brice 4096 oct. 15 11:32 R
-rw-r--r-- 1 brice brice 4461 oct. 15 11:32 README.md
-rw-r--r-- 1 brice brice 165 oct. 15 11:32 RELEASE
drwxr-xr-x 2 brice brice 4096 oct. 15 11:32 sbin
drwxr-xr-x 2 brice brice 4096 oct. 15 11:32 yarn
```

4. Démarrer le serveur maître Apache Spark


```
brice@wambdevps:~/spark$ start-master.sh --port 7072 --webui-port 8082
starting org.apache.spark.deploy.master.Master, logging to /opt/spark/logs/spark-brice-org.apache.spark.deploy.master.Master-1-wambdevps.out
```

5. Exécuter le script de travail esclave

```
brice@wambdevps:~/spark$ start-worker.sh spark://wambdevps:7072
starting org.apache.spark.deploy.worker.Worker, logging to /opt/spark/logs/spark-brice-org.apache.spark.deploy.worker.Worker-1-wambdevps.out
```

6. Accéder à Spark Master avec <http://127.0.0.1:8082>

← → ↻ 127.0.0.1:8082

 **Spark Master at spark://wambdevps:7072**

URL: spark://wambdevps:7072
Alive Workers: 1
Cores in use: 4 Total, 0 Used
Memory in use: 6.6 GiB Total, 0.0 B Used
Resources in use:
Applications: 0 Running, 0 Completed
Drivers: 0 Running, 0 Completed
Status: ALIVE

▼ Workers (1)

Worker Id	Address	State	Cores	Memory	Resources
worker-20221031160012-192.168.100.7-38343	192.168.100.7:38343	ALIVE	4 (0 Used)	6.6 GiB (0.0 B Used)	

▼ Running Applications (0)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------

▼ Completed Applications (0)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------