

# Hadoop Single Node

## Start a Ubuntu VM

create a ubuntu vm on google cloud engine

## Installation

### 1.Install Java

```
sudo apt update
sudo apt install default-jdk
java -version
```

### 2. Install Spark

```
wget https://downloads.apache.org/spark/spark-3.2.4/spark-3.2.4-bin-hadoop3.2.tgz
tar -zxvf spark-3.2.4-bin-hadoop3.2.tgz
```

## Start Spark

```
cd spark-3.2.4-bin-hadoop3.2
./sbin/start-master.sh
```

response should be similar to : " starting org.apache.spark.deploy.master.Master, logging to /home/user/spark-2.3.4-bin-hadoop2.7/logs/spark-user-org.apache.spark.deploy.master.Master-1-hostname.out"

Check the log file

```
nano logs/spark-clarazhang0625-org.apache.spark.deploy.master.Master-1-spark.out
```

```
Spark Command: /usr/lib/jvm/java-11-openjdk-amd64/bin/java -cp /home/clarazhang0625/spark-3.2.4-bin-hadoop3.2/conf/:/home/clarazhang
=====
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
23/07/09 17:29:30 INFO Master: Started daemon with process name: 5373@spark
23/07/09 17:29:30 INFO SignalUtils: Registering signal handler for TERM
23/07/09 17:29:30 INFO SignalUtils: Registering signal handler for HUP
23/07/09 17:29:30 INFO SignalUtils: Registering signal handler for INT
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.spark.unsafe.Platform (file:/home/clarazhang0625/spark-3.2.4-bin-hadoop3.2/jars/spa
WARNING: Please consider reporting this to the maintainers of org.apache.spark.unsafe.Platform
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
23/07/09 17:29:31 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
23/07/09 17:29:31 INFO SecurityManager: Changing view acls to: clarazhang0625
23/07/09 17:29:31 INFO SecurityManager: Changing modify acls to: clarazhang0625
23/07/09 17:29:31 INFO SecurityManager: Changing view acls groups to:
23/07/09 17:29:31 INFO SecurityManager: Changing modify acls groups to:
23/07/09 17:29:31 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Se
23/07/09 17:29:32 INFO Utils: Successfully started service 'sparkMaster' on port 7077.
23/07/09 17:29:32 INFO Master: Starting Spark master at spark://spark.us-central1-a.c.solar-melody-349123.internal:7077
23/07/09 17:29:32 INFO Master: Running Spark version 3.2.4
23/07/09 17:29:32 INFO Utils: Successfully started service 'MasterUI' on port 8080.
23/07/09 17:29:32 INFO MasterWebUI: Bound MasterWebUI to 0.0.0.0, and started at http://spark.us-central1-a.c.solar-melody-349123.in
23/07/09 17:29:32 INFO Master: I have been elected leader! New state: ALIVE
23/07/09 17:45:13 INFO Master: Registering worker 10.128.0.30:34655 with 2 cores, 2.8 GiB RAM
23/07/09 17:59:11 ERROR Master: RECEIVED SIGNAL TERM
```

```
nano logs/spark-clarazhang0625-org.apache.spark.deploy.master.Master-1-spark.out
```

find the master url, e.g. spark://spark.us-central1-a.c.solar-melody-349123.internal:7077

```
./sbin/start-slave.sh spark://spark.us-central1-a.c.solar-melody-349123.internal:7077
```

Similar to the previous step check the log file

```
GNU nano 4.8 spark-3.2.4-bin-hadoop3.2/logs/spark-clarazhang0625-org.apache.spark.de
Spark Command: /usr/lib/jvm/java-11-openjdk-amd64/bin/java -cp /home/clarazhang0625/spark-3.2.4-bin-hadoop3.2/conf/:/home/clarazhang
=====
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
23/07/09 17:45:10 INFO Worker: Started daemon with process name: 5507@spark
23/07/09 17:45:10 INFO SignalUtils: Registering signal handler for TERM
23/07/09 17:45:10 INFO SignalUtils: Registering signal handler for HUP
23/07/09 17:45:10 INFO SignalUtils: Registering signal handler for INT
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.spark.unsafe.Platform (file:/home/clarazhang0625/spark-3.2.4-bin-hadoop3.2/jars/spa
WARNING: Please consider reporting this to the maintainers of org.apache.spark.unsafe.Platform
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
23/07/09 17:45:11 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
23/07/09 17:45:11 INFO SecurityManager: Changing view acls to: clarazhang0625
23/07/09 17:45:11 INFO SecurityManager: Changing modify acls to: clarazhang0625
23/07/09 17:45:11 INFO SecurityManager: Changing view acls groups to:
23/07/09 17:45:11 INFO SecurityManager: Changing modify acls groups to:
23/07/09 17:45:11 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Se
23/07/09 17:45:12 INFO Utils: Successfully started service 'sparkWorker' on port 34655.
23/07/09 17:45:12 INFO Worker: Worker decommissioning not enabled.
23/07/09 17:45:12 INFO Worker: Starting Spark worker 10.128.0.30:34655 with 2 cores, 2.8 GiB RAM
23/07/09 17:45:12 INFO Worker: Running Spark version 3.2.4
23/07/09 17:45:12 INFO Worker: Spark home: /home/clarazhang0625/spark-3.2.4-bin-hadoop3.2
23/07/09 17:45:12 INFO ResourceUtils: =====
23/07/09 17:45:12 INFO ResourceUtils: No custom resources configured for spark.worker.
23/07/09 17:45:12 INFO ResourceUtils: =====
23/07/09 17:45:12 INFO Utils: Successfully started service 'WorkerUI' on port 8081.
23/07/09 17:45:12 INFO WorkerWebUI: Bound WorkerWebUI to 0.0.0.0, and started at http://spark.us-central1-a.c.solar-melody-349123.in
23/07/09 17:45:12 INFO Worker: Connecting to master spark.us-central1-a.c.solar-melody-349123.internal:7077...
23/07/09 17:45:13 INFO TransportClientFactory: Successfully created connection to spark.us-central1-a.c.solar-melody-349123.internal
23/07/09 17:45:13 INFO Worker: Successfully registered with master spark://spark.us-central1-a.c.solar-melody-349123.internal:7077
```

Then start the spark

```
/home/clarazhang0625/spark-3.2.4-bin-hadoop3.2/bin/spark-shell
```

The response should be:

