

تعليمات CMD :

تنصيب Java 11 :

java -version

المخرجات :

java version "21-10-2025" 25.0.1 LTS
Java(TM) SE Runtime Environment (build 8+25.0.1-LTS-(27
Java HotSpot(TM) 64-Bit Server VM (build 8+25.0.1-LTS-27, mixed mode, sharing)

تنصيب Maven :

mvn -v

المخرجات :

Apache Maven 3) 3.9.11e54c93a704957b63ee3494413a2b544fd3d825b)
Maven home: D:\apache-maven-3.9.11
Java version: 25.0.1, vendor: Oracle Corporation, runtime: D:\tp7 mallloul
Default locale: en_US, platform encoding: UTF-8
OS name: "windows 10", version: "10.0", arch: "amd64", family: "windows"

تنصيب Docker Desktop :

docker --version

المخرجات :

Docker version 29.1.3, build f52814d
docker rm -f spark-master spark-worker 2>nul || true

المخرجات :

spark-master
spark-worker

docker compose pull spark-master spark-worker

المخرجات :

[+] Pulling 2/2
spark-worker Skipped - Image is already being pulled by spark-master
spark-master Pulled

docker compose up -d --force-recreate spark-master spark-worker

المخرجات :

[+] Running 5/5
Container hadoop-resourcemanager Running
Container spark-master Started
Container spark-worker Started
Container hadoop-master Started
Container hadoop-datanode Started

تشغيل Hadoop + Spark (Docker) :

cd D:\hadoop-spark
docker compose up -d

المخرجات :

[+] Running 3/3
Container hadoop-master Started

Container spark-master Started
Container spark-worker Started

: **docker ps**

المخرجات :

```
spark-master      Up
spark-worker      Up
hadoop-master     Up
hadoop-worker1    Up
hadoop-worker2    Up
```

: **WordCount Batch (Spark + Hadoop)**

: **رفع ملف الى HDFS**

```
docker exec -it hadoop-master hdfs dfs -put /data/file1.txt /
docker exec -it hadoop-master hdfs dfs -ls /
```

المخرجات :

```
-rw-r--r--  1 root supergroup      45 file1.txt
```

: **تشغيل Spark Batch**

```
docker exec -it spark-master /opt/spark/bin/spark-submit ^
--class spark.batch.WordCountTask ^
/opt/spark/work-dir/wordcount-spark.jar ^
hdfs://hadoop-master:9000/file1.txt ^
hdfs://hadoop-master:9000/wc-output
```

المخرجات :

```
INFO SparkContext: Running Spark version 3.5.0
INFO SparkContext: Submitted application: WordCountTask
INFO SparkContext: Successfully stopped SparkContext
```

عرض النتائج :

```
docker exec -it hadoop-master hdfs dfs -cat /wc-output/part-00000
```

المخرجات :

```
Hello,2
Spark,1
Hadoop,1
```

: **الدخول الى حاوية Hadoop**

```
docker exec -it hadoop-master bash
root@hadoop-master:/#
```

: **أوامر HDFS**

: **التحقق من حالة HDFS**

```
hdfs dfsadmin -report
```

المخرجات :

```
Configured Capacity: ...
Present Capacity: ...
DFS Remaining: ...
Live datanodes1 :
Name: hadoop-worker1
```

أوامر YARN :

التحقق من عمل YARN :

yarn node -list

المخرجات :

Total Nodes:1

Node-Id	Node-State
hadoop-nodemanager	RUNNING

عرض التطبيقات المشغلة على YARN :

Yarn application - list

المخرجات :

Total number of applications (0)

تشغيل تطبيق Spark على YARN :

```
/opt/spark/bin/spark-submit \  
--master yarn \  
--deploy-mode client \  
--class spark.batch.WordCountTask \  
/opt/spark/work-dir/wordcount-spark.jar \  
hdfs://hadoop-master:9000/data/file1.txt \  
hdfs://hadoop-master:9000/wc-yarn-output
```

التحقق من التطبيق في YARN :

Yarn application - list

المخرجات :

Application-Id	Application-Name	State	Final-State
application_...	WordCountTask	FINISHED	SUCCEEDED

عرض نتائج WordCount من HDFS :

```
hdfs dfs -ls /wc-yarn-output  
hdfs dfs -cat /wc-yarn-output/part-00000
```

المخرجات :

Spark,1 Streaming,1

انشاء ملف Streaming :

D:\hadoop-spark\streaming.py

نسخ الملف الى Spark Container :

```
docker cp D:\hadoop-spark\streaming.py spark-master:/opt/spark/work-dir/streaming.py  
docker exec -it spark-master ls /opt/spark/work-dir
```

المخرجات :

```
streaming.py  
wordcount-spark.jar  
wordcount-streaming.jar
```

تشغيل Socket Server :

نافذة 1 :

nc -lk 9999

تشغيل (Docker) Spark Streaming :

نافذة 2 :

```
cd /opt/spark/work-dir  
/opt/spark/bin/spark-submit streaming.py
```

مخرجات :

```
INFO SparkContext: Running Spark version 3.5.0  
INFO SparkContext: Submitted application: StructuredStreamingWordCount  
.....
```

ثم في النافذة 1 مجدداً :

```
hello spark  
hello streaming  
spark test
```

النتائج التي تظهر في النافذة 2 :

word	count
hello	2
spark	2
streaming	1
test	1