

# Máster en Big Data

---

## Tecnologías de Almacenamiento

### 7. Hands-On: Ejecución Apache Spark

## Índice

1. Introducción .....	3
2. Entorno .....	3
3. Ejecución Spark .....	3
4. SparkUI .....	<b>¡Error! Marcador no definido.</b>

## 1. Introducción

El objetivo de este Hands-On es aprender a ejecutar un job de Spark mediante un ejemplo ya precompilado

## 2. Entorno

Para la realización de este hands-on, vamos a desplegar un sandbox de MapR en VMWare.

Una vez arrancado el nodo, las credenciales son:

Usuario: mapr

Password: mapr

## 3. Ejecución Spark

La ubicación de Spark dentro de la distribución MapR la podemos encontrar en:

`/opt/mapr/spark/spark-2.1.0/`

Se pide ejecutar el programa SparkPi, contra yarn en modo cliente, con una memoria por ejecutor de 1g.

El jar ya compilado lo podemos encontrar en: `/opt/mapr/spark/spark-2.1.0/examples/jars/spark-examples_2.11-2.1.0-mapr-1710.jar`

Y el package y nombre de la clase es: `org.apache.spark.examples.SparkPi`

El programa acepta por parámetro el número de paralelización que queremos darle, vamos a pasarle, por ejemplo, 4

Engancha aquí el comando para enviar el Job y una captura de pantalla con la salida

```
/opt/mapr/spark/spark-2.1.0/bin/spark-submit --class org.apache.spark.examples.SparkPi --master yarn --num-executors 8 --deploy-mode client /opt/mapr/spark/spark-2.1.0/examples/jars/spark-examples_2.11-2.1.0-mapr-1710.jar 4
```