Instalación de Docker. Generar un docker-spark-cluster

Néstor Batista Díaz

INSTALAR DOCKER

INSTALAR EL CLUSTER DE SPARK
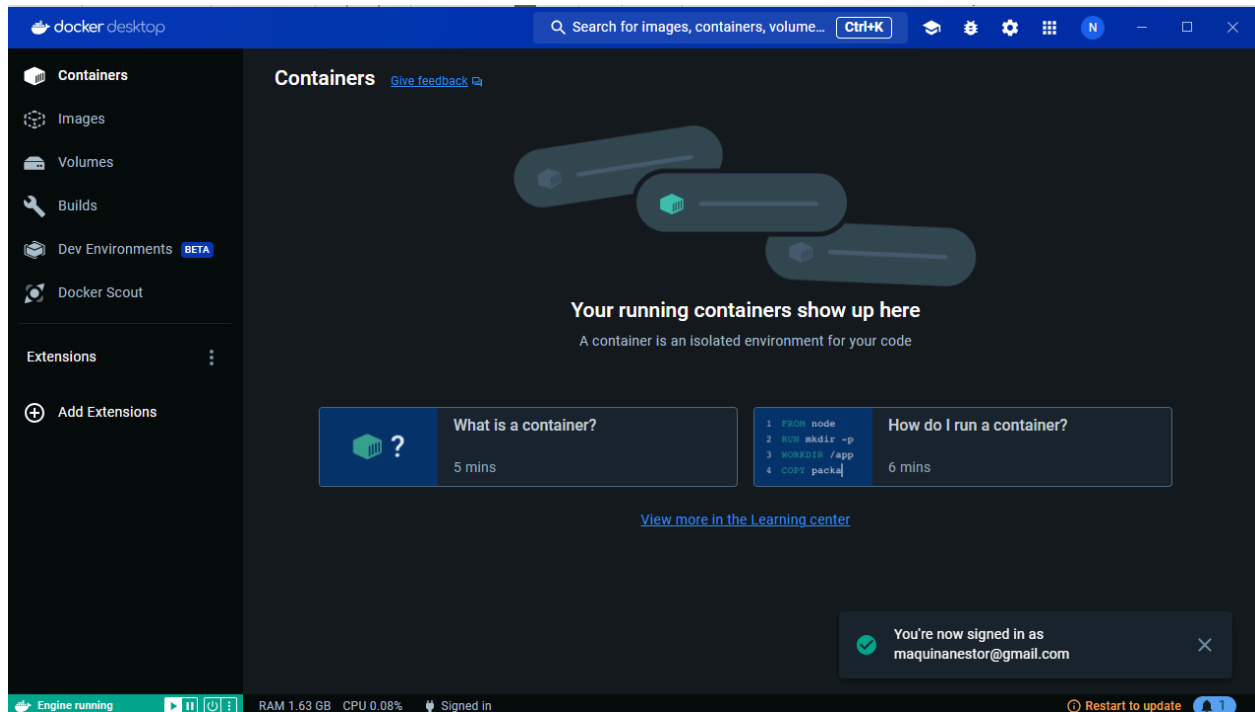
Descargamos el proyecto de github

Descomprimimos

Abrimos CMD y ejecutamos el siguiente código en la carpeta del proyecto

Build the image

```
docker build -t cluster-apache-spark:3.0.2 .
```

Ejecutamos Docker compose

Comprobamos que funciona

☐ ⌄ 🗇 **docker-spark-cluster-master**			Running (4/4)	0.31%		6 minutes ago	◼ ⋮ 🗑
☐ 📦 **demo-database-1** eb64b000e0be 📋	postgres:11.7-alpine		Running	0%	5432:5432 ⧉	6 minutes ago	◼ ⋮ 🗑
☐ 📦 **spark-master-1** 08ec0e71104d 📋	cluster-apache-spark:3.0.2		Running	0.09%	7077:7077 ⧉ 9090:8080 ⧉ **Show less**	6 minutes ago	◼ ⋮ 🗑
☐ 📦 **spark-worker-a-1** 7f9537b14e3f 📋	cluster-apache-spark:3.0.2		Running	0.1%	7000:7000 ⧉ 9091:8080 ⧉ **Show less**	6 minutes ago	◼ ⋮ 🗑
☐ 📦 **spark-worker-b-1** 649117c0d67c 📋	cluster-apache-spark:3.0.2		Running	0.12%	7001:7000 ⧉ 9092:8080 ⧉ **Show less**	6 minutes ago	◼ ⋮ 🗑

← → C 🛡 ⌄ localhost:9090   Aᴬ 🔖 ⌄ | Q⁺ Buscar en Bing ⌄

Spark 3.0.2   **Spark Master at spark://08ec0e71104d:7077**

URL: spark://08ec0e71104d:7077
Alive Workers: 2
Cores in use: 2 Total, 0 Used
Memory in use: 2.0 GiB Total, 0.0 B Used
Resources in use:
Applications: 0 Running, 0 Completed
Drivers: 0 Running, 0 Completed
Status: ALIVE

▾ Workers (2)

Worker Id	Address	State	Cores	Memory	Resources
worker-20240416172524-172.18.0.4-7000	172.18.0.4:7000	ALIVE	1 (0 Used)	1024.0 MiB (0.0 B Used)	
worker-20240416172524-172.18.0.5-7000	172.18.0.5:7000	ALIVE	1 (0 Used)	1024.0 MiB (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