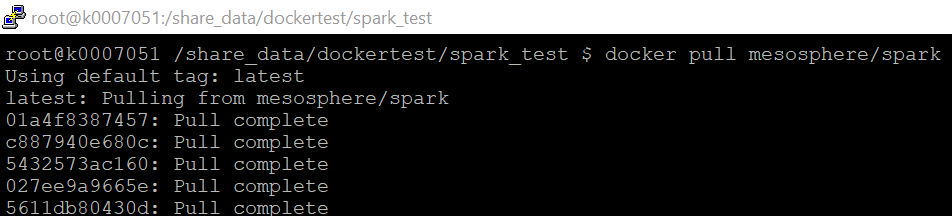
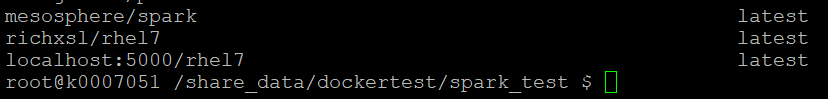
**Running Scala application on Docker image then to Kubernetes**

**docker pull mesosphere/spark**



**docker images or docker images -a**



root@k0007051 /share\_data/dockertest/spark\_test $ docker images | grep mesosphere/spark

mesosphere/spark latest 5c25c7985707 11 months ago 1.42GB

docker build -f Dockerfile.for.spark\_test -t k0007051:5000/spark-test .

docker run -v /share\_data:/share\_data k0007051:5000/spark-test

**Dockerfile.for.spark\_test**

/share\_data/dockertest/spark\_test $ nano Dockerfile.for.spark\_test

# Use an official Spark/Scala runtime as a parent image

#FROM scala/spark

FROM richxsl/rhel7

FROM mesosphere/spark

# Set the working directory to /test

#WORKDIR /share\_data/dockertest/spark\_test

# Set up the predefined perm mounts

#VOLUME /share\_data/dockertest/spark\_test

# Copy the current directory contents into the container at /share\_data

#ADD . /share\_data/dockertest/spark\_test

# Install any needed packages specified in requirements.txt

#RUN pip install --trusted-host pypi.python.org -r requirements.txt

# Define environment variable

ENV NAME World

#ENV JAVA\_HOME /bin/java

ENV SPARK\_HOME /opt/spark/dist

ENV SCALA\_HOME /share\_data/scala-2.11.8

ENV PATH $PATH:${JAVA\_HOME}/bin:${SPARK\_HOME}/bin:${SCALA\_HOME}/bin

# Run spark-submit and the jar file when the container launches

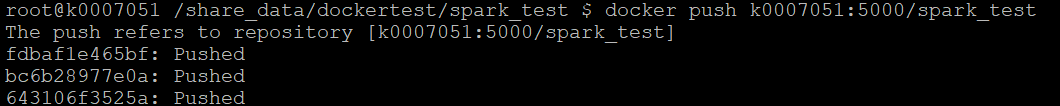
CMD ["spark-submit", "/share\_data/sample1\_2.11-2.3.2.jar"]

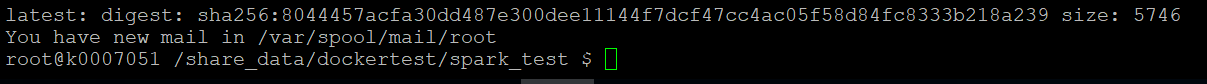
#CMD ["/bin/bash"]

docker build -f Dockerfile.for.spark\_test -t k0007051:5000/spark-test .

docker run -v /share\_data:/share\_data k0007051:5000/spark-test

docker push k0007051:5000/spark-test





**Run Docker image k0007051:5000/spark-test on Kubernetes**

Create the YAML file

root@k0007051 /share\_data/dockertest/spark\_test $ cat spark\_test.yaml

spark\_test.yaml

apiVersion: v1

kind: Pod

metadata:

name: spark-test

spec:

containers:

- image: k0007051:5000/spark-test

name: spark-test

volumeMounts:

- name: share-data

mountPath: /share\_data

volumes:

- name: share-data

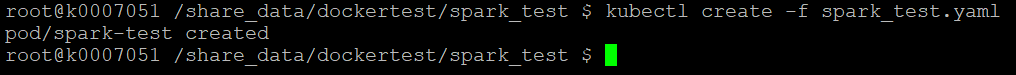
hostPath:

path: /share\_data

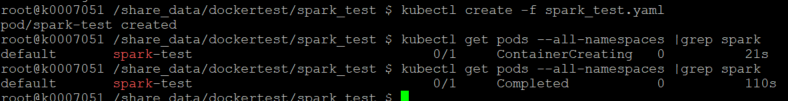
type: Directory

restartPolicy: Never

kubectl create -f spark\_test.yaml



kubectl get pods --all-namespaces |grep spark



kubectl logs spark-test | head



kubectl delete pod spark-test

**Troubleshoot notes:**

1. To find the spark path

docker start k0007051:5000/spark-test

docker attach k0007051:5000/spark-test

You are now in the Docker container. Look for Spark path

Spark home should be: SPARK\_HOME= /opt/spark/dist

<Put this in SPARK\_HOME in Dockerfile.for.spark\_test>

1. To save docker images

docker save -i k0007051:5000/spark-test

1. To load docker images

docker load k0007051:5000/spark-test /share\_data

1. List the docker images

docker ps -a |grep spark