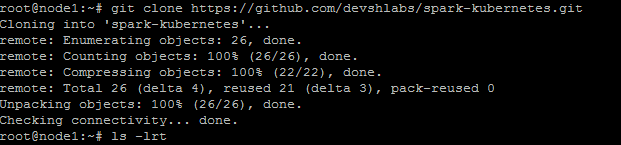
Kuberent-spark

<https://developer.sh/posts/spark-kubernetes-guide>

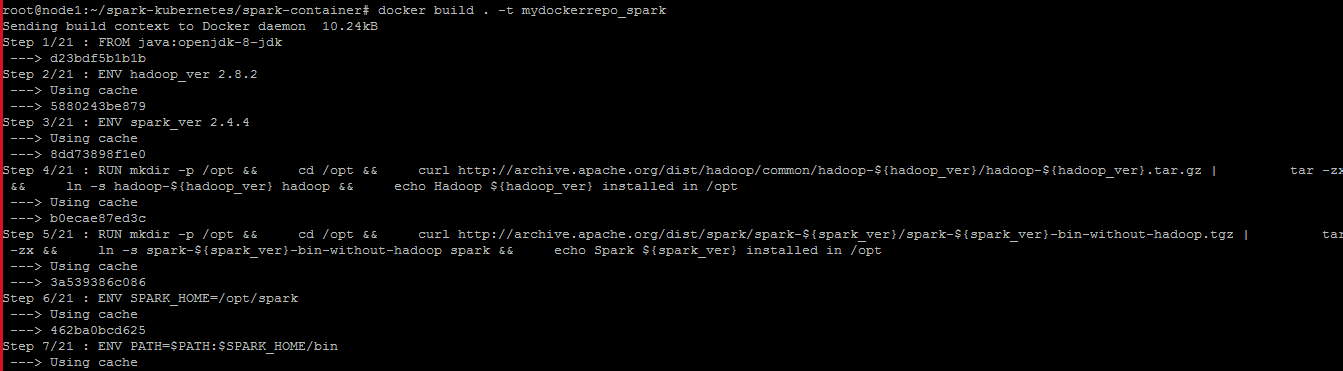
git clone <https://github.com/devshlabs/spark-kubernetes.git>



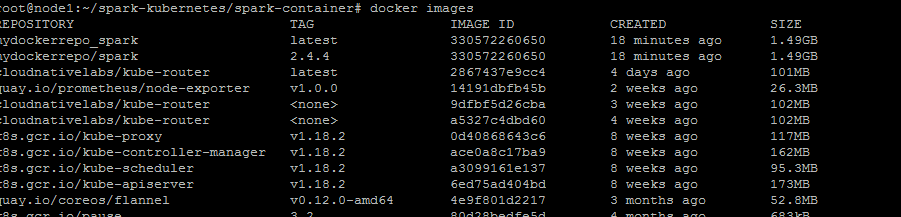
cd spark-kubernetes/

cd spark-container/

docker build . -t mydockerrepo\_spark

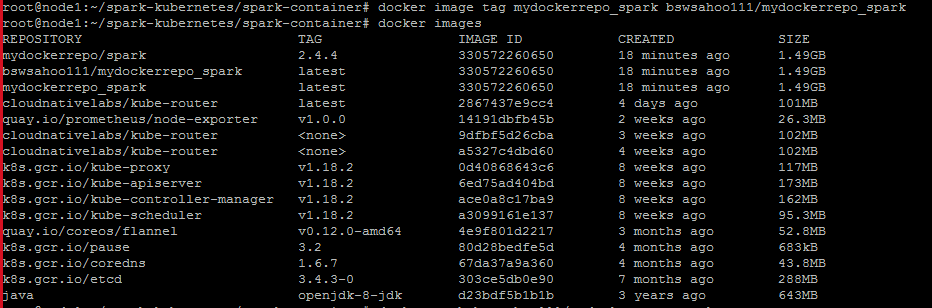


docker images



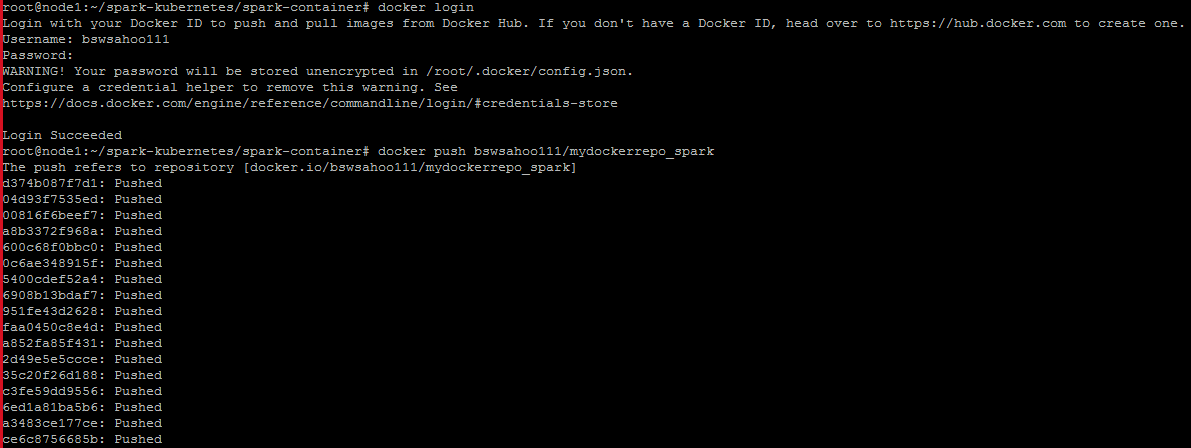
docker image tag mydockerrepo\_spark bswsahoo111/mydockerrepo\_spark

docker images



docker login

docker push bswsahoo111/mydockerrepo\_spark

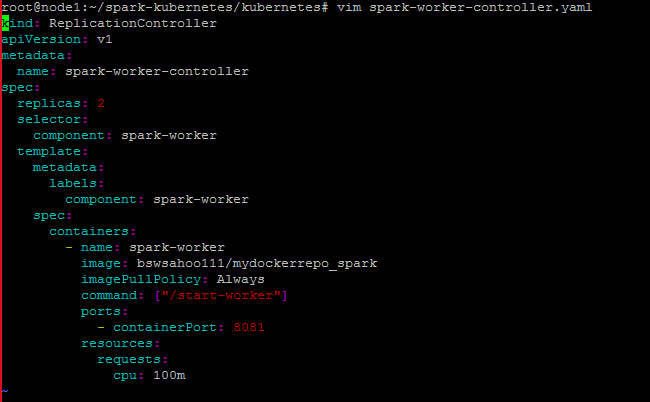


cd spark-kubernetes/kubernetes

ls –lrt

vim spark-worker-controller.yaml

change the image name to bswsahoo111/mydockerrepo\_spark

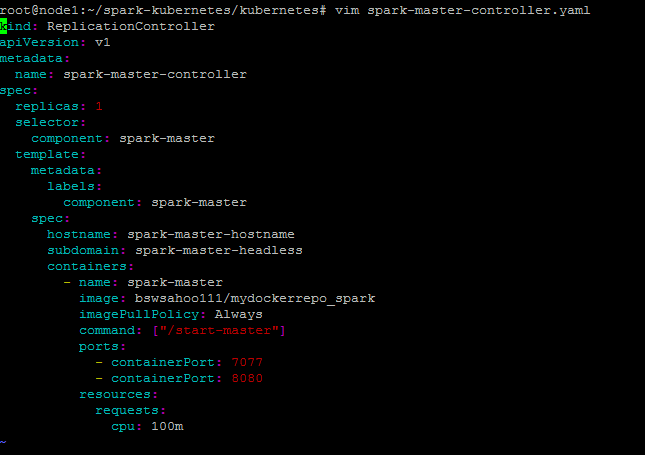


Vim spark-master-controller.yaml

Changed the image to bswsahoo111/mydockerrepo\_spark and removed below line

hostname: spark-master-hostname

subdomain: spark-master-headless

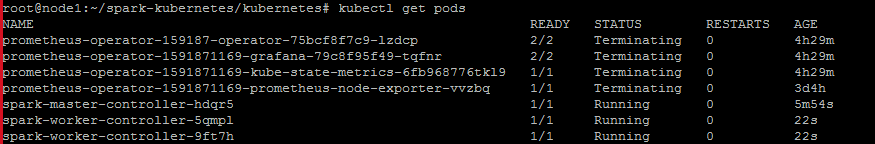


kubectl apply -f spark-worker-controller.yaml

kubectl apply -f spark-master-service.yaml

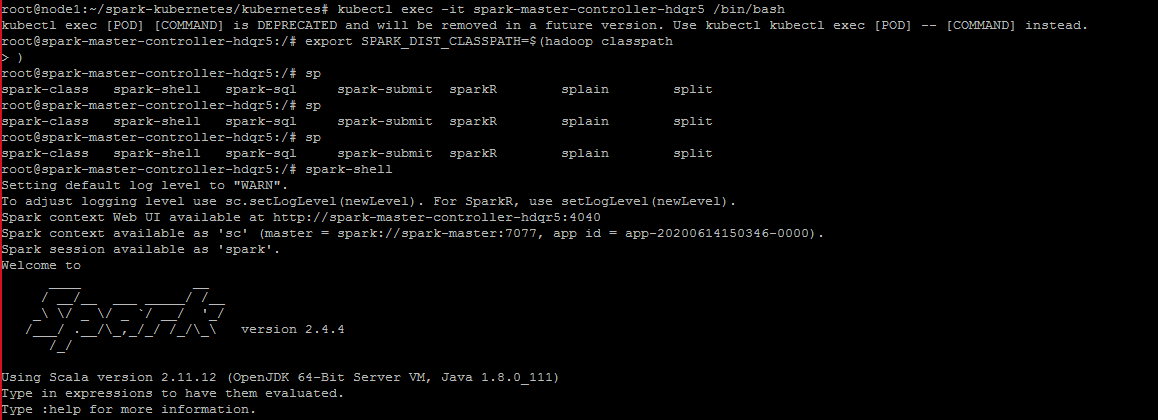
kubectl apply -f spark-master-controller.yaml

kubectl get pods



kubectl exec -it spark-master-controller-hdqr5 /bin/bash

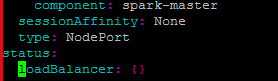
export SPARK\_DIST\_CLASSPATH=$(hadoop classpath)



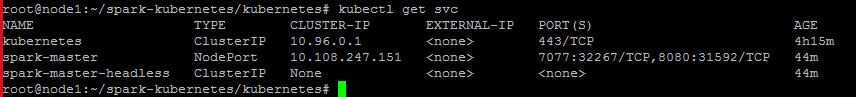
To open the public IP for spark

kubectl edit svc spark-master

Replace ClusterIp to Nodeport like below



kubectl get svc



<http://130.211.222.98:31592/>

