



Exercise 4.3: Basic Node Maintenance

In this section we will cause some of our pods to be evicted from a node and rescheduled elsewhere. This could be part of basic maintenance or a rolling OS update.

1. Create a deployment, then scale to create plenty of pods.

```
student@lfs458-node-1a0a:~$ kubectl create deployment maint --image=nginx
deployment.apps/maint created

student@lfs458-node-1a0a:~$ kubectl scale deployment maint --replicas=20
deployment.apps/maint scaled
```

2. Use the terminal on the second node to get a count of the current docker containers. It should be something like 29, plus a line for status counted by **wc**. The main system should have something like 25 running, plus a line of status. Your numbers may be different.

```
student@lfs458-node-1a0a:~$ sudo docker ps | wc -l
26

student@lfs458-worker:~$ sudo docker ps | wc -l
30
```

3. In order to complete maintenance we may need to move containers from a node and prevent new ones from deploying. One way to do this is to **drain**, or **cordon**, the node. Currently this will not affect DaemonSets, an object we will discuss in greater detail in the future. Begin by getting a list of nodes. Your node names will be different.

```
student@lfs458-node-1a0a:~$ kubectl get nodes
```

| NAME | STATUS | ROLES | AGE | VERSION |
|------------------|--------|--------|-----|---------|
| lfs458-worker | Ready | <none> | 4h | v1.17.1 |
| lfs458-node-1a0a | Ready | master | 4h | v1.17.1 |

4. Modifying your second, worker node, update the node to **drain** the pods. Some resources may not drain, expect an error which we will work with next. Note the error includes **aborting command** which indicates the drain did not take place. Were you to check it would have the same number of containers running, but will show a new taint preventing the scheduler from assigning new pods.

```
student@lfs458-node-1a0a:~$ kubectl drain lfs458-worker
node/lfs458-worker cordoned
error: unable to drain node "lfs458-worker", aborting command...

There are pending nodes to be drained:
  lfs458-worker
error: DaemonSet-managed pods (use --ignore-daemonsets to ignore):
calico-node-vndn7, kube-proxy-rjjsl

student@lfs458-node-1a0a:~$ kubectl describe node |grep -i taint
```

| | |
|---------|---------------------------------------------|
| Taints: | <none> |
| Taints: | node.kubernetes.io/unschedulable:NoSchedule |

5. As the error output suggests we can use the **--ignore-daemonsets** options to ignore containers which are not intended to move. We will find a new error when we use this command, near the end of the output. The node will continue to have the same number of pods and containers running.

```
student@lfs458-node-1a0a:~$ kubectl drain lfs458-worker --ignore-daemonsets
node/lfs458-worker already cordoned
WARNING: ignoring DaemonSet-managed Pods: kube-system/calico-node-dr279, kube-system/kube-proxy-qhc4f
evicting pod "maint-5bcc549596-b66fr"
evicting pod "maint-5bcc549596-2h5zg"
evicting pod "nginx-85ff79dd56-8h72c"
<output_omitted>
pod/maint-5bcc549596-xfbb6 evicted
pod/maint-5bcc549596-zmpdq evicted
pod/maint-5bcc549596-p2nnm evicted
node/lfs458-worker evicted
```

6. Run the command again. This time the output should both indicate the node has already been cordoned, then show the node has been drained. Not all pods will be gone as daemonsets will remain. Note the command is shown on two lines. You can omit the backslash and type the command on a single line.

```
student@lfs458-node-1a0a:~$ kubectl drain lfs458-worker \
--ignore-daemonsets --delete-local-data
node/lfs458-worker already cordoned
WARNING: ignoring DaemonSet-managed Pods: kube-system/calico-node-dr279, kube-system/kube-proxy-qhc4f
node/lfs458-worker drained
```

7. Were you to look on your second, worker node, you would see there should be fewer pods and containers than before. These pods can only be evicted via a special taint which we will discuss in the scheduling chapter.

```
student@lfs458-worker:~$ sudo docker ps | wc -l
5
```

8. Update the node taint such that the scheduler will use the node again. Verify that no nodes have moved over to the worker node as the scheduler only checks when a pod is deployed.

```
student@lfs458-node-1a0a:~$ kubectl uncordon lfs458-worker
node/lfs458-worker uncordoned

student@lfs458-node-1a0a:~$ kubectl describe node |grep -i taint
Taints:          <none>
Taints:          <none>
```

```
student@lfs458-worker:~$ sudo docker ps | wc -l
5
```

```
student@lfs458-node-1a0a:~$ sudo docker ps | wc -l
61
```

9. Clean up by deleting the maint deployment.

```
student@lfs458-worker:~$ kubectl delete deployment maint
deployment.apps "maint" deleted
```