

Oracle Container Engine for Kubernetes

Level 200

Jamal Arif Oracle Cloud Infrastructure October, 2019



Safe harbor statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.

The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Objectives

After completing this lesson, you should be able to understand:

- OKE Custom Create
- Bastion Host for Private Clusters
- Using OCI LB with OKE
- Persistent Storage for OKE
 - Using OCI Block Volume
 - Using OCI File System Storage

Pre-requisites: Container Engine (OKE) 100 Level



OKE 'Custom Create' in OCI

Pre-requisites for creating a K8s Cluster

- Monthly universal Credits have limit of 3 clusters per OCI region with 1000 nodes in a cluster and Pay-asyou-go or Promo accounts have a limit for One Cluster (by default).
- Must also have compute Instance Quota (Required) to launch k8s worker nodes in an AD or across ADs for HA
- Required Policy in the root compartment of your tenancy
 allow service OKE to manage all-resources in tenancy
- To launch a K8s cluster, user must be either part of the Admin group or a group to which a policy grants the appropriate Container Engine for Kubernetes permissions.
- Policies can be created for users which are not part of the admin group
- For Example: To enable users in group 'dev-team' to perform any operation on cluster-related resources → allow group dev-team to manage cluster-family in tenancy

Note: Polices must also grant the group 'dev-team' Networking permissions of VCN_READ and VCN_CREATE, SUBNET_READ and SUBNET_CREATE, COMPARTMENT_INSPECT, INTERNET_GATEWAY_CREATE, NAT_GATEWAY_CREATE, ROUTE_TABLE_UPDATE, SECURITY_LIST_CREATE: Details here

Custom Create Pre-reqs: Basic Virtual Cloud Network Config

- An Existing VCN with following
 - Internet Gateway
 - Nat Gateway (Private Worker Nodes)
 - Route table
 - with default route to IGW for Load balancer subnets and public worker node subnets
 - with default route to NAT GW for private worker node subnets
 - K8s worker node subnets
 - three public AD specific subnets in different ADs or a regional subnet for High Availability of public workers
 - three private AD specific subnets in different ADs or a regional subnet for *High Availability* of private workers
 - LBs Subnets A regional public subnet for OCI public load balancer
 - Separate Security Lists for K8s worker nodes subnets and LB subnets

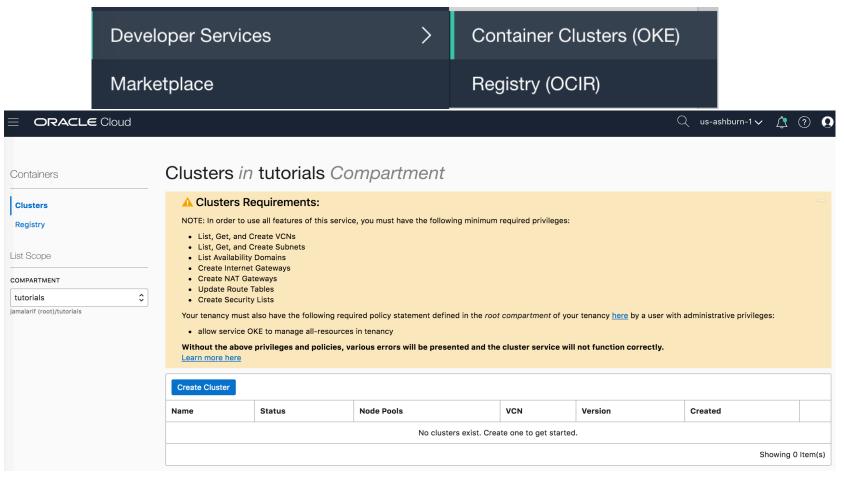
Custom Create Pre-reqs: Basic Virtual Cloud Network Config

Security Lists for K8s worker nodes subnets

- Stateless ingress and egress rules that allow all traffic between the different worker node subnets
- stateless ingress and egress rules that allow all traffic between worker node subnets and load balancer subnets
- an egress rule that allows all outbound traffic to the internet
- For PUBLIC worker nodes only ingress rules to allow Container Engine for Kubernetes to access worker nodes on port 22 from CIDR blocks mentioned <u>here</u>.

OKE Custom Create

Step 1: Navigate to Menu → Developer Services → Container Clusters (OKE) → Create Cluster



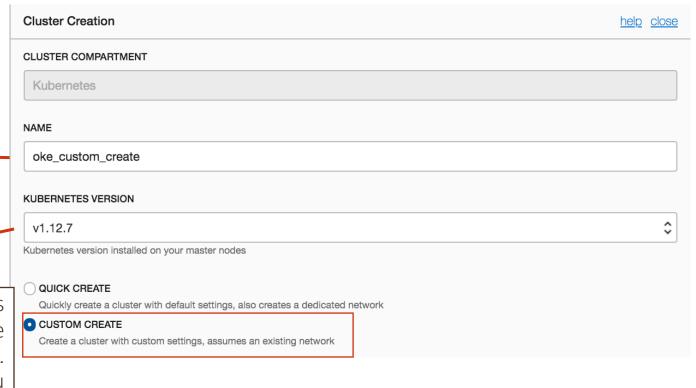


OKE Custom Create

Step 2: Cluster Creation

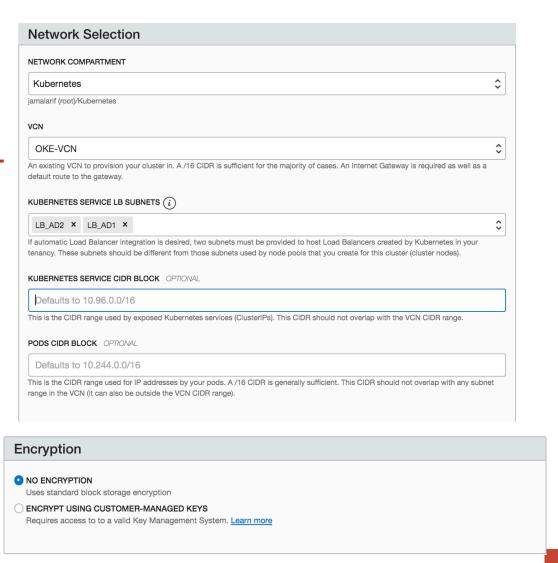
Name of the Cluster

The version of Kubernetes to run on the master nodes and worker nodes of the cluster. Either accept the default version or select a version of your choice. Amongst other things, the Kubernetes version you select determines the default set of admission controllers that are turned on in the created cluster (the set follows the recommendation given in the <u>Kubernetes documentation</u> for that version).



Step 2: Cluster Creation

- Provide VCN you have created/already available for OKE.
- Provide LB subnets (public subnets)
- Service CIDR blocks and PODs CIDR block are optional and chosen by default.
- Specify whether to encrypt Kubernetes secrets at rest in the etcd key-value store for the cluster using the Key Management service.



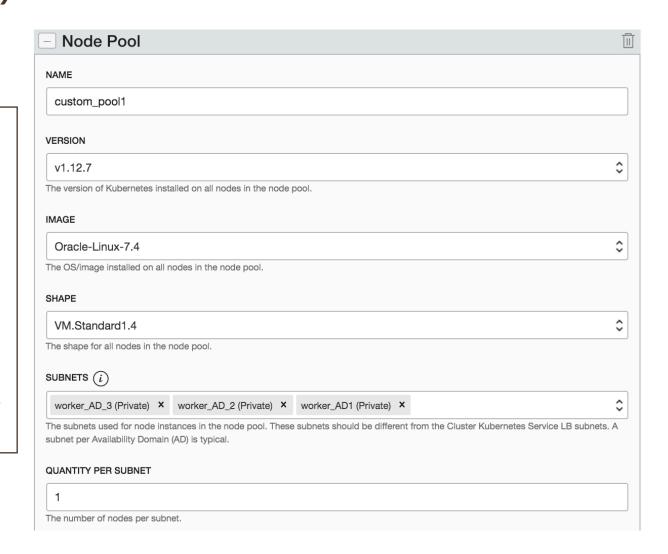
Step 2: Cluster Creation

Version: The version of Kubernetes to run on each worker node in the node pool. By default, the version of Kubernetes specified for the master nodes is selected.

Image: The image to use on each node in the node pool. An image is a template of a virtual hard drive that determines the operating system and other software for the node.

Shape: The compute shape to use for each node in the node pool.

Quantity per Subnet: The number of worker nodes to create for the node pool in each private subnet.

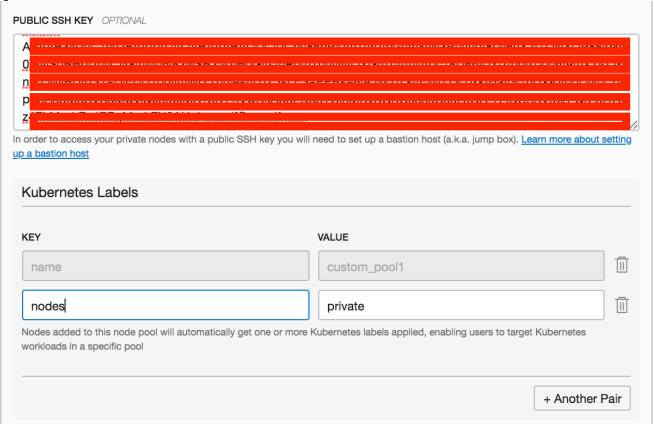




Step 2: Cluster Creation

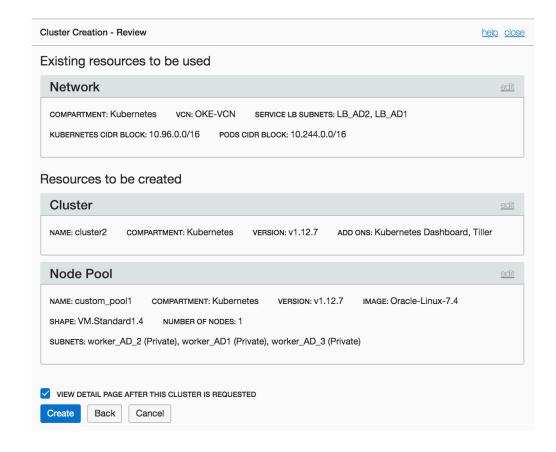
Public SSH Key: (Optional) The public key is installed on all worker nodes in the cluster, and you can use this key to access the worker nodes (Connect via Bastion Host since worker nodes are in Private subnets)

Kubernetes Labels: One or more labels (in addition to a default label) to add to worker nodes in the node pool to enable the targeting of workloads at specific node pools.





Step 2: Cluster Creation





Connecting to Private Worker Nodes via Bastion Host

How to connect to private worker nodes via a bastion host

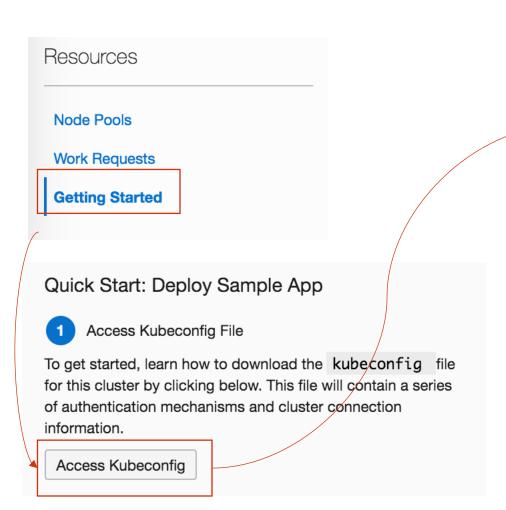
- Create a regional public subnet in OKE virtual cloud network
- Create a Bastion Host in the regional subnet (Oracle Linux)
- Verify bastion host subnet has security list rules to connect to worker nodes subnet
- Use ssh proxy command to connect to the worker nodes
- \$ ssh -t -o ProxyCommand='ssh -i /pathtosshprivatekey/ opc@<Bastion Host public IP> -W %h:%p %r' -i /pathtosshprivatekey/ opc@<private instance IP>
 - This command allows us to "tunnel" through the bastion host to our private instance, while maintaining SSH keys locally on our laptop. Storing private SSH keys on a public server such as a Bastion host is not recommended.

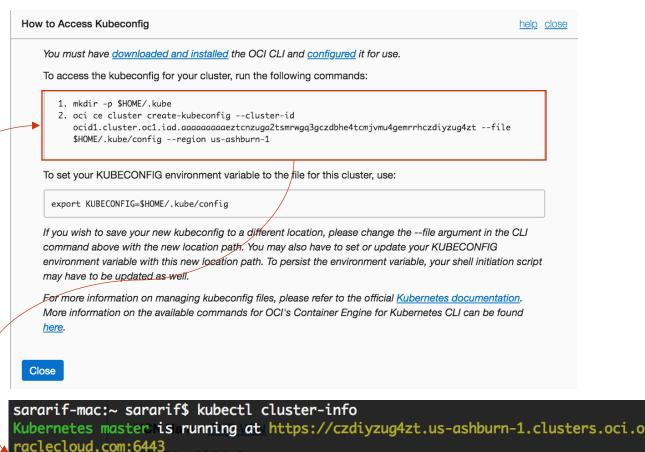


Running a basic app on OKE & exposing it externally using an OCI Load Balancer

Connecting to Cluster using Kubectl

Step 1: Navigate to Menu → Developer Services → Container Clusters (OKE) → 'Your Cluster'





KubeDNS is running at https://czdiyzug4zt.us-ashburn-1.clusters.oci.oraclecloud

.com:6443/api/v1/namespaces/kube-system/services/kube-dns:dns/proxy



Running a basic nginx on OKE

 Create a yaml file 'nginx.yaml' with deployment details of a basic nginx app

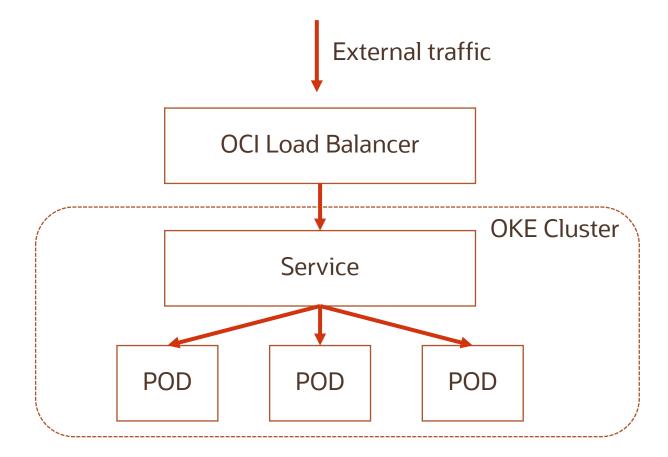
- \$kubectl apply -f nginx.yaml
- It creates a deployment with 2 replica pods each running an nginx container with the image specified in the yaml file

```
[sararif-mac:wordpress sararif$ kubectl get deployments
                   DESIRED
                             CURRENT
                                      UP-TO-DATE
                                                   AVAILABLE
                                                               AGE
                             2
nginx-deployment 2
                                                               110s
[sararif-mac:wordpress sararif$ kubectl get pods
                                                     RESTARTS
                                    READY
                                           STATUS
                                                                AGE
nginx-deployment-5c689d88bb-wl18b
                                   1/1
                                            Running
                                                                114s
nginx-deployment-5c689d88bb-xxcm2
                                    1/1
                                            Running
                                                                114s
sararif-mac:wordpress sararif$
```



Using OCI Load Balancers to Load Balance traffic

- The nginx app created in last step is running on OKE cluster, however we are unable to access the app from outside the OKE cluster.
- Kubernetes 'service' can be used to expose the app. While exposing the service, we can use the type 'LoadBalancer' to make use of OCI Load Balancer to load balance the traffic across the two pods.



Using OCI Load Balancers to Load Balance traffic

- Create a service and expose it via type LoadBalancer
 - \$kubectl expose deployment nginx-deployment --type=LoadBalancer --port=80
- A service is created and an OCI Load Balancer is launched. External IP is OCI LB listener Public IP.
- Security lists are automatically edited to allow traffic connectivity from LB subnet to worker node subnet.

```
[sararif-mac:wordpress sararif$ kubectl expose deployment nginx-deployment --type=LoadBal
ancer --port=80
service/nginx-deployment exposed
sararif-mac:wordpress sararif$ kubectl get services
NAME
                                   CLUSTER-IP
                                                                                 AGE
                   TYPE
                                                  EXTERNAL-IP
                                                                  PORT(S)
                                  10.96.0.1
                                                                                 3h36m
                   ClusterIP
                                                                  443/TCP
kubernetes
                                                  <none>
nginx-deployment
                   LoadBalancer
                                   10.96.241.71
                                                  129.213.74.72
                                                                  80:30680/TCP
                                                                                 92s
[sararif-mac:wordpress sararif$
```

129.213.74.72

Welcome to nginx!

If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to <u>nginx.org</u>. Commercial support is available at <u>nginx.com</u>.

Thank you for using nginx.



Security lists are added...

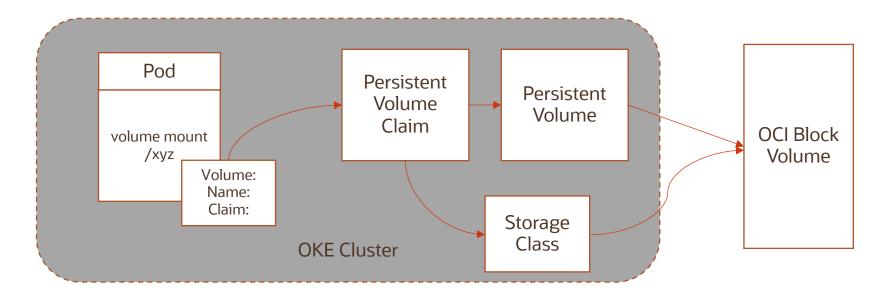
Load Balancers in Kubernetes Compartment **Create Load Balancer** Name State **OCID IP Address Overall Health** Created d2886e68-a1df-11e9-b50e-0a580aed4e8f ...4vtmyq Show Copy 129.213.74.72 (Public) ? Unknown Tue, 09 Jul 2019 00:23:49 GMT Active Showing 1 Item(s) < Page 1 > TCP TCP traffic for ports: 30680 No 172.16.20.0/24 ΑII 30680 No **TCP** TCP traffic for ports: 30680 172.16.21.0/24 30680 ΑII 172.16.10.0/24 TCP ΑII 30680 TCP traffic for ports: 30680 No No 172.16.11.0/24 TCP ΑII 30680 TCP traffic for ports: 30680 TCP No 172.16.12.0/24 ΑII 30680 TCP traffic for ports: 30680



Using OCI Block Volume as Persistent Volume Claim for Kubernetes Pods

Why Persistent storage for containers?...

- Container storage via a container's root file system is ephemeral, and can disappear upon container deletion and creation.
- To provide a durable location to store data and prevent it from being lost, you can create and use persistent volumes to store data outside of containers
- You can define and apply a persistent volume claim to your cluster, which in turn creates a persistent volume that's bound to the claim.
- A claim is a block storage volume in the underlying laaS provider (OCI BV in current scenario)



Creating a PVC...

Following yaml defines a PVC that requests 50GB of persistent storage.

```
apiVersion: v1
      kind: PersistentVolumeClaim
      name: mysqlclaim
       storageClassName: "oci"
       selector:
         matchLabels:
            failure-domain.beta.kubernetes.io/zone: "US-ASHBURN-AD-1"
10
        accessModes:
11
        - ReadWriteOnce
12
       resources:
13
         requests:
14
           storage: 50Gi
```

• \$kubectl create -f https://raw.githubusercontent.com/jamalarif/oci/master/oke/wordpress/mysql pvc claim.yml



Using PVC with K8s deployments...

- Use that created PVC while creating a new deployment. For instance -
- \$ kubectl create -f https://raw.githubusercontent.com/jamalarif/oci/master/oke/wordpress/my sql_deployment.yml

```
sararif-mac:wordpressexample sararif$ kubectl create -f https://raw.githubusercontent.com/jam]
alarif/oci/master/oke/wordpress/mysql_deployment.yml
deployment.extensions/mysql created
sararif-mac:wordpressexample sararif$ kubectl get deployments

NAME DESIRED CURRENT UP-TO-DATE AVAILABLE AGE
mysql 1 1 112s
```

```
oke-czdiyzug4zt-nrwmyrrgqzd-st6r6otoeua-0
                               Start Stop Reboot Move Resource Apply Tag(s) Actions >
                                  Instance Information
                                   Availability Domain: fyhg:US-ASHBURN-AD-1
                                                                                                                                  OCID: ...hc6oaq Show Copy
                                  Fault Domain: FAULT-DOMAIN-1
                                  Region: iad
                                                                                                                                  Launched: Mon. 08 Jul 2019 20:58:38 LITO
                                  Shape: VM.Standard1.4 @
                                                                                                                                  Compartment: iamalarif (root)/Kubernetes
                                  Virtual Cloud Network: OKE-VCN
                                                                                                                                  Launch Mode: NATIVE
                                  Maintenance Reboot:
                                  Primary VNIC Information
                                  Private IP Address: 172.16.10.2
                                  Network Security Groups: None Edit
                                  This instance's traffic is controlled by its firewall rules in addition to the associated Subnet's security lists and the VNIC's network security group
                                  NIC Attachment Type: VFIO
                                                                                                                                  Firmware: UEFI 64
                                                                                                                                  Boot Volume Type: PARAVIRTUALIZED
                                  Remote Data Volume: PARAVIRTUALIZED
                                Attached Block Volumes
Attached Block Volumes (1
                                               OCID: ...5cir7a Show Copy
                                                                                      Attachment Access: Read/Write
                                                                                                                                                                               Availability Domain: fvhg:US-ASHBURN-AD-1
Attached VNICs (1)
                                                Attachment OCID: ...rtan3q Show Copy Block Volume Compartment: Kubernetes
```

```
n: extensions/v1beta1
cind: Deployment
 name: mysql
labels:
  app: mysql
   app: mysql
     app: mysql
      - image: mysql:5.6
        name: mysql
          - name: MYSQL_ROOT_PASSWORD
                name: mysql
                 key: password
               ntainerPort: 3306
        name: mysql
          - name: mysql-persistent-storage
                   Path: /var/lib/mysql
            ne: mysql-persistent-storage
         claimName: mysqlclaim
```

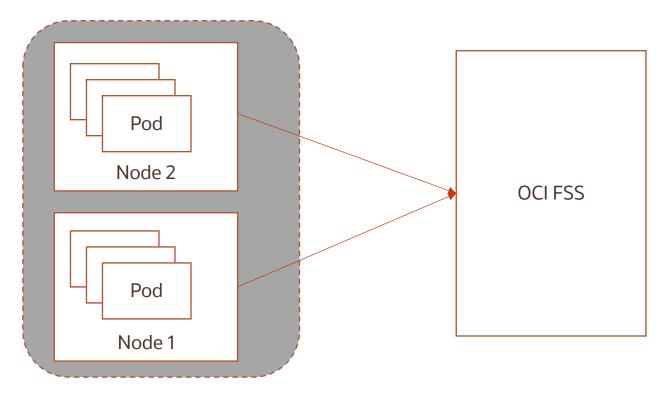




Using a network file system as persistent storage

Why FSS as persistent storage?

- Container storage via a container's root file system is ephemeral, and can disappear upon container deletion and creation.
- Containers that need to maintain state can use OCI File System Storage
- OCI FSS can be shared across multiple worker nodes, allowing multiple PODs to have access to the same persistent storage device.



Creating a PVC using OCI FSS...

Create a storage Class

```
1 kind: StorageClass
2 apiVersion: storage.k8s.io/v1beta1
3 metadata:
4    name: fsstest1
5    provisioner: oracle.com/oci-fss
6    parameters:
7    # Insert mount target OCID from the FSS here
8    mntTargetId: ocid1.mounttarget.oc1.iad.aaaaaa4n
```

Create a Persistent Volume and a persistent volume claim from this storage class.

```
apiVersion: v1
kind: PersistentVolume
netadata:
 name: oke-fsspv
 storageClassName: fsstest1
 storage: 100Gi
 accessModes:
 - ReadWriteMany
 mountOptions:
 - nosuid
 nfs:
# Replace this with the IP of your FSS file system in OCI
 server: 172.16.5.4
# Replace this with the Path of your FSS file system in OCI
 path: "/okefss-1"
 readOnly: false
```

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
   name: oke-fsspvc
spec:
   storageClassName: fsstest1
   accessModes:
   - ReadWriteMany
   resources:
    requests:
   # Although storage is provided here it is not used for FSS file systems
    storage: 100Gi
   volumeName: oke-fsspv
```



Use the volume with a Pod...

- Create a Pod and mount to PVC.
- Use node labels for launch multiple Pods across different Nodes and test persistent storage across pods.
 - \$kubectl label node 172.16.10.2 nodeName=worker1
 - \$kubectl label node 172.16.11.2 nodeName=worker2
 - Create Pods with yaml as below (Note: Volume details and nodeSelector)

```
piVersion: v1
kind: Pod
metadata:
 name: nginx-fss
spec:
 containers:
 - name: web
   image: nginx
   volumeMounts:
    - name: nfs
      mountPath: "/usr/share/nginx/html/"
   ports:
     - containerPort: 80
       name: http
 volumes:
 - name: nfs
   persistentVolumeClaim:
     claimName: oke-fsspvc
     readOnly: false
 nodeSelector:
   nodeName: worker1
```

```
oiVersion: v1
cind: Pod
etadata:
name: nginx-fss2
pec:
containers:
 - name: web
  image: nginx
   volumeMounts:
    - name: nfs
      mountPath: "/usr/share/nginx/html/"
   ports:
     - containerPort: 80
       name: http
volumes:
 - name: nfs
   persistentVolumeClaim:
     claimName: oke-fsspvc
     readOnly: false
   nodeName: worker2
```



Use the volume with a Pod...

- Use the following command to write some data in the shared persistent storage from both PODs and test if it is persistent and available across both pods.
 - \$kubectl exec -it nginx-fss bash

```
[sararif-mac:fss1 sararif$ kubectl exec -it nginx-fss bash
[root@nginx-fss:/# cat /usr/share/nginx/html/hello.txt
HI THERE, I AM POD 1 on WORKER Node 1
Good to hear from POD1, this is POD2 from workernode 2
[root@nginx-fss:/# exit
exit
[sararif-mac:fss1 sararif$ kubectl exec -it nginx-fss2 bash
[root@nginx-fss2:/# cat /usr/share/nginx/html/hello.txt]
HI THERE, I AM POD 1 on WORKER Node 1
Good to hear from POD1, this is POD2 from workernode 2
[root@nginx-fss2:/# exit
exit
sararif-mac:fss1 sararif$
```

Summary

After completing this lesson, you should be able to understand:

- How to create an OKE cluster using Custom Create
- How to connect to worker nodes of a private cluster via a Bastion Host
- How to expose an app using k8s service and use OCI LB to load balance the traffic
- Why Persistent Storage is required for containers and how you can
 - Using OCI Block Volume as persistent storage
 - Using OCI File System Storage as persistent storage

ORACLE

Oracle Cloud always free tier:

oracle.com/cloud/free/

OCI training and certification:

cloud.oracle.com/en_US/iaas/training/certification
education.oracle.com/oracle-certification-path/pFamily-647

OCI hands-on labs:

ocitraining.qloudable.com/provider/oracle

Oracle learning library videos on YouTube:

youtube.com/user/OracleLearning

