ORACLE

Oracle ZFS Storage in OCI Quick Start Guide

Configuration of an Oracle ZFS Storage Instance in Oracle Cloud Infrastructure (OCI)

December 2021 | Version 1.9

PURPOSE STATEMENT

This document provides step-by-step instructions for configuring an Oracle ZFS Storage instance in OCI.

DISCLAIMER

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document and information contained herein may not be disclosed, copied, reproduced or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. 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, and timing of any features or functionality described in this document remains at the sole discretion of Oracle.

Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

TABLE OF CONTENTS

INTRODUCTION

Oracle is uniquely positioned to provide products and services that run 24/7 either on-premises or in the cloud and so, has the expertise to optimally run our own products in Oracle's own cloud.

Oracle ZFS Storage in OCI Marketplace provides cloud-based NAS storage and replication services to enable on-premises ZFS Storage customers to migrate data and apps from on-premises to OCI. Oracle ZFS Storage instances provide both protocol services and performance for data migration, replication, and sharing.

The Oracle ZFS Storage image in OCI can be configured as a Bare Metal (BM) or Virtual Machine (VM) instance to support the following use cases:

- Migrate data to OCI over NFS, NFSv4, SMB or cross protocols with AD integration using an Oracle ZFS Storage BM or VM instance as a storage gateway
- Share data from ZFS BM or VM instance in OCI over NFS, SMB, or cross protocols back to on-premise
- Replicate data to ZFS BM or VM instance in OCI as a replication target and also reverse the replication back to on-premise
- Migrate and host applications workloads using similar protocols as your on-premise deployments

Consider that you can evaluate the ZFS Storage product and its unique features as a cloud-based storage solution that could potentially provide value or solve infrastructure problems at a very low cost.

Sharing data and replicating data can be hosted in the following ways:

- Cloud to Cloud
- On-premise to Cloud
- Cloud to on-premise

After you have reviewed the above supported shapes, review the following summary of recommended shapes and recommended number of NFS and SMB clients.

Network Bandwidth Expectations for NFS/SMB Clients

Shape	Memory	Network Speeds	Maximum Client Bandwidth	Typical Sustained Bandwidth	Number of Clients
VM.Standard2.4	60GB	4.1 Gbps	256 MB/s	192 MB/s	Tens
VM.Standard2.8	120GB	8.2 Gbps	512 MB/s	384 MB/s	Hundred
VM.Standard2.16	240GB	16.4 Gbps	1025 MB/s	768 MB/s	Few Hundred
VM.Standard2.24	320GB	24.6 Gbps	1537 MB/s	1150 MB/s	Hundreds
BM.Standard2.52	768GB	25x2 Gbps	3125 MB/s	2343 MB/s	Thousands

Notes:

- Typical sustained workload mix with 50% read / 50% write.
- Number of clients depends on the desired throughput available to each client. If more throughput is needed per client then fewer clients should be used.
- A bare metal (BM) or virtual machine (VM) instance requires only one volume for operation. You can add more volumes to increase storage capacity for your needs.
- Maximum block volume capacity is 960TB.
- Detailed shape specifications are available at OCI Shapes.

What's New

This release includes the following new features:

- Clustered ZFS in OCI Marketplace instances
    - Provides a highly available operating environment providing file and storage services in the event of a node failure.
    - An active/passive configuration of two instances is supported.

- The passive node can detect that the active node is unavailable and take over as the active node. The data pools from the active node are exported and imported on the passive node so that it becomes the active node.

Known Issues

- Bare Metal shapes will generate a spurious network problem that can be ignored. The problem will be seen under the ZFS Storage in OCI Maintenance -> Problems tab and description will read: *The driver is suffering from a performance error detected in the driver. A(n) unsupported error has been detected during driver's attach context causing a(n) performance service*. (30773285 - OCI ZFS on BM network interfaces show errors on 2.52 shapes)
- Virtual Machine instances will show network devices speed as 1Gb even though it will use the full bandwidth allowed by the compute shape. (32749253 - VNICs speed is mentioned as 1G at CLI/BUI though VNIC effective bandwidth is more)
- If a new OCI VNIC is added to a running ZFS Storage in OCI VM, a reboot is required before the network device can be used. (32518670 - Adding an additional vnic to the OCI zfssa VM fails)
- The primary network interface used for iSCSI boot should not be modified. Use secondary network interfaces instead. (33001957 - Adding a secondary IP address to bnxt0 panics the zfssa BM instance)
- Bare Metal shapes will generate an *unknown* interface on the original bnxt0. This *unknown* entry does not impact the system other than having noise in the datalink/interface list. (33693103 - unknown1 interface post BM vio cluster setup).
    - The workaround to remove the *unknown* entry is as follows:

      ```
      Instance-A:> svc.deleteComposite('interface:bnxt0')
      Instance-A:> svc.deleteComposite('datalink:bnxt0')
      Instance-A:> maintenance system restart
      ```
    - Note that restart only restarts the management software.
    - Be aware system might reboot after running this command.

Image Configuration Summary

- Two images are available:
    - One is for a virtual machine (VM) instance, which currently is: ORACLE-ZFS-STORAGE-VM-8.8.38-1.2.38.4957.1X
    - One is for a bare metal (BM) instance, which currently is: Oracle-ZFS-Storage-BM-8.8.38-1.2.38.4957.1x
- A boot volume and data volume are the minimum requirements for configuring a ZFS Storage instance

Overview of Configuration Steps

This guide describes the steps to configure Oracle ZFS Storage as a compute instance in Oracle's Cloud Infrastructure (OCI) and contains the following sections:

1. Import OCI Image
2. Configure OCI Compute Instance
3. Configure Block Storage
4. Configure ZFS Storage
5. Share an SMB Filesystem

For more information, see the following references:

- Oracle ZFS Storage Appliance - Release OS8.8.x - General ZFS Storage administration information
- APIs for ZFS Storage in OCI - Final section provides additional management APIs used developed specifically for Oracle ZFS Storage in OCI version.

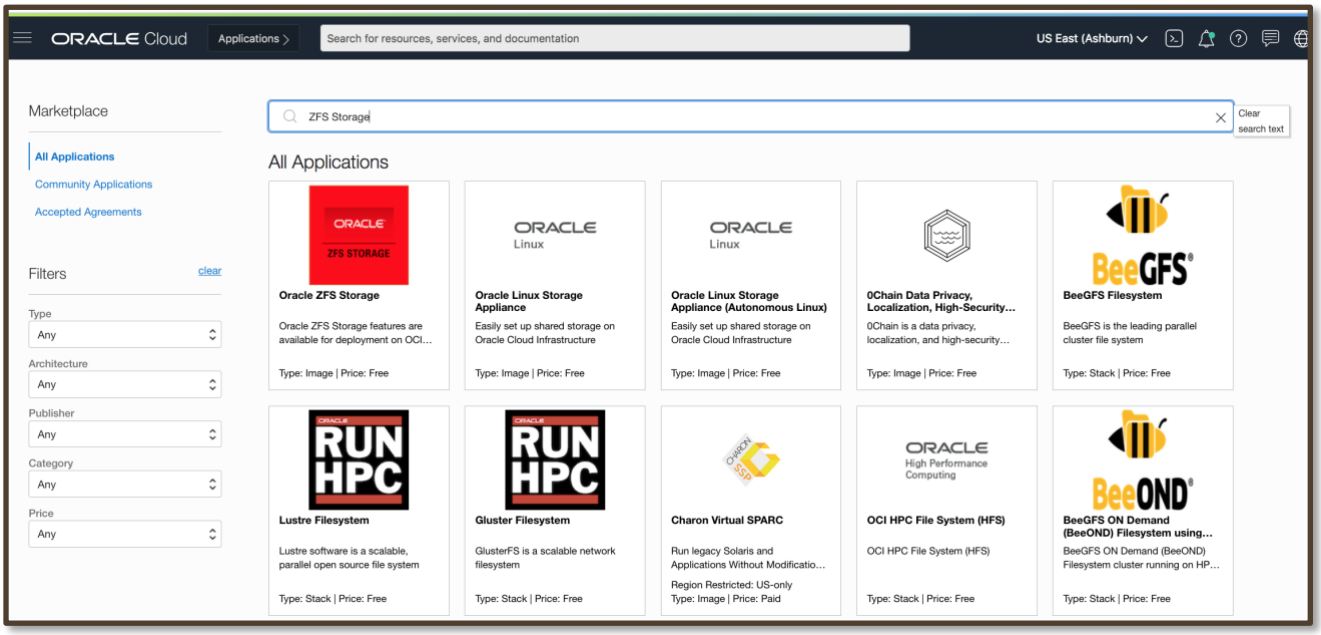
FIRST STEPS

1. The first step is to get an Oracle Cloud Infrastructure account.
   https://www.oracle.com/cloud/
   This guide assumes a usable compartment, virtual cloud network (VCN) and subnet has already been created and setup for use. An administrator for your OCI tenancy will authorize resources in a specified compartment for you to use.
   The following information will be needed to configure the OCI compute instance.
   1. OCI Compartment ID
   2. VCN Compartment and Name
   3. Subnet Compartment and Name

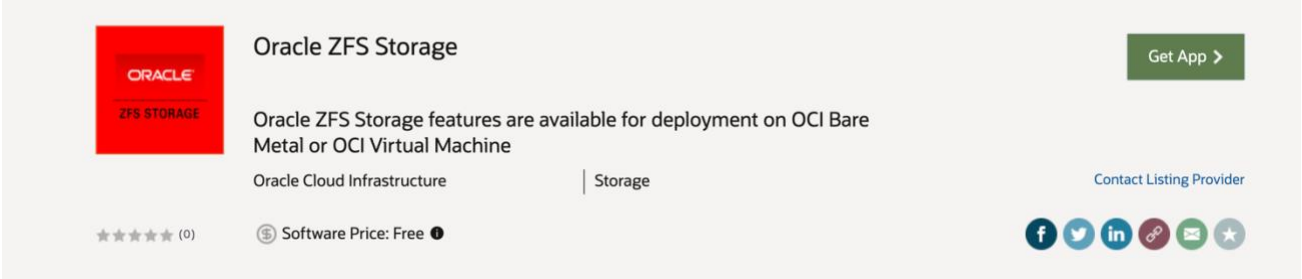
   You will also need a `ssh` client to do the initial configuration and know how to configure the `ssh` client to use `ssh` key authentication.

1. IMPORT OCI IMAGE

1. Log into your OCI tenancy and region.
2. In the left hamburger menu, click on Marketplace.
3. In the main screen, click on Marketplace.
4. In the main search screen, enter ZFS Storage.
5. When the Oracle ZFS Storage image appears, click on the link.



6. From the Oracle ZFS Storage screen, click get app.

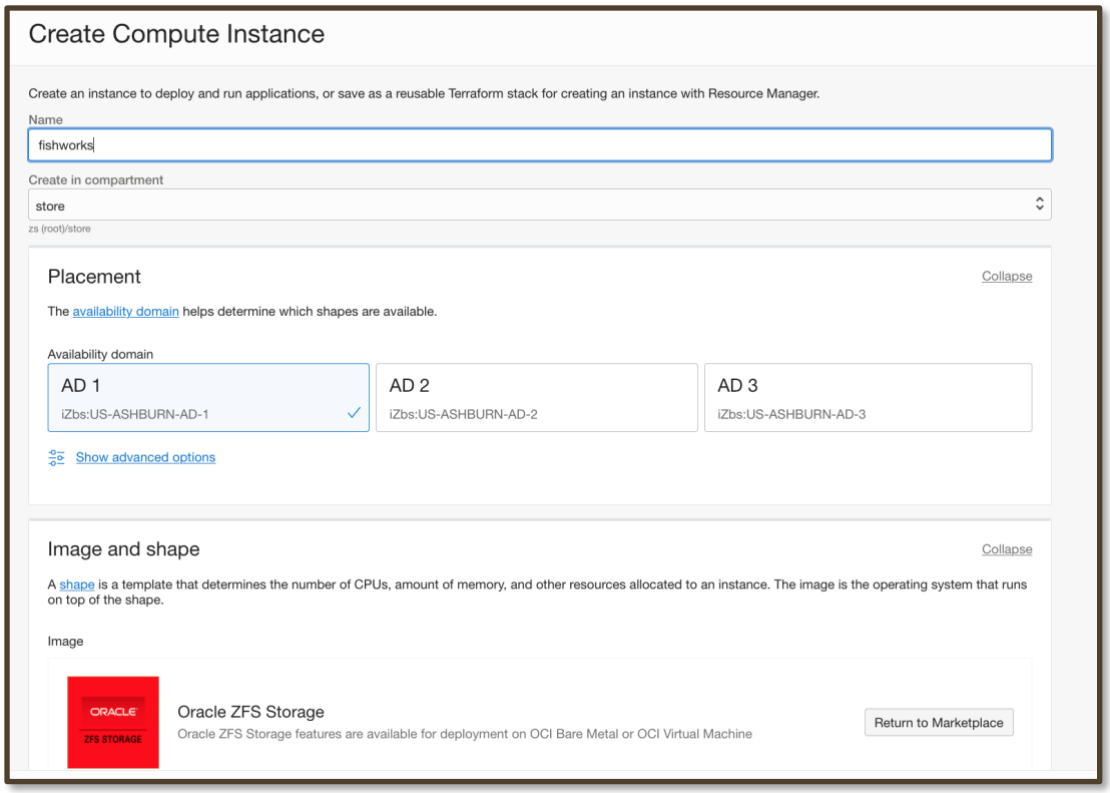


7. The virtual machine (PV image) is selected by default. Select BM image if you are configuring a bare metal instance.

   Review the Overview below, review and accept the terms and conditions. Then, click Launch Instance.

2. CONFIGURE OCI COMPUTE INSTANCE

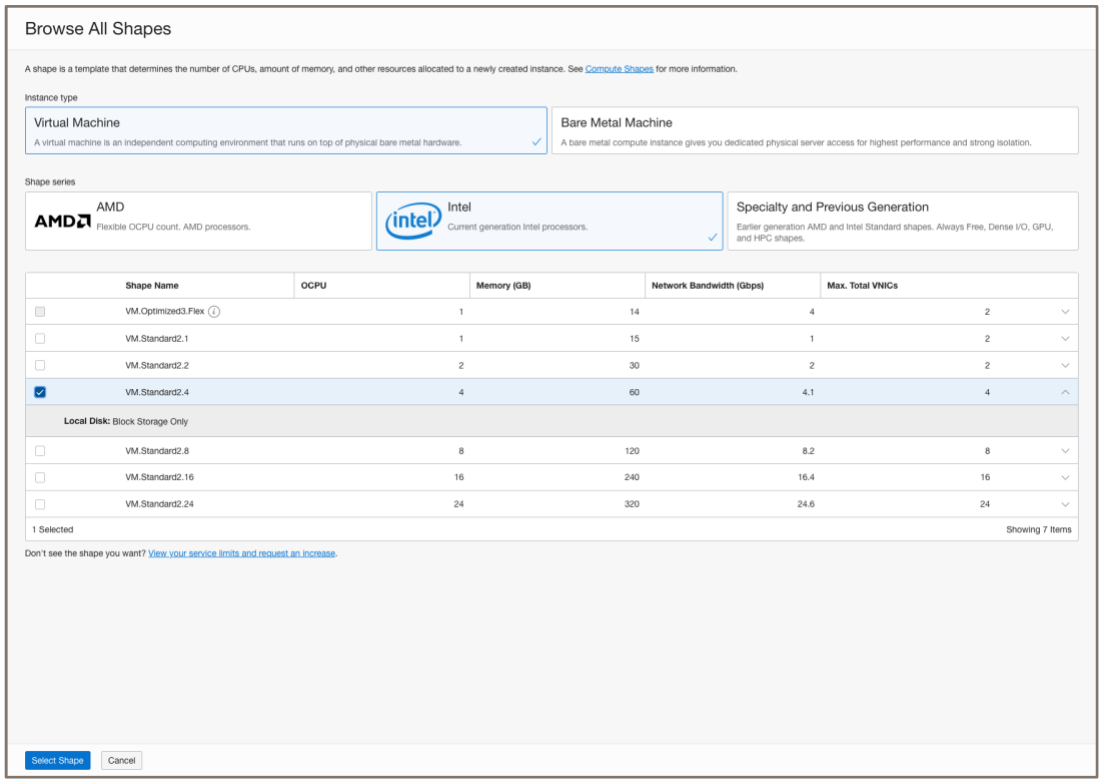
1. Enter the name of the compute instance.
2. Click the Change Shape button. For example, to change to a virtual machine (VM) shape from a bare metal (BM) shape or vice versa.



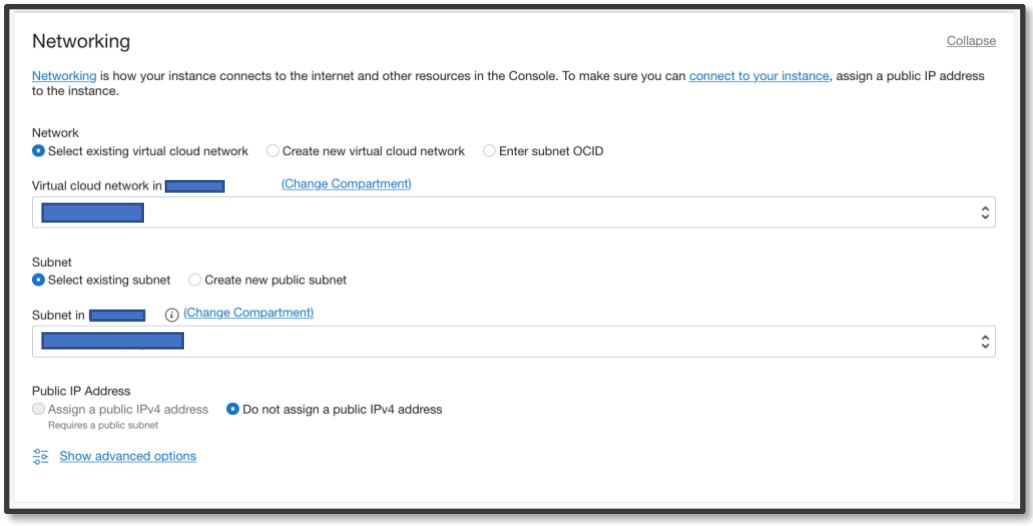
3. From the Browse All Shapes screen, select the Virtual Machine shape or the Bare Metal shape.
   - For a Bare Metal image, select BM.Standard2.52
   - For a Virtual Machine image, select VM.Standard2.4, VM.Standard2.8, VM.Standard2.16, or VM.Standard2.24

   In the example below, the VM shape is selected.

4. Next, select the Intel shapes box and then the VM.Standard2.4 shape.



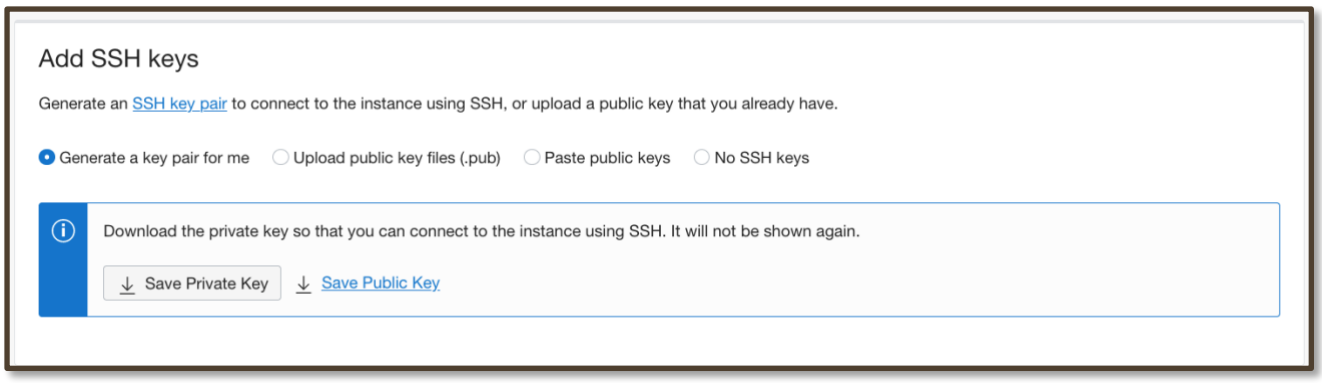
5. Click Select Shape.
6. Configure networking settings for the instance.
   Ask your OCI tenancy administrator what Network and Subnet to use.



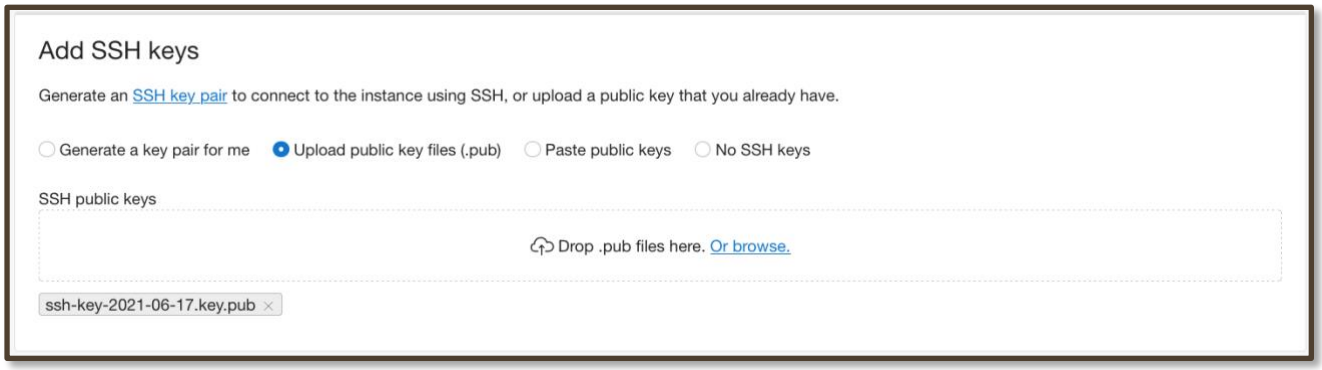
7. Generate an SSH key and save the key, choose an existing public key file, or paste the contents of a public key.
   Generate or locate your ssh keys. Existing keys can be found in the your .ssh directory.

```
$ ls .ssh
config          id_rsa          id_rsa.pub          known_hosts
```

If you already generated keys for the opc user, locate opc/opc.pub.



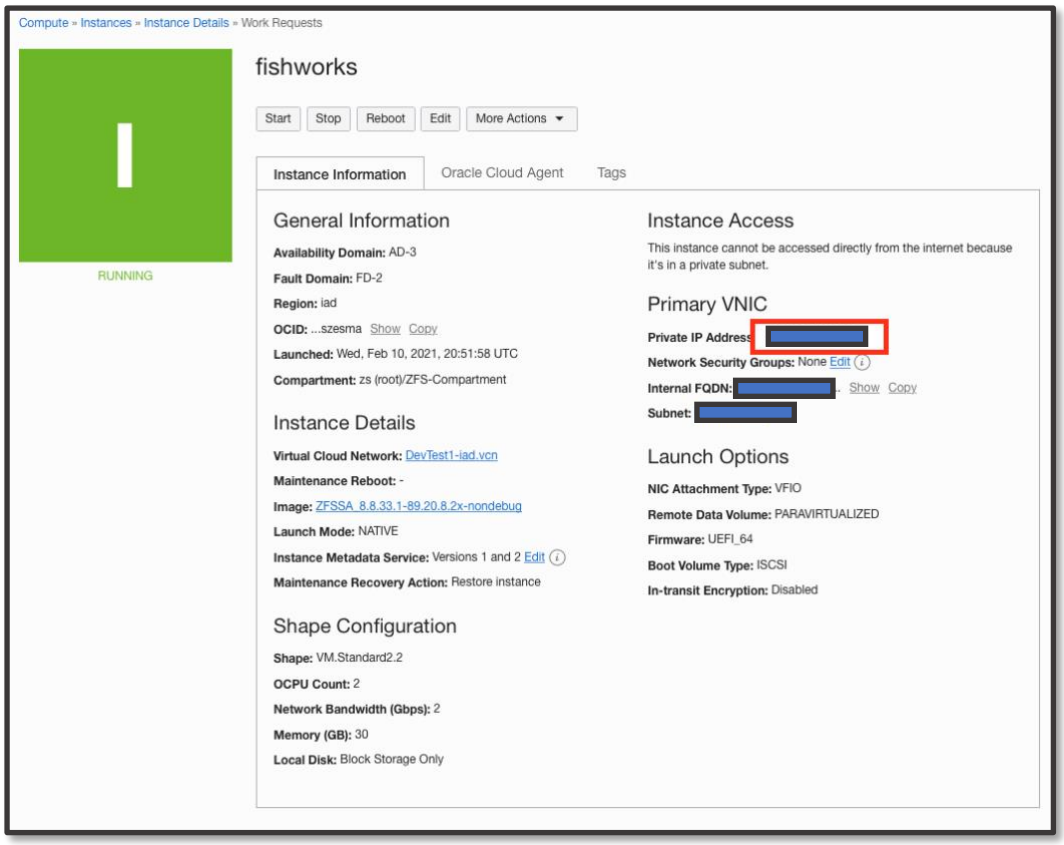
For example, click Upload public key files and select it from the browse link.



8. Boot volume configuration.
   Accept both default values. By default, a 300GB boot volume is created. For most use cases, this size should be acceptable.

9. Select *Create* to launch the compute instance.

10. Wait for the instance to go into the running state and copy the *Primary VNIC Private IP Address.*



11. Open a terminal and `ssh` to the ZFS Storage Appliance *Private IP Address* to set the opc user password to enable access to the BUI.
    - The instance includes the `opc` user by default. The opc account has all authorizations enabled and can be used to configure the storage appliance. If root user access is needed, see https://support.oracle.com/knowledge/Sun%20Microsystems/2811414_1.html.
    - You can transition to a full administrative-capability root account once you have logged in as the **opc user** if you need full administrative access to the instance.
    - Use your `ssh` credentials to log in as the `opc` user to your newly running instance.
    - There could be a slight delay before you can `ssh` into the running instance.

    For example:

```
ssh -i <path_to_private_key_file> opc@203.0.113.29

fishworks:> configuration users
fishworks:configuration users> select opc
fishworks:configuration users opc> set initial_password
Enter new initial_password: ********
Re-enter new initial_password: ********
              Initial_password = (set) (uncommitted)
fishworks:configuration users opc> commit
fishworks:configuration users> exit
```

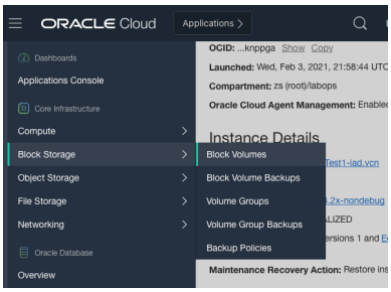
3. CONFIGURE BLOCK STORAGE

In this section, you will do the following steps:

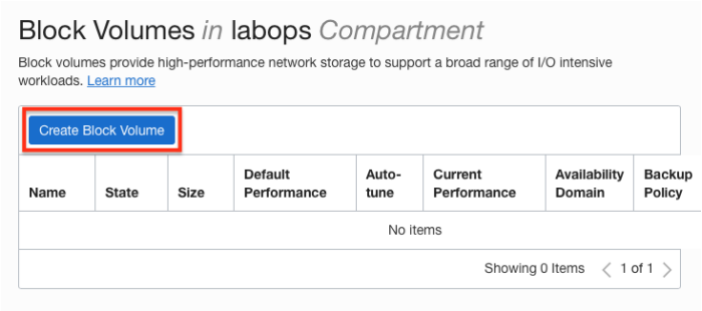
- Create block volume or volumes
- Attach block volumes to ZFS Storage instance
- Add iSCSI target

This section will create a block volume in OCI and attach it to a ZFS Storage Appliance. Once that is done a storage pool is created on the appliance.

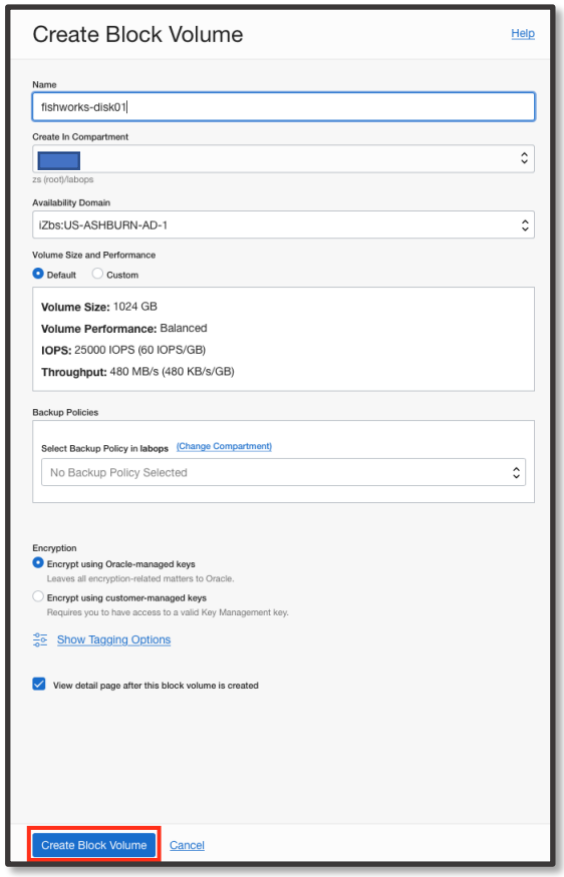
1. Go to the *Block Volumes* page for your compartment.



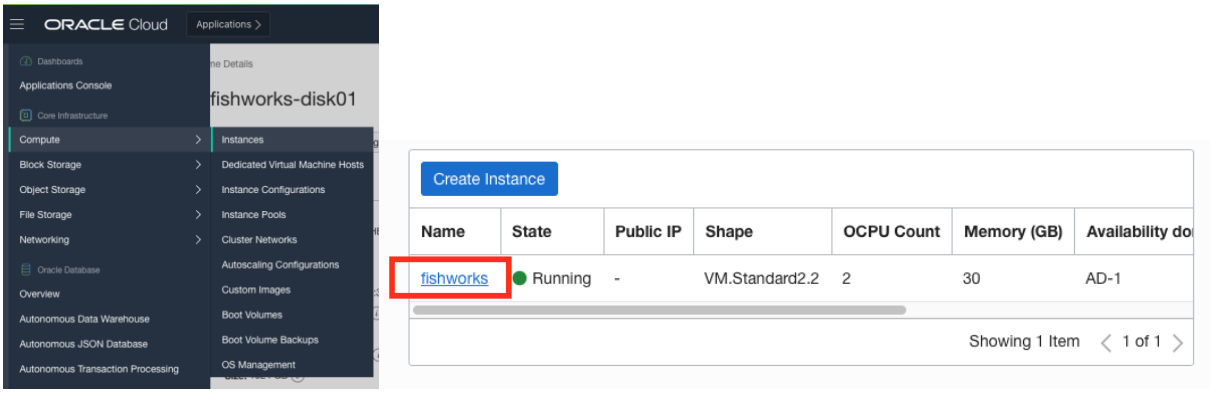
2. Select *Create Block Volume.*



3. Set create block volume properties and select *Create Block Volume.*

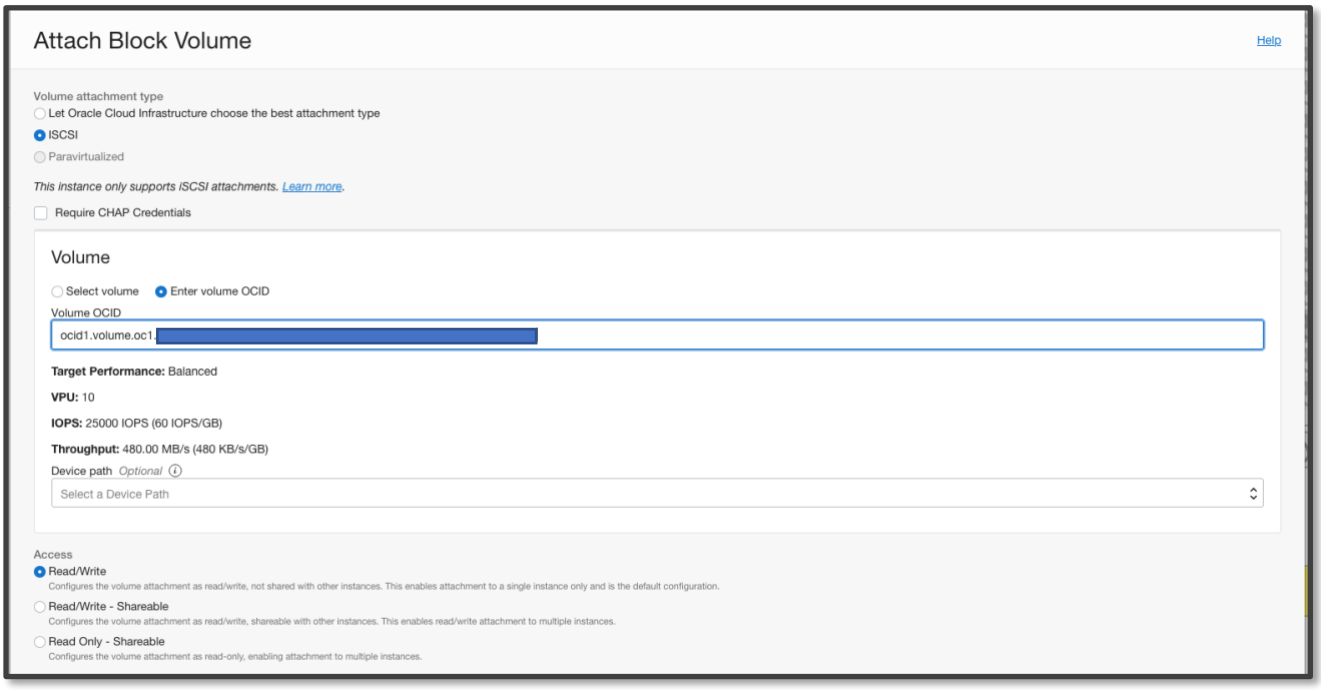


4. Attach storage to compute instance. Go back to Compute Instances and select the ZFS storage appliance compute instance.



5. Scroll down to Resources and select *Attached Block Volumes* and then select *Attach Block Volume.*

6. Set Attach Block Volume Properties.
   - For Bare Metal shapes select iSCSI, for VM shapes select iSCSI or paravirtualized.
   - Select the block volume from your compartment by volume name or volume OCID.
   - Select Read/Write
   - Click Attach

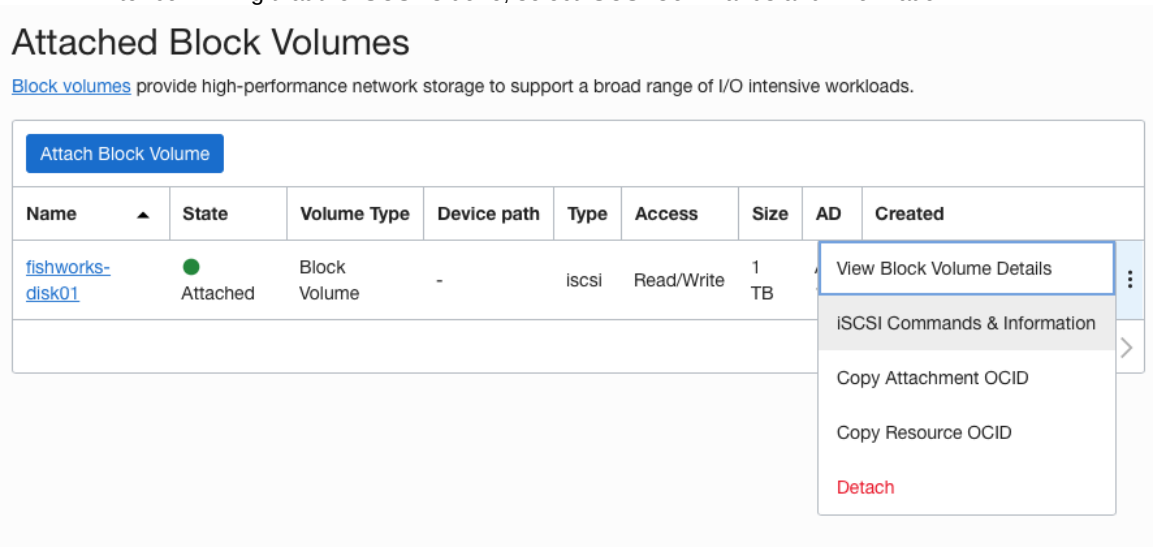


Repeat step 1 through 6 to add more block volumes for data disks.

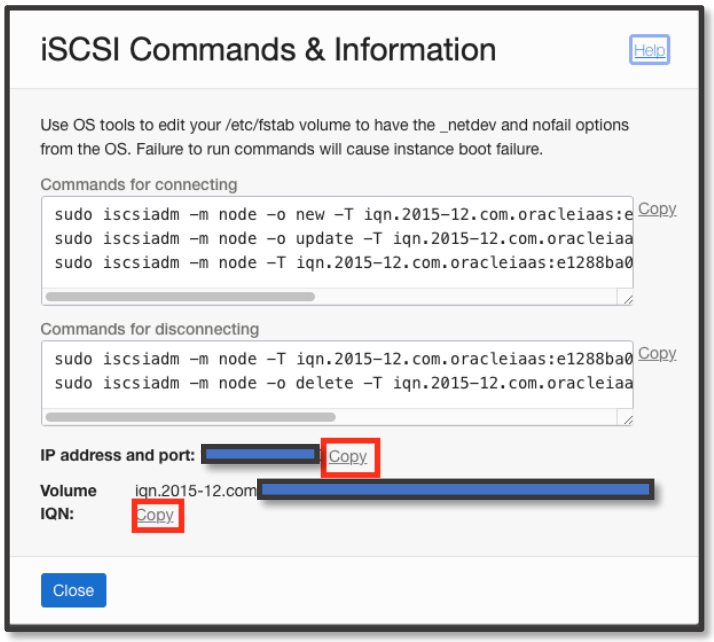
4. CONFIGURE STORAGE FOR ZFS STORAGE INSTANCE

1. Add iSCSI target to ZFS Storage in OCI.

   After confirming that the iSCSI is done, select iSCSI Commands and Information.

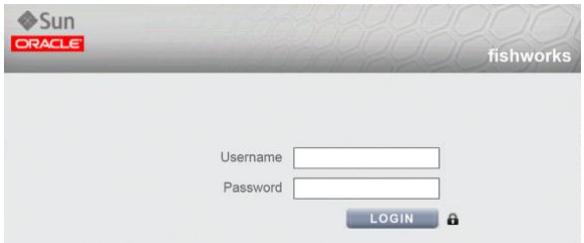


2. Get iSCSI *IP Address and port* and *IQN*.



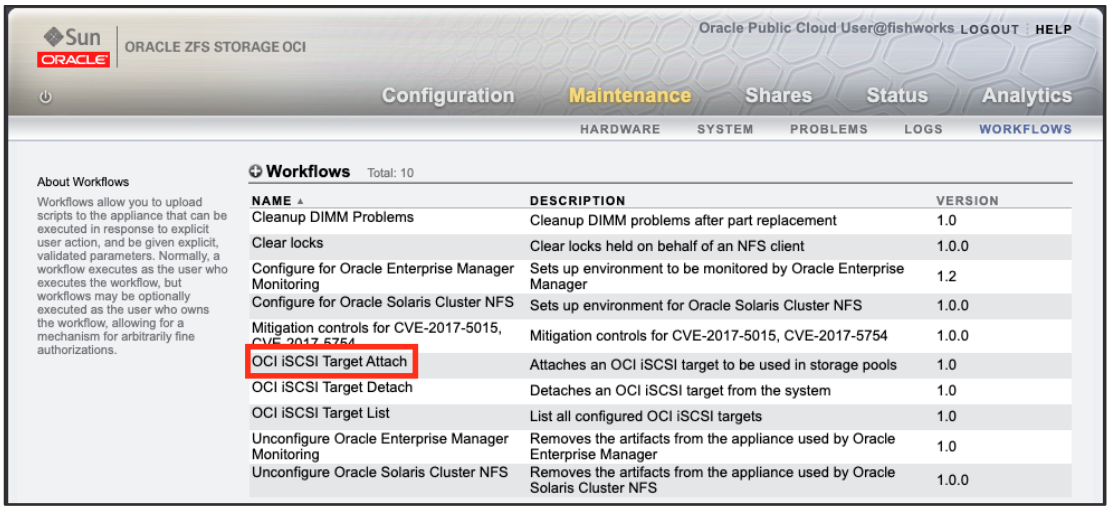
3. Open a web browser and navigate to the ZFS Storage URL. For example: https://203.0.113.29:215

   This will be the IP address of the *Primary Private Address* on port 215.
   Accept the web browser security warning. Then the login screen appears.

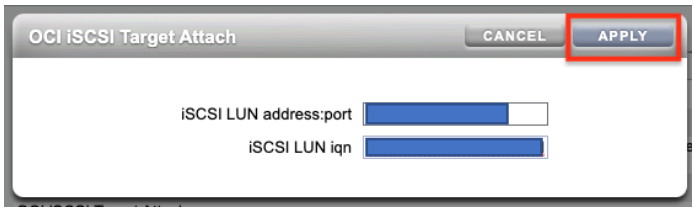


   Type in the *Username* opc and the *Password* created on the ssh terminal and click *LOGIN*.
   This guide will use the BUI for the remaining configuration.

4. Go the Oracle ZFS Storage Appliance BUI. Select *Maintenance* and then select *Workflows* and then select the arrow on *OCI iSCSI Target Attach* workflow to execute the workflow.



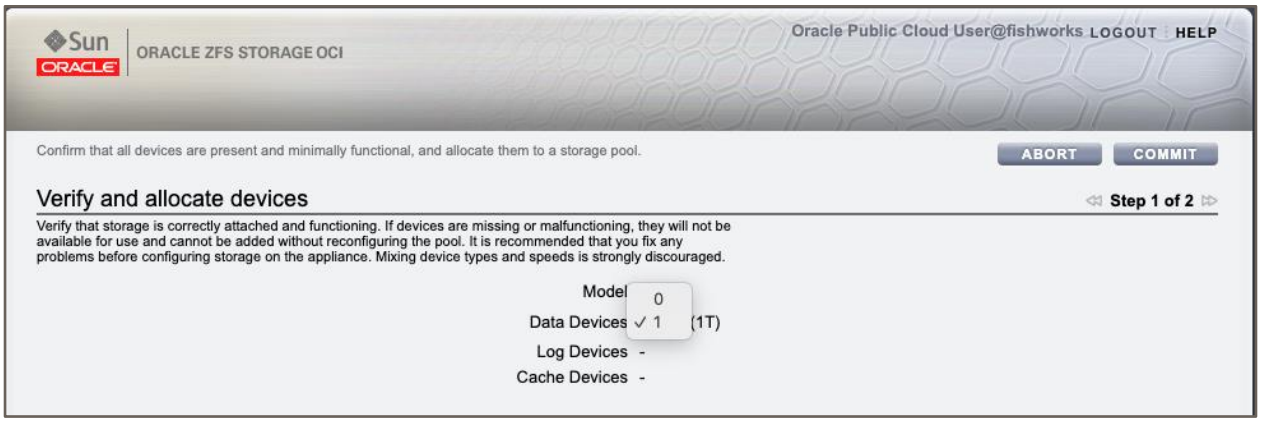
5. Enter Address and IQN from the OCI iSCSI Commands & Information into the workflow and then click apply.



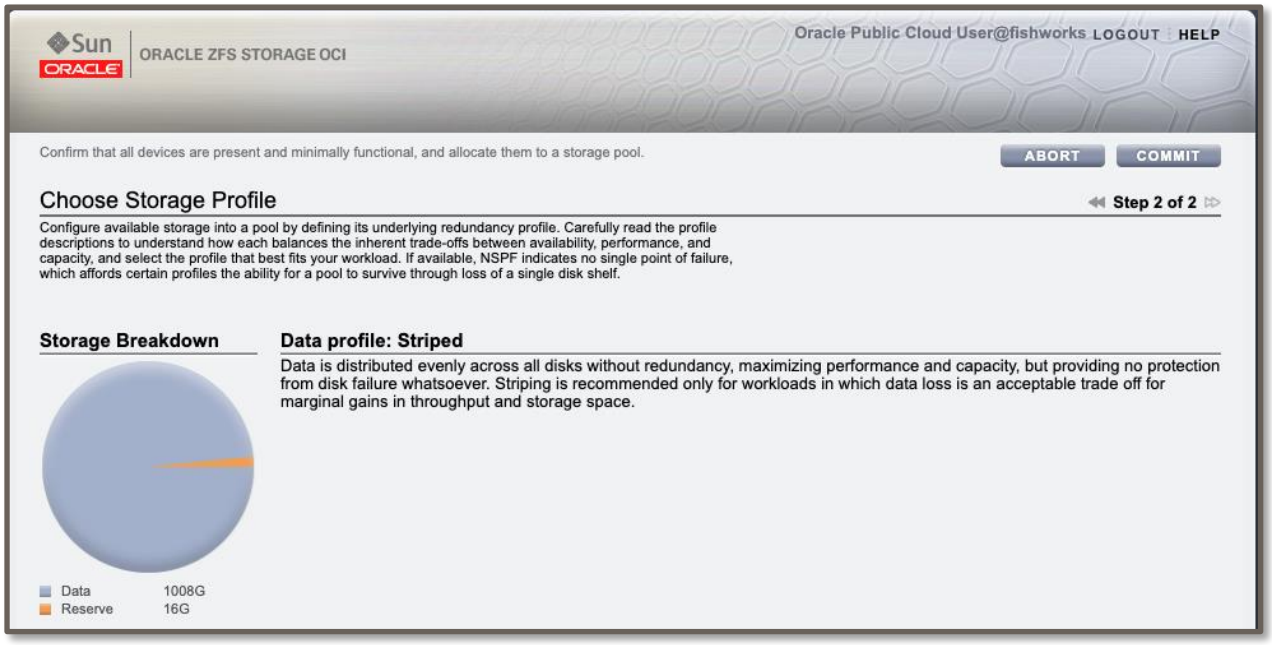
6. Go to *Configuration Storage* and create a storage pool by clicking the ⊕ icon next to *Available Pools*.

7. Name the new storage pool and click APPLY.

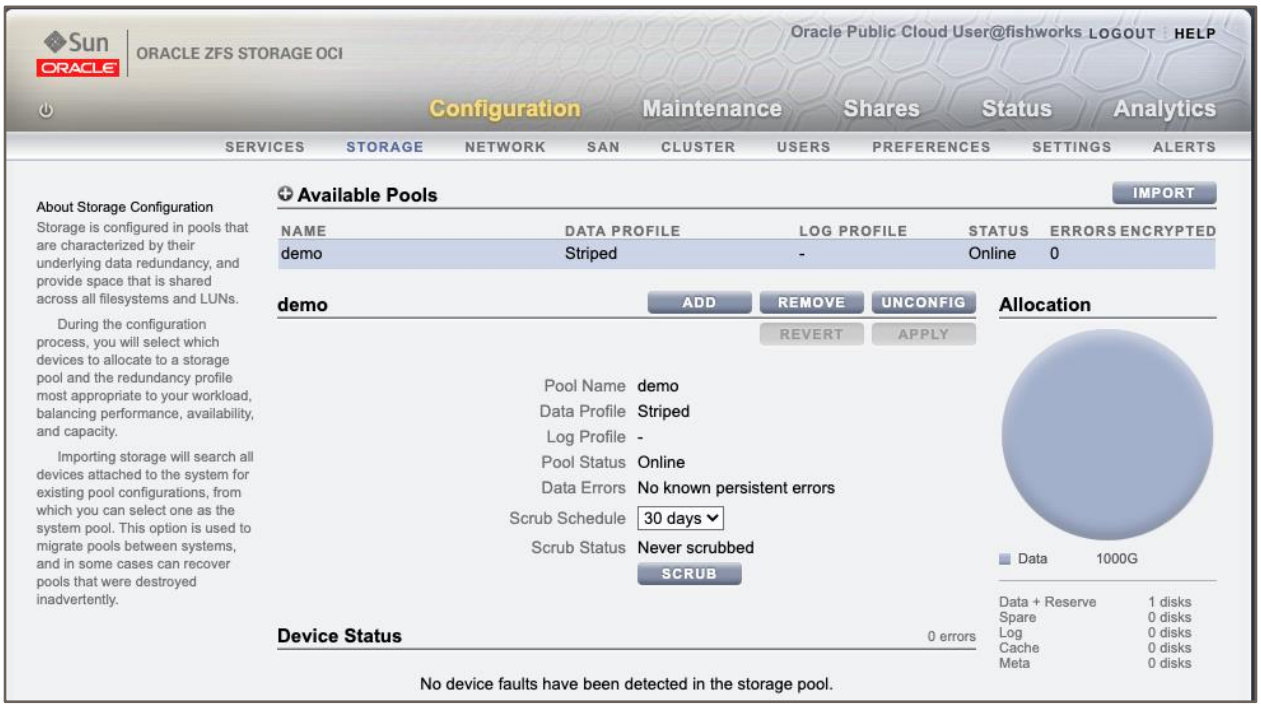
8. Use the drop-down list next to *Data Device*s to select a disk to allocate to the pool, then click *COMMIT*.



9. Select COMMIT to accept the striped data profile.



10. The *Configure Storage* page is redisplayed with a description of the new pool.



5. CLUSTER CONFIGURATION OVERVIEW

This section describes how to configure OCI instances into a clustered pair of instances that supports active/passive configuration, which means one instance is active providing data services and one instance is passive but available for operation if the active instance becomes unavailable.
Active/passive configuration behavior:

- Primary data pool or pools are configured and running on the active instance
- If the active instance fails, the primary data pool(s) are exported and imported on the passive instance
- The passive instance becomes the active instance until the active instance is recovered
- Estimated failover time between instances is 70-90 seconds

- Orchestration software transitions the following components when takeover occurs back to the active instance:
    - Secondary IP networks
    - Public IP networks
    - Storage volumes

VIO Clustering

A virtual cluster link (VIO) is used to cluster two ZFS Storage instances. Network IP addresses are used for the clustering transport. Heartbeats occur over VIO link using primary VNIC and IP. The cluster quorum is determined by OCI compute instance metadata properties.

- Primary OCI VNIC is used for the VIO link as well as storage volume I/O and OCI API calls.
- This VNIC should be setup on a private subnet with no access to NAS clients or storage administrators.
- Secondary OCI VNICs supply access to NAS clients and storage administrators.
- Configuration changes are synchronized across instances.

Clustered Instance Terminology

A resource is a physical or virtual object that is present and possibly active on one or both cluster heads. Resources are managed by storage administrators who can set which instance owns the resource when CLUSTERED.

Term	Description
Resource Type	
Singleton	Known by both instances but only active on one instance. (Storage Pools and NAS IP)
Private	Only available and active on one instance. (Administration Network Interface)
Replicate	Resource known by both heads. (Service configuration)
Symbiote	Follows other resources (Replications actions follow storage pool)
Clustered State	
Unconfigured	Clustering is not configured.
Owner	Clustering is configured. This active instance owns the storage and data resources.
Stripped	Clustering is configured. This passive instance does not control any shared resources.
Clustered	Clustering is configured in an active/active configuration. Not supported in this release.

Clustered Configuration Operation

- OCI API commands are issued from each clustered ZFS instance to manage OCI compute, storage, and network resources.
- OCI principal authentication is used to issue OCI API commands.
- All ZFS cluster resources must be in the same OCI availability domain and the same dynamic group.
- All storage volumes will be mounted as shareable on both ZFS instances.
- Network interfaces configured as singletons must use secondary IP addresses so they can be migrated.

Requirements for Clustered ZFS Instances

Review the following sections to identify requirements for running a clustered ZFS instance configuration.

ZFS Compute Instance Requirements

- Both instances must be in the same OCI Availability Domain, are the same shape type, and running the same OCI version.
- Both ZFSSA compute instances should be in different OCI Fault Domains
- Both ZFS instances must be part of a dynamic group with permissions to manage instances, network attachments, and storage attachments in the group. For more information, see the following references:
    - Create a Dynamic Group and Matching Rules
    - Write Policies for Dynamic Groups
        - Allow dynamic-group vio.dg to manage instance-family in compartment store:store.dev:vio

        - Allow dynamic-group vio.dg to manage all-resources in compartment store:store.dev:vio
        - The OCI extended metadata properties available to each compute instance will have a 'zfssa' dictionary property that shall be exclusively used by the VIO subsystem.
- The passive or peer instance (B) must be launched with no ssh keys and no user configuration script. The system will only be accessible via the system console until it has been configured as part of a cluster.

Storage

All storage volumes should be attached as sharable to both ZFS Storage instances.

Network Requirements

At least two network interfaces are required. The first network interface is used exclusively for cluster I/O, OCI backend disk access, and OCI API calls. The second interface is used for NAS clients. It is recommended to add two more interfaces that are used for administrative access to each ZFS instance.

- Primary NIC on both peers must be on the same subnet. All cluster traffic occurs over the primary NIC.
- Primary OCI VNIC should not allow access to storage administrators or NAS clients.
- Secondary OCI VNICs must be attached to both systems in order to provide appliance administration or storage services.
- Network interfaces configured as singletons must use secondary IP addresses so they can be migrated.
- Any additional network interfaces but be configured on the same subnet on both peers.
- NAS clients expecting HA must use OCI VNICs with secondary IP addresses or public IP addresses configured as cluster singleton resources.
- Each OCI VNIC pair needs to be on the same subnet.

OCI Network Configuration Example

The following table describes a typical network configuration for an active/passive cluster named zs1.

- Primary OCI VNIC is used for the VIO link, storage I/O and OCI API calls
- Secondary OCI VNICs are used for NAS clients and storage administrators
- Singletons must use secondary IP addresses so they can be migrated

VNIC	ZFS INSTANCE A (ZS1-A)	ZFS INSTANCE B (ZS1-B)	USAGE
vtionet3 primary	Unused	zs1-b	Optional Private B admin access (A unused)
vtionet2 primary	zs1-a	Unused	Optional Private A admin access (B unused)
vtionet1 secondary	zs1-pool1 (if owner)	zs1-pool1 (if owner)	NAS client access to pool1 used by owner
vtionet1 primary	Unused	Unused	Do not use because NAS clients cannot use passive system

ZFS instances A and B in the zs1 cluster will have datalinks configured for vtionet1, vtionet2, and vtionet3 as described in the following table.

DATALINK	INTERFACE	CLUSTER TYPE	ACCESS
vtionet2	Only configure zs1-a	Private	instance-A
vtionet3	Only configure zs1-b	Private	instance-B
vtionet1	Only configure zs1-pool1	Singleton	OWNER

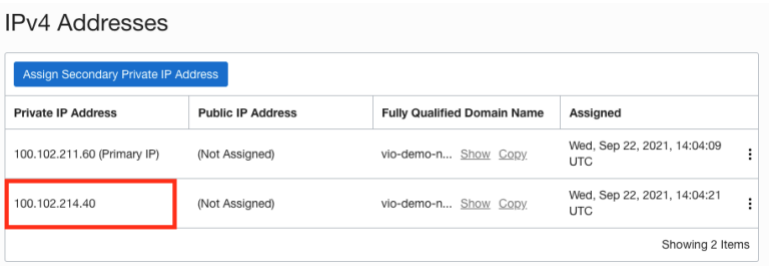
How to Configure Clustered Instances Manually

After your primary ZFS in OCI Marketplace instance is configured as described in the previous sections steps a-x, deploy a second ZFS in OCI Marketplace instance (with only a system pool configured).
In the following example, the primary (active) instance is instance-A and the secondary (passive) instance is instance-B. The instances are *peers* and some steps below will prompt you for the peer IP or ocid, for example.
Prerequisite steps:

- Deploy two instances:
  - Primary (active) instance-A configured with block volume storage configured as data pool(s)
  - Secondary (passive) instance configured with system pool only
  - No current requirement that instances are the same shape
  - Both instances must be in the same availability domain
- Identify secondary instance-b information
  - OCI instance ID
  - IP address
- Open two terminal windows, one for each instance.

1. Launch instance-A normally with ssh key, user configuration script, and attach at least one secondary VNIC and storage.
2. Configure instance-A network.
   - Secondary VNIC (vtionet1) should use the secondary IP address, not the primary address.
   - The secondary IP address will be migrated to instance-B during takeover.

   In this example, the IP address 100.102.214.40 must be used. If the Primary IP is used, then clients will not have access to the cluster when instance-A is passive.



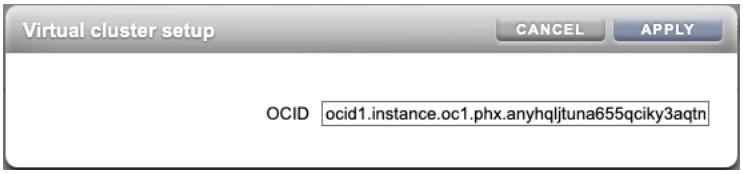
   After the network is configured, verify access to the secondary IP (100.102.214.40)

   Note: The instance may have to be rebooted to have access to vtionet1, vtionet2, and vtionet3, if they were added after the instance started.

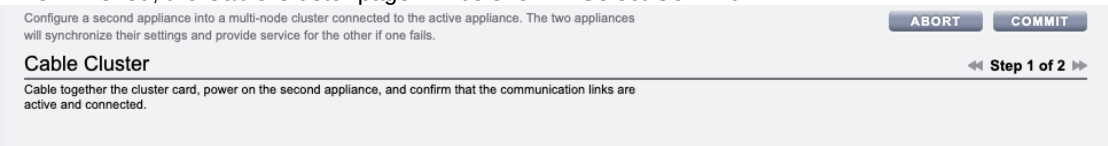
3. Launch instance-B with no ssh keys and no user configuration script.
   - Primary subnet must be the same as instance-A.
   - Instance-B must have vtionet1, vtionet2 and vtionet3 on the same subnet as instance-A.
   - Instance-B will go into initial configuration and listen for cluster connections.

4. Configure the instance-A cluster links using either the BUI or the CLI.
   - See step 5 for BUI configuration
   - See step 6 for CLI configuration

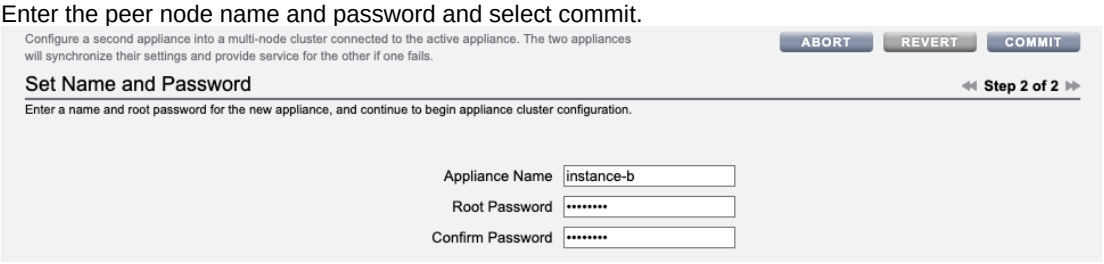
5. BUI: If running cluster setup in the BUI, fill in the peer field with the peer's OCID.
   - From the BUI, select the Configuration tab and go to CLUSTER.
   - Select Setup and enter the peer OCID in the dialog.



- When finished, the Cable Cluster page will be shown. Select Commit.



- Enter the peer node name and password and select commit.



6. CLI: If using the CLI and running a newer release, use the `linkinit` command.
   - Go to configuration cluster and run the `linkinit` command and then verify that the links are active.

```
instance-A:configuration cluster> linkinit ocid1.instance.oc1.iad…
instance-A:configuration cluster> links
 lio_dev/vtionet0 = AKCIOS_ACTIVE
```

7. CLI: Set up the cluster configuration.
   As peer instances, setting up the cluster will cause instance-B to load the configuration from instance-A for its initial configuration. This means instance-B will have all the same service configuration and the same network settings as instance-A.

```
instance-A:configuration cluster> setup

 Cable Cluster. Cable together the cluster card, power on the second appliance,
 and confirm that the communication links are active and connected.

 instance-A:configuration cluster setup cabling> done
 instance-A:configuration cluster setup identity> set nodename=instance-B
                       nodename = instance-b
 instance-A:configuration cluster setup identity> set password
 Enter new password:
 Re-enter new password:
                       password = (set)
 instance-A:configuration cluster setup identity> done
 instance-A:configuration cluster> show
 Properties:
                  state = AKCS_OWNER
            description = Active (takeover completed)
               peer_asn = 17861b69-6452-4e71-9cf6-bf096faa963e
          peer_hostname = instance-b
             peer_state = AKCS_STRIPPED
       peer_description = Ready (waiting for failback)

 Children:
                 resources => Configure resources
```

8. Set vtionet2 as private to instance-A. This step allows administrative (BUI/CLI/REST) access when instance-A is in passive (STRIPPED) cluster state.

```
instance-A:configuration cluster resources> show
Resources:

RESOURCE         OWNER           TYPE       LABEL         CHANGES  DETAILS
net/vtionet1     instance-a      singleton  vtionet1      no       192.88.99.40
net/vtionet2     instance-a      singleton  instance-A    no       192.88.99.52
net/vtionet3     instance-a      singleton  instance-B    no       192.88.99.25
zfs/p            instance-a      singleton                no       49G

instance-A:configuration cluster resources> select net/vtionet2
instance-A:configuration cluster resources net/vtionet2> set type=private
                          type = private (uncommitted)
instance-A:configuration cluster resources net/vtionet2> commit
```

9. (Optional) Test failover by rebooting instance-A so that instance-B automatically takeover cluster operation. For example:

```
$ ssh opc@192.88.99.52
Instance-A:> maintenance system reboot
This will reboot the appliance. Are you sure? (Y/N) y
Connection to 192.88.99.52 closed.
```

10. Configure instance-B peer network interface.
    - The IP address for vtionet3 will point to peer instance-A if it was originally configured on peer instance-A.
    - The first step is to make it point to the IP address for the peer instance-B system.
    - Then set the cluster type to private so that Peer instance-B can be accessible when it is passive.
      Note: You can use the console connection to debug network issues.

    For example:

```
$ ssh opc@192.88.99.25
instance-B:> configuration net interfaces "vtionet3"
instance-B:configuration net interfaces vtionet3> set v4addrs=100.102.217.56/20
                        v4addrs = 100.102.217.56/20 (uncommitted)
instance-B:configuration net interfaces vtionet3> commit

instance-B:configuration cluster resources> ls
Resources:

RESOURCE         OWNER           TYPE       LABEL         CHANGES  DETAILS
net/vtionet1     Instance-A      singleton  vtionet1      no       100.102.214.40
net/vtionet3     Instance-B      singleton  instance-B    no       100.102.217.56
zfs/p            Instance-A      singleton                no       49G

Instance-B:configuration cluster resources> select net
net/vtionet1  net/vtionet3
Instance-B:configuration cluster resources> select net/vtionet3
Instance-B:configuration cluster resources net/vtionet3> set type=private
                            type = private (uncommitted)
Instance0B:configuration cluster resources net/vtionet3> commit
```

11. Log in to instance-A using Instance-A private IP address.

```
% ssh opc@192.88.99.52
Last login: Thu Sep 23 16:05:37 2021 from 10.154.107.29
Instance-A:> configuration cluster
Instance-A:configuration cluster> ls
aksh: warning: terminal type "xterm-256color" unknown; using "vt100"
Properties:
                          state = AKCS_STRIPPED
                    description = Ready (waiting for failback)
                       peer_asn = 06237502-5a1d-4cbc-a5a7-abb7b28b1e28
                  peer_hostname = Instance-B
                     peer_state = AKCS_OWNER
               peer_description = Active (takeover completed)

Children:
                             resources => Configure resources
```

12. After confirming that takeover by instance-B is successful, initiate a takeover on instance-A, if you want instance-A to be the active instance.

```
Instance-A:configuration cluster> takeover
Continuing will immediately take over the resources assigned to the cluster
peer. This may result in clients experiencing a slight delay in service.
```

```
Are you sure? (Y/N) y
```

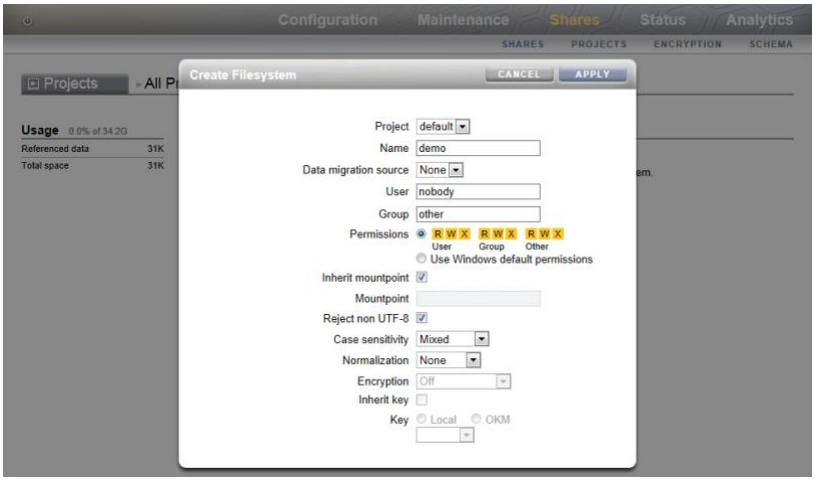
6. SHARE AN SMB FILESYSTEM

Complete the following steps to set up a simple filesystem share over Server Message Block (SMB) with Windows user access.

1. **Navigate to the Shares screen.**

   Click the add item icon next to Filesystems to create a new filesystem.

2. **Name the filesystem and change the permissions for Group and Other to allow anyone to read, write, and execute on the filesystem.**
   In this example, the filesystem is named demo. The filesystem is part of the default project. Click APPLY to save the changes.

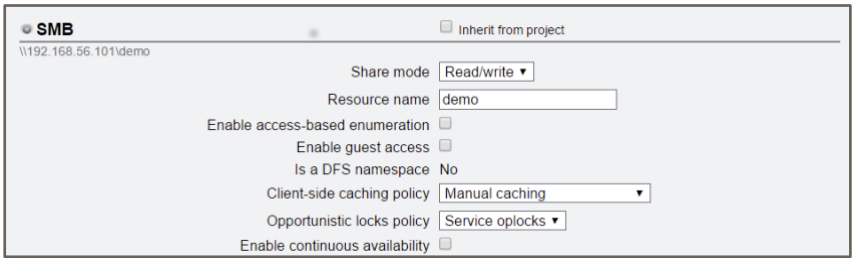


3. In the Shares screen, mouse over the entry for the new filesystem and click the edit icon to edit the filesystem attributes.

4. Click Protocols.

5. In the SMB section, clear the checkbox for Inherit from project, select Read/Write in the Share mode drop-down list, and set the Resource Name.

   In this example, the Resource Name is demo. Click APPLY to save the changes.



6. Select the Configuration tab to access the Configuration Services screen.

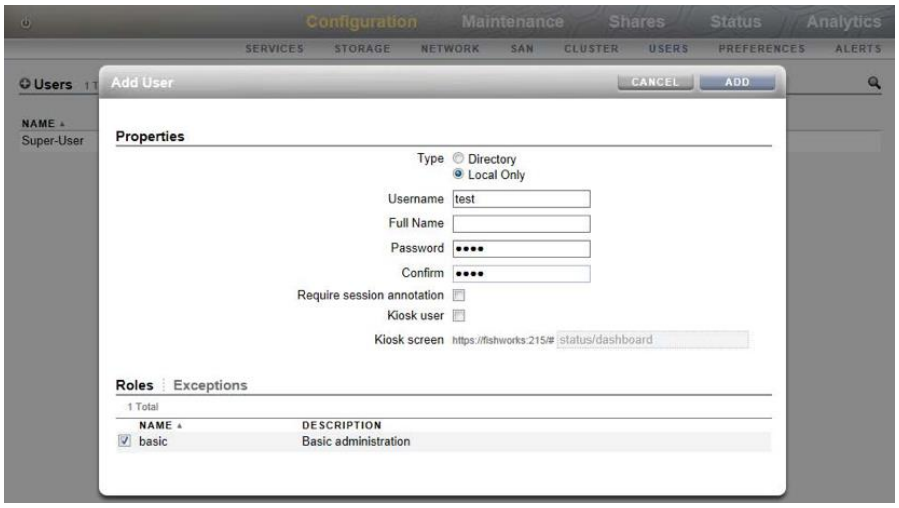
7. Enable the SMB service by clicking the power icon.

   The state will change from Disabled to Online.

8. Configure a user with access to the filesystem share.

   a. Click USERS in the navigation bar, and click the add item icon next to Users to create a new user.

   b. Select Local Only, set the Username and Password, and click ADD. Log out of the BUI by clicking LOGOUT near the top of the screen.



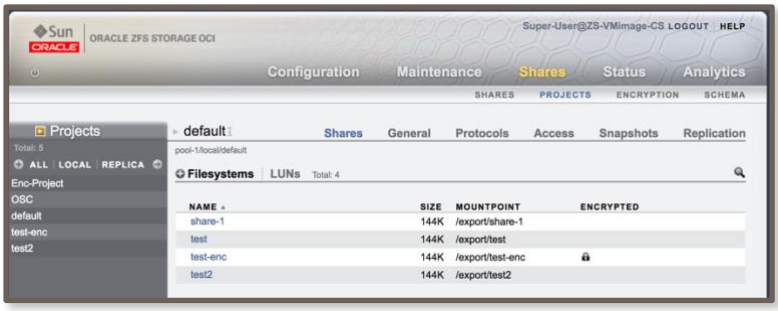
9. From a Windows client, connect to the IP address of your ZFS Storage instance, and log in with the credentials you set in step 8 to access the shared filesystem.

7. SHARE AN NFS FILESYSTEM

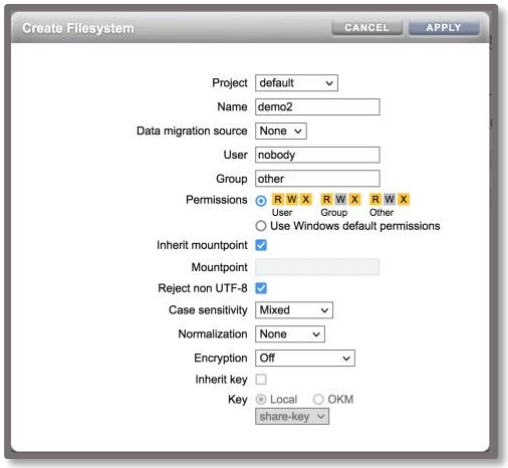
Complete the following steps to set up a simple filesystem over NFS to share with an NFS client or clients.

1. **Navigate to the Shares screen.**

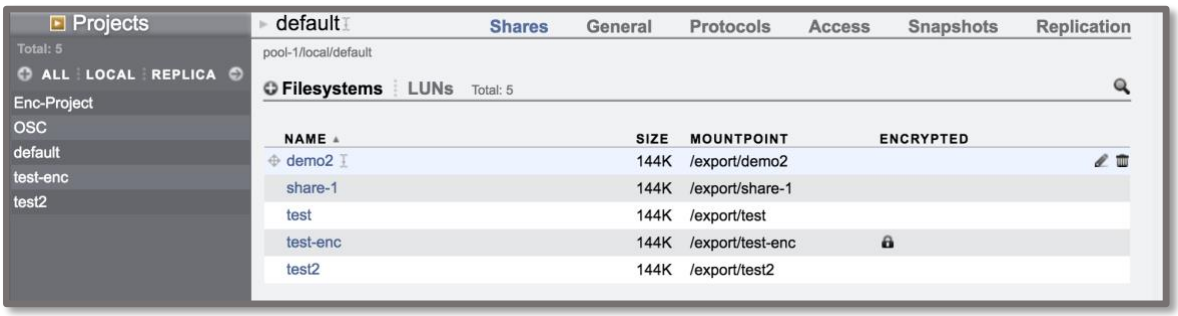
   Click the add item icon next to Filesystems to create a new filesystem. Projects provide an administrative point for filesystems so you can set properties at the project level that are inherited by filesystems within the project. The system includes the default project.



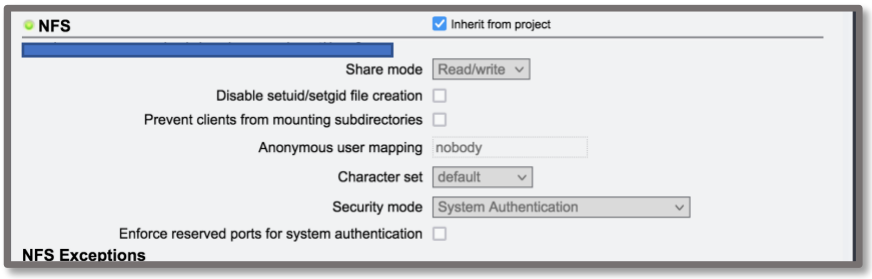
2. **Name the filesystem and change the permissions to match the user/group requirements.**
   In this example, the filesystem is named demo2. The filesystem is part of the default project. Click APPLY to save the changes.



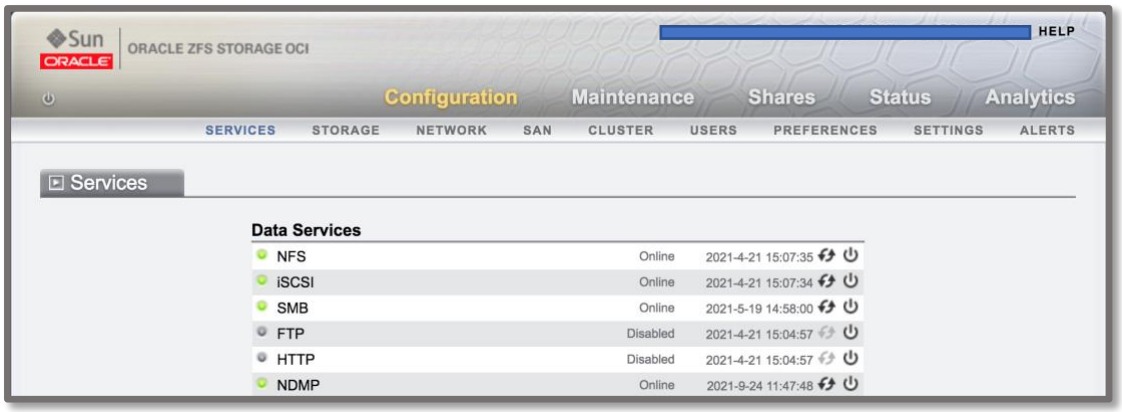
3. In the Shares screen, mouse over the entry for the new filesystem and click the edit icon to edit the filesystem attributes.



4. Click Protocols.
5. In the NFS section, set the Share mode to Read/write in the pulldown menu, if it is not inherited from the project. Click APPLY.



6. Select the Configuration tab to access the Configuration Services screen.
7. Enable the NFS service by clicking the power icon if it is not already enabled.



8. Mount the filesystem over NFS with syntax similar to the following:

```
% mount -t nfs ZFS-IP:/export/demo2 /mnt
```

UPGRADING YOUR ZFS STORAGE INSTANCE

When new ZFS Storage images are available from the OCI Marketplace listing, you can upgrade your running instances. For information about upgrading your ZFS Storage instances, see the following doc: Oracle Support Document 2817714.1 (How to Upgrade a ZFS Storage on Oracle Cloud Infrastructure (OCI) Marketplace Deployment) can be found at: https://support.oracle.com/epmos/faces/DocumentDisplay?id=2817714.1.

ZFS IN OCI INSTANCE BEST PRACTICES

Network Best Practices

- The primary OCI VNIC should be used for iSCSI traffic.
- A secondary OCI VNIC should be created for NAS traffic.

- A secondary IP on the secondary OCI VNIC should be used for NAS traffic. Since the secondary IP address can moved, it allows for easier migration of NAS traffic to different ZFS Storage in OCI instances.
- For information on configuring a secondary OCI VNIC, see Managing VNICs.

ZFS Storage in OCI Network Routing

- It is recommended to set the multihoming model to strict.
- Create a network IPv4 route on the primary network interface with the destination set to 169.254.0.0/16 for iSCSI traffic to increase network throughput.

ZFS Storage in OCI Network Datalinks

- Link Speed, Link Duplex and Flow Control should all be set to Auto.
- Link speed for VM instances will be reported as 1GB but will actually use the full amount of bandwidth allocated to the instance. (See known issues)
- All network datalinks should have the MTU set to 9000 for best performance.

ZFS Storage in OCI Network Interfaces

- The primary network interface used for iSCSI traffic should not be modified because it can cause a system panic. (See known issues)
- Consider using separate subnets for storage administrators and NAS clients for enhanced security.
- NAS client interfaces should uncheck 'Allow Administration' for enhanced security.

Block Storage Best Practices

System Boot Disk

- System disk contains read only OS image, logs, core dumps and configurations.
- Configuration data can be backed up using 'Maintenance System Configs'
- Does not include OS image, logs, core dumps, replication or share data.
- Logs and core dumps can be saved using 'Maintenance System Bundles'
- Entire system disk can be backed up using OCI boot volume backups.

Storage Pools

- Pool disks contain all configuration data under 'Shares'
- All disks in each pool should be same size especially if they are under 800GB.
- All data disks in each pool should have the same performance settings.
- Suggest creating a volume group containing all data disks for each storage pool.
- Block volume backups must use volume groups to keep pool data consistent.
- For best system resource usage recommend only one pool per VM.
- All data disks provided by OCI have multiple copies so stripped pools provide data protection. ZFS will detect bit rot but data will have to be restored from backup if bit rot is detected.
- Consider backing up data disks or using a parity or mirrored storage profile to protect against bit rot or a block volume outage.

Backup of ZFS Configuration

We recommend that after your ZFS Storage in OCI instance is configured, that you create a backup of the configuration with the following steps:

- From the Appliance BUI, go to Maintenance→System.
- Under the Configurations section, click Backup.
- This will create a backup of the Appliance configuration, that can be downloaded and stored separately for recover purposes.

For information about the configuration backup content, what is included and what is not included, see Backing Up the Configuration.

Block Volume Backups

OCI Block Volume Service allows you to create snapshots of both boot volume and block volumes.

- Boot Volume snapshots
  - https://docs.oracle.com/en-us/iaas/Content/Block/Tasks/backingupabootvolume.htm
- Block Volume Backups
  - https://docs.oracle.com/en-us/iaas/Content/Block/Concepts/blockvolumebackups.htm

SECURITY REFERENCES

For information about setting permissions on shares and recommended security practices, see the following references:

- Access Control Lists for Filesystems
- Oracle® ZFS Storage Appliance Security Guide, Release OS8.8.x

APIS FOR ZFS STORAGE IN OCI

Initial Configuration

The on-premise ZFS Storage Appliance uses a manual based initial configuration over the serial console that sets up the initial network settings and the root password. Console configuration is still supported if the ssh keys are not supplied to the instance during launch. A serial console connection can be made by using the *Console Connection* on the OCI compute instance page.

If user SSH keys are defined for the instance then they will be automatically applied to the opc and root users during the initial configuration. The password will be set to a long random value to disable password based access. Initial access will only be available via ssh until the password is set.

The initial install will try and use DHCP settings to set the initial configuration properties usually populated from the console. If DHCP does not provide all the values then a *Console Connection* still needs to be mode to do the initial install.

Root User Configuration

You will need to configure the root user to perform some tasks such as taking a configuration backup, configurating replication, or even logging in remotely as root or using the su to root command.
To enable root login over ssh, from the Appliance BUI, go to the Configuration tab to reach the Configuration Services screen. Under Remote Access, select the ssh service. From the ssh service screen, enable Permit root login.
For more detailed configuration information, see <MOS Doc ID 2811414.1>.

Initial Setup

Initial setup for on-premise ZFS Storage Appliance is a guided manual setup that configures storage, DNS, naming services, NTP, network and phone home.

OCI compute instances have a special metadata called "user_data" field that can be used for configuring the compute instance on initial boot. If this field is applied then the ZFS Storage in OCI will run the configuration script. See Working with CLI Scripting in the Oracle® ZFS Storage Appliance Administration Guide

ZFS Storage in OCI will also look for "config_data" metadata and if it exists it will be used instead of "user_data". The reason for this is OCI does not allow "user_data" to be modified, and "config_data" can be modified and deleting "config_data" after the system has been configured can help stop leaking of configuration information through OCI compute metadata.

The install script can define or override initial configuration properties that are usually populated by DHCP to guarantee that manual configuration via the *Console Connection* is not needed.

DNS is usually configured during initial configuration but the script can add additional DNS server. NTP is auto configured for the OCI environment and enabled by default so it does not need to be configured. If more than one NIC is used it is recommenced to change the routing multihoming to strict.

Example User Data Script to configure a system. This script sets up initial configuration and also set up static routing for iSCSI. Note: Since OCI only allows one user script the configuration properties are defined as comments so that when the CLI script is ran they will be ignored.

```
#@hostname=fishworks
#@domain=example.com
#@ip_addr=203.0.113.29
#@ip_mask=255.255.255.0
#@router=203.0.113.29
#@dns_servers=203.0.113.125
```

```
script

print('setting up DNS...');
run('top configuration services dns');
run('create');
set('address', '203.0.113.126');
run('commit');

run('top configuration net routing');
prop('multihoming', 'strict');
run('commit');
.
```

REST API

The ZFS Storage Appliance has a REST API for managing all aspects of an on-premise ZFS Storage Appliance but some additional functionality is required for integrate with cloud orchestration services. This REST API expands the ZFS REST API to manage OCI resources that are not available from the on-premise ZFS Storage REST API.

The new REST API endpoint will be available from the /api/oci/v2 resource. This resource path will be hidden and its only supported use will be by Oracle IT within OCI. BUI and CLI interfaces for these resources will not be made available since their purpose is for use with cloud orchestration software.

Instance Information

OCI compute instances are able to query OCI to get information about itself. The instance endpoint simply returns the instance information returned from the OCI service: http://169.254.169.254/opc/v1/instance/

The content is defined by the OCI instance information service defined at https://docs.cloud.oracle.com/iaas/Content/Compute/Tasks/gettingmetadata.htm

The purpose of supplying instance information is to enable cloud monitoring and orchestration use cases.

Example Request

```
GET /api/oci/v2/instance
```

Example Response

```
{
    "instance": {
        "compartmentId":
"ocid1.compartment.oc1..aaaaaaaadrcsphnhoji76iyemczsoh2ggmokaaka7fxzm772hrjpxqfne
slq",
        "displayName": "fishworks",
        "timeCreated": 1573235633999,
        "agentConfig": {
            "managementDisabled": true,
            "monitoringDisabled": true
        },
        "canonicalRegionName": "us-phoenix-1",
        "shape": "BMStandard2.52",
        "availabilityDomain": "iZbs:PHX-AD-1",
        "id":
"ocid1.instance.oc1.phx.anyhqljsuna655qccpnt2hmojteyoljyh43fvfalzuljgurzogxrkj5tc
ghq",
        "state": "Running",
        "definedTags": { },
        "region": "phx",
        "faultDomain": "FAULT-DOMAIN-1",
        "image":
"ocid1.image.oc1.phx.aaaaaaaaxqr2he5lcipvthyxru2lllstb3jzgfg6gyqxl7xb45qt6evn2lwa
",
```

```
    }
}
```

iSCSI Targets

A hardware based ZFS Storage Appliance does not act as an iSCSI initiator and has no API to manage the available iSCSI targets. The iscsi-targets resource allows clients to manage the iscsi-targets attached to the ZFS instance.

Important! Dual attachment of LUNs to multiple ZFS instances is not supported. All iSCSI LUNs should be attached to only one system at a time. Failure to do so will result in data corruption!

The purpose of this resource it to enable initial setup and migration if iSCSI LUNs between ZFS instances.

iSCSI Target Resource Properties

PROPERTY	TYPE	DESCRIPTION
addr	string	iSCSI target address:port
iqn	string	iSCSI target iqn
pool	immutable string	The pool name if the LUN associated with the iSCSI target is part of a zfs pool.

List iSCSI Targets

The get command will return the currently attached iSCSI LUNs available for use as storage pools on the ZFS Storage instance. The available number of iSCSI targets should match the number of storage devices available for pools. If a device is configured as part of a storage pool, the pool name will be provided.

Example Request

```
GET /api/oci/v2/iscsi-targets
```

Example Response:

```
{
    "targets": [{
        "iqn": "iqn.2015-12.com.oracleiaas:48558e46-6cf0-4b27-b8c2-f8bd69812305",
        "addr": "169.254.2.5:3260",
        "pool": "p1"
    }, {
        "iqn": "iqn.2015-12.com.oracleiaas:bff0f183-736e-49c3-9478-b1352407262f",
        "addr": "169.254.2.6:3260",
        "pool": "p1"
    }]
}
```

Create iSCSI Targets

This command will attach all the specified OCI iSCSI targets to the ZFS Storage instance. The most efficient method to bring all storage online is to attach all LUNs using a single command. When the command returns the LUNs associated with the iscsi targets should be available on the ZFS Storage system. To import or clear any pools associated with the new devices, use the pool resource documented below.

Generate a list of all volumes attached to a compute instance that can be used as the POST data for the command, run the following:

```
INSTANCE_ID=ocid1.instance.oc1.phx.anyhqljruav6g7ic63hz2jkm5ywxyjz3y23xh
oci compute volume-attachment list --instance-id $INSTANCE_ID \
    --query 'data [*].{iqn:"iqn",addr:"ipv4"}'
```

Example Request

```
POST /api/oci/v2/iscsi-targets
[{
     "iqn":"iqn.2015-12.com.oracleiaas:48558e46-6cf0-4b27-b8c2-f8bd69812305",
     "addr":"169.254.2.5:3260"
}, {
    "iqn":"iqn.2015-12.com.oracleiaas:bff0f183-736e-49c3-9478-b1352407262f",
    "addr":"169.254.2.6:3260"
}, {
    "iqn":"iqn.2015-12.com.oracleiaas:6a9b6b07-aaf3-44ce-82ba-b2ae59b11d4e",
    "addr":"169.254.2.7:3260"
}]
```

The response will be a JSON object containing a "target" property whose value is the list of iscsi-targets that were created.

Remove a single iSCSI Target

The delete command on a specified IQN resource will only remove the specified iSCSI LUN. This is useful for modifying an existing system.

```
DELETE /api/oci/v2/iscsi-targets/<iqn>
```

Remove All iSCSI Targets

The delete command will remove all iSCSI LUNs from the system. Any LUN that is part of a pool will not be deleted.

Example Request

```
DELETE /api/oci/v2/iscsi-targets
```

Additional Configuration

Apply a configuration from a mirrored system disk. This will have the same effect as saving a system configuration backup and applying it to the system.

Property	Type	Description
addr	string	iSCSI addr of target containing cloned boot disk
iqn	string	iSCSI iqn of target containing cloned boot disk
locked	boolean	Flag determining if system configuration is locked. When locked a warning will be given on login that system configuration changes will not be propagated to the destination system. Storage pools will also not be imported if the system reboots to avoid dual import when using shared disks.
version	string	system version to import (optional, default is to use latest version on system disk)
guid	string	System pool guid (optional, default will be to use the first importable system disk guid found)
status	immutable string	One of STARTING, RUNNING, FAILED, COMPLETED showing full job status.
config_status	immutable string	Set to COMPLETED or FAILED once the system config is ready for client I/O. Some background tasks may still be running such as collecting lock data for phone-home support.
started	immutable string	job start time in iso8601 formatted UTC date
finished	immutable string	job end time in iso8601 formatted UTC date
message	immutable string	Human readable status and error messages for debugging.

Lock Configuration

Lock configuration on original source compute instance.

```
PUT /api/oci/v2/config
{
    "locked":true
}
```

Start Configuration Import Job

Example command to start a configuration import job. The addr and iqn properties are the values of a iscsi settings of the system disk attached to the ZFS instance.

The configuration from the old ZFS attached system disk will be applied to the new ZFS system.

```
PUT /api/oci/v2/config
{
    "addr": "169.254.2.18:3260",
    "iqn": "iqn.2015-12.com.oracleiaas:36bcc330-c1a0-4494-ad66-76c88cb1d044"
}
```

Get Configuration Job Properties

Get details of a running config job.

```
GET /api/oci/v2/config
```

```
{
    "config": {
        "addr": "169.254.2.18:3260",
        "iqn": "iqn.2015-12.com.oracleiaas:36bcc330-c1a0-4494-ad66-
76c88cb1d044",
        "status": "COMPLETED",
        "started": "2020-09-25T16:50:27",
        "config_status": "COMPLETED",
        "finished": "2020-09-25T16:51:50",
        "version": "ak-nas-2013.06.05.8.23.3_13.7-2",
        "message": "2020-09-25T16:51:50 Restored system configuration\n2020-
09-25T16:51:50 Waiting for background logs to be copied\n2020-09-25T16:51:50
completed",
        "guid": "17894726420108290233"
    }
```

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com.
Outside North America, find your local office at oracle.com/contact.

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

SOLUTION BRIEF
September 2021
Author: zfssa-storage-feedback_ww_grp@oracle.com