ORACLE
ZFS STORAGE APPLIANCE

Oracle ZFS Storage Appliance with Oracle Private Cloud Appliance

Configuration Best Practices for Supporting Oracle Database

ORACLE WHITE PAPER | APRIL 2016

ORACLE

Table of Contents

Introduction

This paper provides guidance in the configuration of the Oracle ZFS Storage Appliance with the Oracle Private Cloud Appliance.

Highlighted in this paper are:

- Connectivity options for Oracle ZFS Storage Appliance resources with the Oracle Private Cloud Appliance 2.0
- Oracle Linux configuration and deployment on the Oracle Private Cloud Appliance (formerly called Oracle Virtual Cloud Appliance)
- Configuration of the Oracle ZFS Storage Appliance for use by Oracle Database 12c running on the Oracle Private Cloud Appliance, including Direct NFS (dNFS) and Oracle Integrated Storage Protocol (OISP)

NOTE: References to Sun ZFS Storage Appliance, Sun ZFS Storage 7000, and ZFS Storage Appliance all refer to the same family of Oracle ZFS Storage Appliances.

NOTE: This document assumes the Oracle Private Cloud Appliance Release 2.0 Software and Oracle ZFS Storage Appliance Software OS8.4.

About Oracle ZFS Storage Appliance

The Oracle ZFS Storage Appliance is designed to provide high performance, flexibility and scalability. The Oracle ZFS Storage Appliance provides multiple connectivity protocols for data access, including: Network File System (NFS), Common Internet File System (CIFS), Internet Small Computer System Interface (iSCSI), InfiniBand (IB), and Fibre Channel (FC). It also supports the Network Data Management Protocol (NDMP) for backing up and restoring data. The Oracle ZFS Storage Appliance architecture also offers the Hybrid Storage Pool (HSP) feature, in which direct random access memory (DRAM), flash and physical disks are seamlessly integrated for efficient data placement (see Figure 1). A powerful performance monitoring tool called DTrace Analytics provides details about the performance of the various components, including network, storage, file systems, and client access. The tool also offers plenty of drill-down options that allow administrators to monitor specific rates of latency, size of transfer, and utilization of resources. The Oracle ZFS Storage Appliance provides a variety of RAID protections to balance the capacity, protection, and performance requirements of the applications, databases, and virtualized environments.

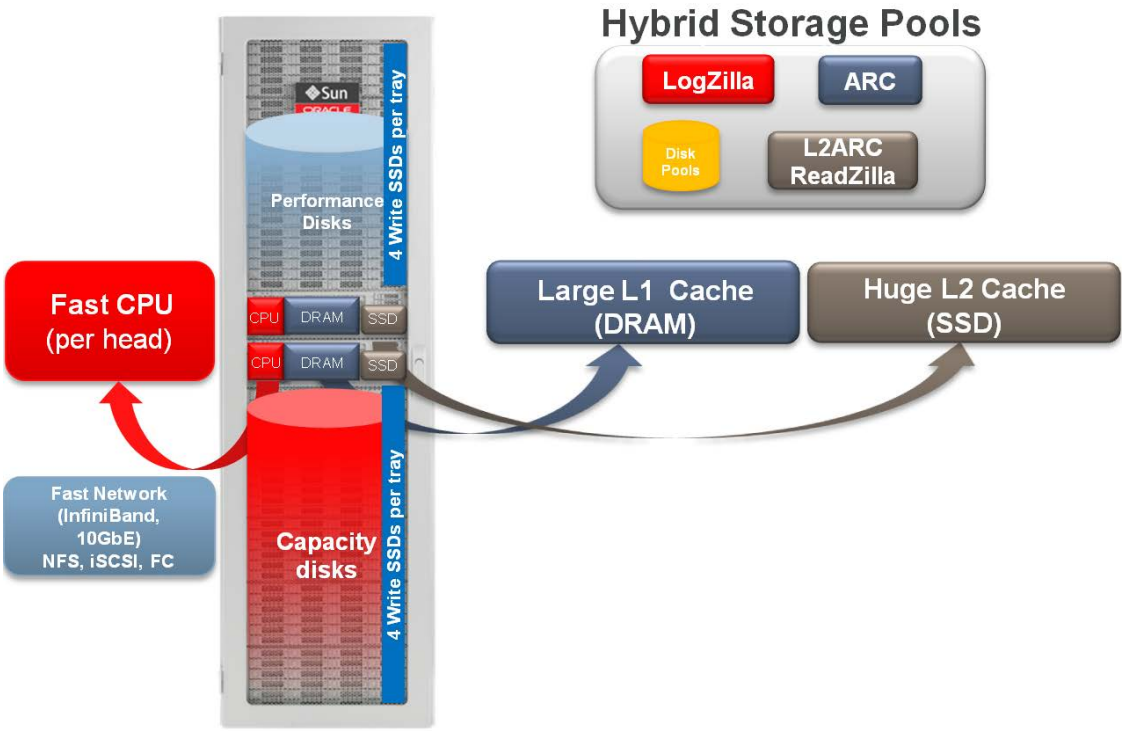

Figure 1. Oracle ZFS Storage Appliance – architecture overview

About Oracle Private Cloud Appliance

Oracle Private Cloud Appliance, formerly named Oracle Virtual Cloud Appliance, is an integrated, "wire once," software-defined converged infrastructure system designed for rapid deployment of a private cloud at an industry-leading price point.

The Oracle Private Cloud Appliance includes compute nodes, management nodes, virtual networking, and integrated Oracle ZFS Storage ZS3-ES for internal storage.

Compute Nodes: Compute nodes include Oracle Server X5-2 systems with Intel Xeon CPUs, high-speed dual inline memory modules (DIMMs), redundant 40 Gb/sec InfiniBand host channel adapters (HCAs), and redundant disks. Each compute node runs Oracle VM Server for x86 to provide server virtualization.

Management Nodes: The Oracle Private Cloud Appliance contains redundant management nodes integrated into the rack and the networking fabric.

Virtual Networking: Oracle Private Cloud Appliance uses ultra high-performance Oracle Fabric Interconnect, a component of the Oracle Virtual Networking family. Each Private Cloud Appliance hardware configuration contains multiple redundant quad data rate (QDR) InfiniBand switches and

Oracle Fabric Interconnect systems that serve as gateways to the data center's Ethernet network and to external storage attached using 10GbE Ethernet, InfiniBand, or Fibre Channel.

Integrated Storage: Oracle Private Cloud Appliance features a fully integrated, enterprise-grade Oracle ZFS Storage ZS3-ES for centrally storing the management environment as well as providing data storage for virtual machines (VMs).

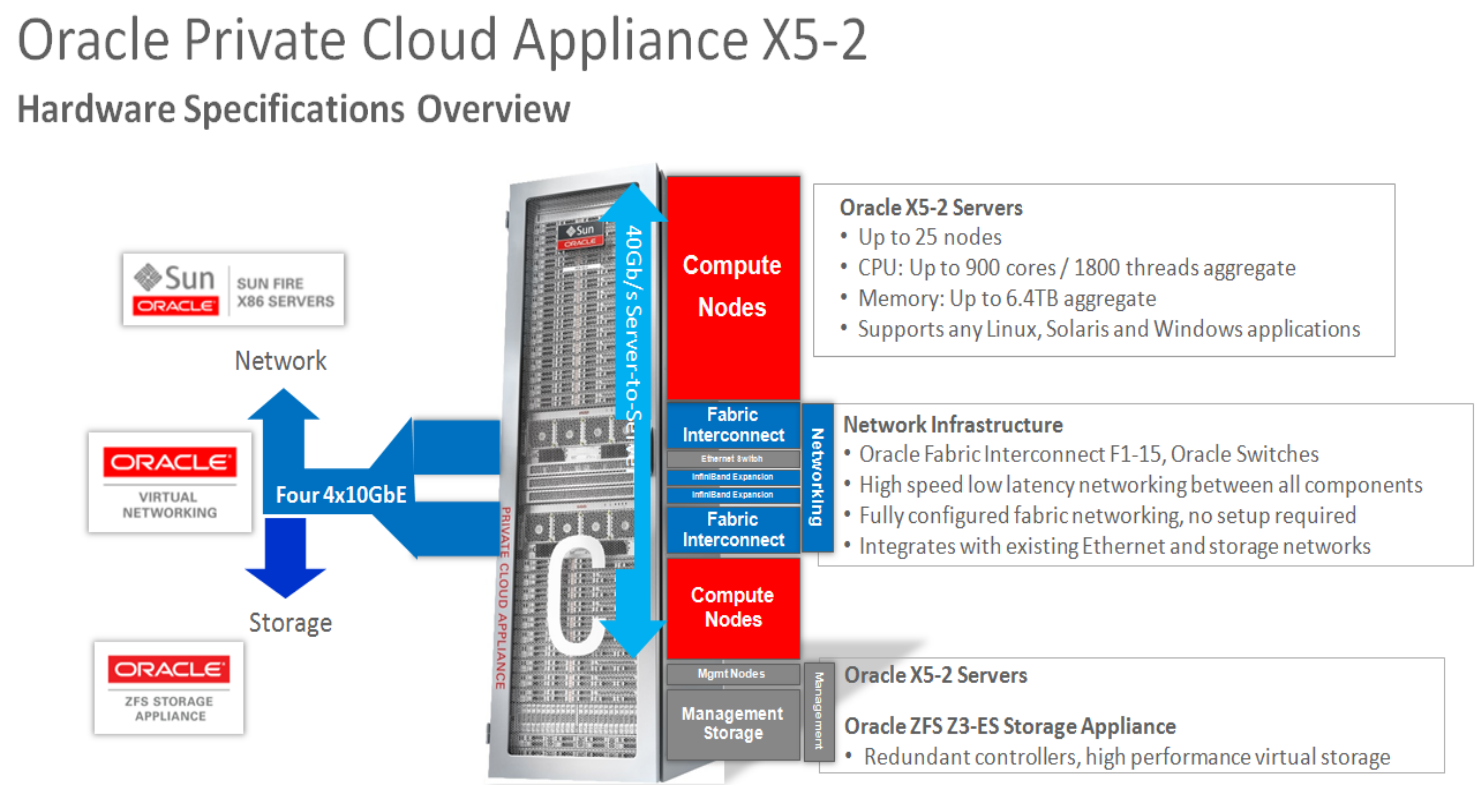

Figure 2. Oracle Private Cloud Appliance – hardware overview

Private Cloud Appliance Storage

Each management node and compute node in the Oracle Private Cloud Appliance has its own local storage consisting of two 1.2TB hard drives mirrored using a hardware RAID controller. These drives contain the Oracle VM hypervisor system and Oracle VM related software. The Oracle Private Cloud Appliance rack contains a pre-installed Oracle ZFS Storage ZS3-ES. The ZS3-ES is connected to the Oracle Private Cloud Appliance management and compute nodes through the integrated InfiniBand fabric and has dual controllers and a single disk tray with 4 800GB SSDs and 20 900GB hard drives, providing a total of 11.3 TB of user storage. The internal Oracle ZFS Storage ZS3-ES is primarily intended for use as a high-performance repository for the necessary Oracle Private Cloud Appliance

management databases and configuration metadata storage, storage of Oracle VM templates and OS images, and other infrastructure storage needed by the Oracle Private Cloud Appliance and Oracle VM environments. While the Oracle ZFS Storage ZS3-ES storage is accessible for other use, it cannot be expanded and can be limiting for application and database workloads.

An external Oracle ZFS Storage Appliance like the ZS3-2 or ZS4-4 is an ideal solution for storage expansion of the Oracle Private Cloud Appliance. It has been co-engineered with the Oracle Private Cloud Appliance to maximize performance and efficiency while reducing deployment risk and total cost of ownership. As an engineered storage expansion, the Oracle ZFS Storage Appliance provides for the Oracle Private Cloud Appliance environment the following capabilities:

- Extremely high performance for applications and workloads deployed on the Oracle Private Cloud Appliance. The Oracle ZFS Storage Appliance is optimized for workloads with intensive input-output operations per second (IOPS), such as online transaction processing (OLTP) databases, as well as for bandwidth-driven workloads including data warehousing, business intelligence analytics, and video processing. The Oracle ZFS Storage Appliance is powerful enough to run a diverse set of workloads concurrently by leveraging the Oracle Private Cloud Appliance InfiniBand network or the 10GbE network.
- High throughput for thousands of virtual machines (VMs). It is capable of booting 16,000 VMs in approximately seven minutes. The Oracle ZFS Storage Appliance provides fast response time and low latency, which allows it to avoid VM boot storms, throughput congestion, hot spots, and I/O saturation.
- Superior storage analytics, which allow you to visualize and drill down into specific workloads to understand where congestion occurs and why. The Oracle ZFS Storage Appliance can even allow you to examine and manage the storage aspects of Oracle Private Cloud Appliance environments all the way down to the VM level.
- Scalable capacity, which can expand up to 3.4 petabytes (PB) and offer multiple configurations to address different application needs.
- Less risks. The Oracle ZFS Storage Appliance reduces risks by automating storage management using Oracle Enterprise Manager, so you have fewer storage systems to integrate and manage. It also lowers risk by providing leading fault-monitoring and self-healing capabilities, and by simplifying setup and management through its DTrace Analytics feature.
- Less complexity. The Oracle ZFS Storage Appliance's large DRAM and flash cache–based architecture is more efficient in serving the I/O from large virtualized environments. In addition, Oracle ZFS Storage Appliance's unique features, such as its Hybrid Columnar Compression, reduce the amount of storage needed for data warehouses built using Oracle Database. As noted with its proposition of less risk, less required storage means less management and less cost of ownership.

Connectivity to an external Oracle ZFS Storage Appliance is either through the Oracle Private Cloud Appliance 40 Gb/sec InfiniBand fabric, through optional Fibre Channel interfaces, or through the Oracle

Private Cloud Appliance 10GbE public network. Connecting to the InfiniBand fabric provides extremely high bandwidth, but the Oracle ZFS Storage Appliance must be within 5 meters (about 15 feet) of the Oracle Private Cloud Appliance. Fibre Channel provides more flexibility than InfiniBand in distance and connectivity, but when compared to 10GbE there are management and cost tradeoffs. The Oracle Private Cloud Appliance comes standard with available public 10GbE ports in the dual Oracle Fabric Interconnect F1-15s that are certified for external 10GbE storage attachment. This paper will illustrate 10GbE connectivity to the Oracle ZFS Storage Appliance.

Connecting the Oracle Private Cloud Appliance to the Data Center Switches Using 10GbE

The Oracle Private Cloud Appliance connectivity to external 10GbE-attached storage must be accomplished through the designated 10GbE Public I/O module ports on the Oracle Fabric Interconnect F1-15 units in the Oracle Private Cloud Appliance rack. These ports should be attached in a redundant manner to two data center 10GB switches. For details about physically connecting the 10GbE infrastructure on the Oracle Private Cloud Appliance to your data center switches, refer to both the Oracle Private Cloud Appliance's Installation Guide and Administration Guide at http://www.docs.oracle. com/cd/E55813_01/.

Choosing and Configuring the Oracle ZFS Storage Appliance

The following section provides best practices for optimizing an Oracle ZFS Storage Appliance for attachment to the Oracle Private Cloud Appliance serving Oracle Database 12c.

Choosing a Controller

The Oracle ZFS Storage Appliance is available in two models: the ZS4-4 and ZS3-2 with the following maximum configurations:

TABLE 1. ORACLE ZFS STORAGE APPLIANCE DETAILS

Features	ZS3-2	ZS4-4
CPU Cores	32	120
DRAM	512 GB or 1 TB	2 TB
Write Optimized Flash	4 TB	10.5 TB
Read Optimized Flash	12.8 TB	12.8 TB
Max Raw Capacity	1.5 PB	3.5 PB
HA/Cluster Option	Yes	Yes
Focus	Mid-Range	Scalability

The Oracle ZFS Storage ZS4-4 is the flagship product that offers maximum levels of scalability, CPU, and dynamic random-access memory (DRAM). With the potential to scale up to 2 TB of DRAM, 10 TB of write-optimized flash, and 12 TB of read-optimized flash, this is a highly scalable platform that can support up to 3.5 PBs of raw storage capacity.

The Oracle ZFS Storage ZS3-2 is a cost-efficient model that can still achieve extremely high levels of throughput and redundancy. It can support up to 1 TB of DRAM, 4 TB of write-optimized flash, 12 TB of read-optimized flash and 1.5 PB of raw storage capacity.

Both of these models are excellent choices for attachment to the Oracle Private Cloud Appliance. When making the decision on which model is best suited for your environment, consider these factors:

- If the Oracle ZFS Storage Appliance will be used 100 percent exclusively for hosting Oracle Database workloads from the Oracle Private Cloud Appliance, write- and read-optimized cache is often recommended to achieve good performance and usability with most non-backup database I/O, incrementally updated backup workloads, cloning for dev/test provisioning, and many other mixed I/O scenarios. Having an Oracle ZFS Storage Appliance with a significant amount of write-optimized flash increases the flexibility for the type of workloads that it can be effectively used for. This may be important for current or future activity planning.
- If running direct transactional database workloads is a primary focus of the system, then having a large amount of DRAM and CPU resources will be a significant benefit.
- The Oracle ZFS Storage ZS3-2 offers exceptional throughput and redundancy at an extremely low price point. The Oracle ZFS Storage ZS3-2 is well suited for smaller Oracle Private Cloud Appliance configurations.

Choosing the Correct Disk Shelves

The Oracle ZFS Storage Appliance offers two options for disk shelves, both with similar price points. The Oracle Storage Drive Enclosure DE2-24C features high-capacity 4 TB disks and the Oracle Storage Drive Enclosure DE2-24P features high-performance 900 GB disks. Each shelf contains 24 disks and both can be configured with the same write-optimized flash options (up to four disks per shelf can be replaced with SSD write flash accelerators). The Oracle ZFS Storage ZS4-4 and ZS3-2 can be customized based on disk shelf and write-optimized flash requirements.

TABLE 2. ORACLE STORAGE DRIVE ENCLOSURE MODEL (DISK SHELF) DETAILS

Disk Shelf	Size/Disk	RPM	IOPS/Disk	MBPS/Disk	Rack Units
DE2-24P	900 GB	10K	160	170	2
DE2-24C	4 TB	7.2K	120	180	4

The high-capacity Oracle Storage Drive Enclosure DE2-24C disk shelf is recommended when ultra-low-latency performance is not necessary. The larger capacity and slightly higher sequential throughput give the Oracle Storage Drive Enclosure DE2-24C a significant advantage for most data warehouse and sequential use cases. The Oracle Storage Drive Enclosure DE2-24P may be considered in situations where the higher IOPS and lower latency would be a significant advantage or if rack space is a limiting factor and the desire is to maximize performance in a small partial-rack configuration.

Choosing a Storage Profile

When selecting a storage profile for Oracle Databases, mirrored or single parity are recommended unless the pool is being used only for Oracle Recovery Manager (RMAN) backup or archival databases.

TABLE 3. STORAGE PROFILE COMPARISON

Storage Profile	Usable Capacity	Advantages	Negatives
Mirrored	42.2%	Maximum Random Read Performance Maximum Protection Maximum Flexibility	Costly
Single Parity	69.3%	General Write Performance Moderate Flexibility	Lower Redundancy
Double Parity	76.7%	Streaming Performance Most Space Efficient	Lower IOPS

Mirrored

Mirrored is a frequently recommended storage profile due to its strong redundancy and robust performance, particularly for database restores or random read performance. Because it generates twice as many virtual devices (vdevs) as a single parity implementation, a mirrored storage pool is capable of handling far more IOPS. This capability gives it the flexibility to achieve exceptional performance with workloads that generate small random I/O, such as direct database OLTP transactions as well as perform well with large, sequential I/O such as traditional RMAN workloads.

The downsides of choosing a mirrored profile are that it consumes more disk space than the other two options and generates more internal bandwidth on writes, which would be impactful in situations where serial attached SCSI (SAS) or Peripheral Component Interconnect (PCI) bandwidth are limiting factors.

Mirrored is recommended when there is an emphasis on achieving optimal OLTP performance, and sequential read performance, such as RMAN restores.

Mirrored is recommended in situations where dev/test database provisioning or direct database workloads are a focus. If the database runs an OLTP workload (characterized by mostly small transactions with a focus on changing and reading existing rows of data) or has an element of write transactions dispersed

throughout the day, then this will generate small random I/O that will place an IOPS load on the storage pool with these use cases. A mirrored profile is best suited to handle heavy IOPS.

Single Parity

Single parity is a middle of the road option that is particularly attractive when usable capacity concerns would make the mirrored storage profile a poor fit.

Single parity implements a narrow 3+1 stripe width and utilizes powerful ZFS features to provide exceptional performance with large streaming I/Os but also enough flexibility to handle some random or smaller I/O workloads.

Because single parity uses a narrow stripe width, it still generates a moderate amount of vdevs (half as many as mirrored) and has flexibility to handle some workloads that generate non-streaming large I/O, but an IOPS-intensive workload such as an incrementally updated backup strategy for an OLTP database should utilize a mirrored profile for data file copies.

Double Parity

Double parity provides the best usable capacity and will perform almost as well as single parity on large streaming I/O by utilizing a wide stripe width. The width varies at the time of storage pool creation, depending on the number of disks in the configuration. The stripe width ranges up to 14 disks. As a result, the number of vdevs in a double parity storage pool is far fewer than with mirrored or single parity. The ability to handle direct IOPS-intensive workloads is severely diminished.

Double parity is only recommended for situations where the Oracle ZFS Storage Appliance is 100 percent dedicated to large sequential workloads such as traditional RMAN backup and restore, media streaming, or very inactive archival databases. Double parity is not recommended if there will possibly be additional use cases such as cloning for dev/test provisioning or utilizing an incrementally updated backup strategy. Mirrored or single parity profiles are more flexible for handling additional use cases that may result in heavier disk IOPS with lower latencies.

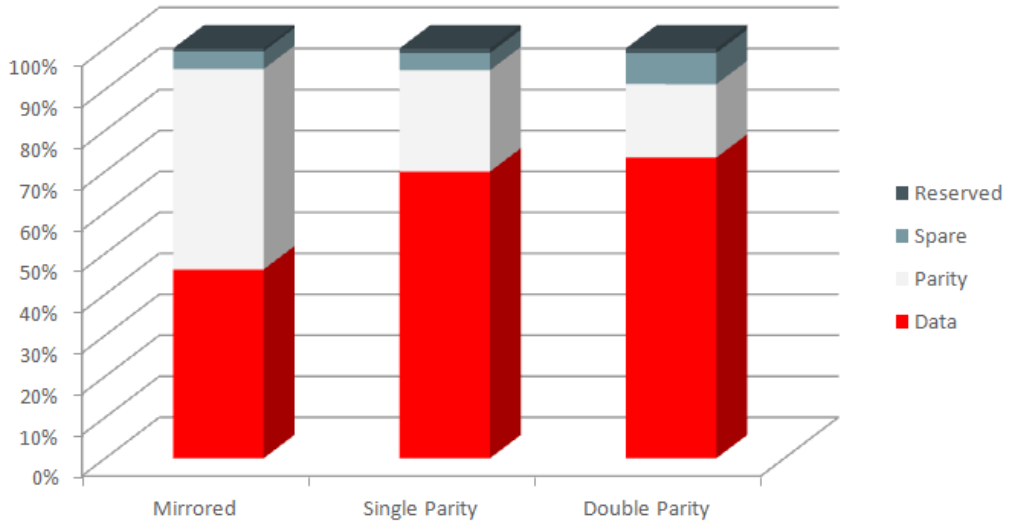

Figure 3. Raw disk capacity distribution

Configuring the Storage Pools

In most situations, Oracle recommends configuring a single storage pool on each controller. Each storage pool should be configured with half of the available hard disk drives (HDDs) in each disk shelf. This allows for maximum performance and redundancy.

Oracle recommends selecting the no single point of failure (NSPF) option when configuring the storage pool. This will ensure that the loss of an entire disk shelf will not compromise the availability of data.

Using Write Flash Accelerators and Read-Optimized Flash

The Oracle ZFS Storage Appliance provides a unique, cost-effective, high-performance storage architecture that is built on a flash-first Hybrid Storage Pool model. A performance on demand approach allows optional write-flash accelerators and read-optimized flash devices to be configured into the storage pool.

Utilizing flash-based caching to unlock the power of Oracle's Hybrid Storage Pool architecture is critical to achieving optimal performance with transactional or mixed I/O workloads. If an incrementally updated backup strategy is implemented to the pools, then write flash devices are recommended.

Using Oracle Direct NFS

Oracle Direct NFS (dNFS) is highly recommended for all database workloads between the Oracle Private Cloud Appliance and the Oracle ZFS Storage Appliance. It is required to achieve optimal performance.

dNFS is a custom NFS client that resides within the database kernel and provides several key advantages:

- Significantly reduces system CPU utilization by bypassing the operating system (OS) and caching data just once in user space with no second copy in kernel space
- Boosts parallel I/O performance by opening an individual transmission control protocol (TCP) connection for each database process
- Distributes throughput across multiple network interfaces by alternating buffers to multiple IP addresses in a round robin fashion
- Provides high availability (HA) by automatically redirecting failed I/O to an alternate address

These features enable increased bandwidth and reduced CPU overhead.

A best practice is to increase the maximum number of NFS server threads on the Oracle ZFS Storage Appliance from the default of 500. To do this, access the Oracle ZFS Storage Appliance browser user interface (BUI), select **Configuration→Services→NFS**, and set the number of threads to 1000.

Oracle Intelligent Storage Protocol

Oracle Intelligent Storage Protocol (OISP) was introduced with dNFS in the 12c version of Oracle Database. OISP requires NFSv4 support, which is native to the Oracle ZFS Storage Appliance. Also,

OISP enables database-aware storage by dynamically tuning record size and synchronous write bias on the Oracle ZFS Storage Appliance, which simplifies the configuration process and reduces the performance impact due to configuration errors. Hints are passed from the Oracle Database kernel to the Oracle ZFS Storage Appliance. These hints are interpreted to construct a workload profile that is used to dynamically optimize storage settings.

OISP is an optional enhancement. Reference My Oracle Support (MOS) Document 1943618.1 for instructions on how to enable OISP.

Configuring the Oracle ZFS Storage Appliance Networking

The Oracle Private Cloud Appliance provides optical 10GbE connections to external storage through the public ports on the two Oracle Fabric Interconnect F1-15 switches in the rack. Configuration of these ports is described in the *Oracle Private Cloud Appliance Administrator's Guide*.

Note: This white paper addresses accessing an external Oracle ZFS Storage Appliance using the Fabric Interconnect F1-15 switches, and does not address the internal rack-mounted Oracle ZFS Storage ZS3-ES that is standard in every Oracle Private Cloud Appliance and which is primarily used for Oracle VM infrastructure.

Each Oracle ZFS Storage Appliance external ZFS controller head has four 10GBase-T (RJ45 copper) 10GbE ports as standard. If optical 10GbE connectivity is required to the Oracle ZFS Storage Appliance, then optional optical 10GbE cards are required in each Oracle ZFS Storage Appliance head. RJ45 small form factor pluggable (SFP) interfaces are also available in the Fabric Interconnect switches.

For network configuration, both the IP network multipathing protocol (IPMP) and Link Aggregation Control Protocol (LACP) IEEE 802.3ad bonding mode are recommended for deployment of the Oracle Private Cloud Appliance with Oracle ZFS Storage Appliance. IPMP provides physical interface failure detection, increased reliability, and improved network performance for systems with multiple interfaces. LACP IEEE 802.3ad provides a method that combines the capacity of multiple full-duplex Ethernet links into a single logical link, which increases bandwidth by aggregating existing datalinks, load balance of inbound and outbound network traffic, and also automatic failover/failback traffic from a failed link over to working links in the aggregation. These two technologies complement each other, and can be deployed together to provide benefits for network performance, bandwidth, and high availability for Oracle Private Cloud Appliance environments with Oracle ZFS Storage Appliance.

For the LACP configuration, ensure that the Oracle ZFS Storage Appliance 10GbE network interfaces and IP data center switches are correctly configured with the following LACP policies:

- LACP policy L3 – Oracle ZFS Storage Appliance and IP network switches
- LACP active mode – Oracle ZFS Storage Appliance and IP network switches
- LACP short timer interval – Oracle ZFS Storage Appliance and IP network switches

For all 10GbE network interfaces used by the Oracle Private Cloud Appliance environment, ensure that 9000 maximum transmission unit (MTU) jumbo frames are enabled on the Oracle ZFS Storage Appliance 10GbE network interfaces and data center IP switches.

To configure IPMP, click on one of the devices to be connected to one or more other devices through LACP. Drag and drop the device to the Datalinks column. A dialog box entitled "Network Datalink" will open. Give the datalink a name, set the MTU size, Link Speed and Link Duplex properties as necessary for the physical link. Then check the box titled "LACP Aggregation". From here, you can click set LACP attributes of the bundle to L3, Active, and Short as previously listed, and click the checkboxes on all of the devices to be bundled together by LACP. With all these devices selected, click APPLY.

Figure 4 presents a sample of an IPMP network configuration recommended for Oracle ZFS Storage Appliance (head 1) connected to an Oracle Private Cloud Appliance environment.

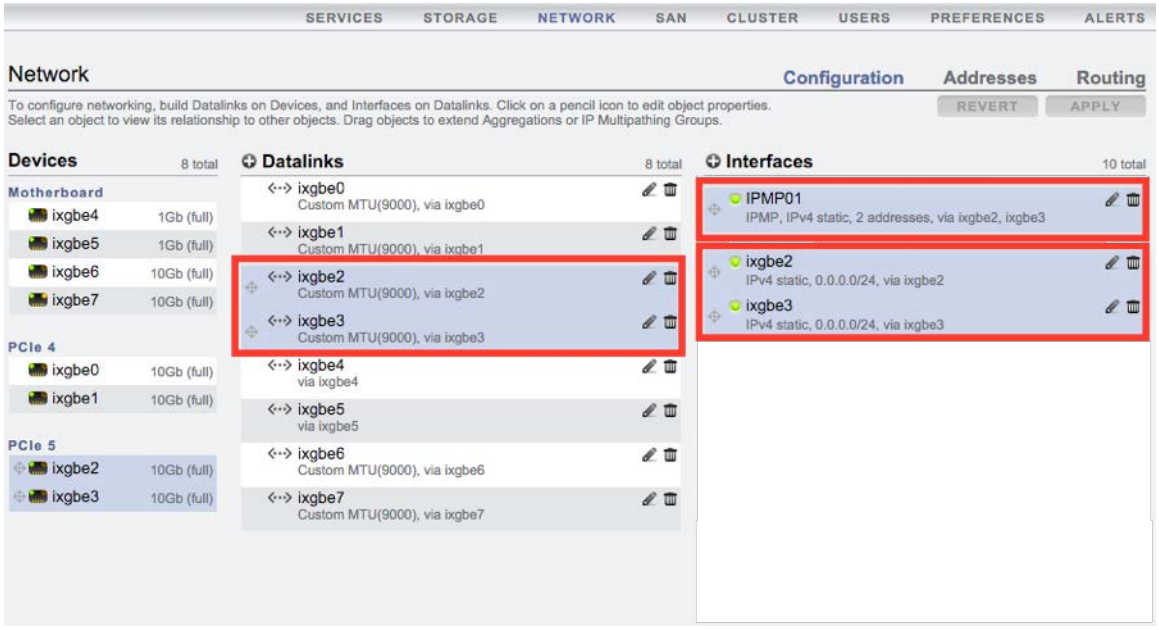

Figure 4. Example of IPMP network configuration on Oracle ZFS Storage Appliance cluster head 1 for Oracle VM 3

To have at least two 10GbE network interfaces as part of the IPMP configuration in active/active mode, or bundled into a single 802.3ad LACP channel, will require at least two10GbE network interfaces per Oracle ZFS Storage Appliance head.

Drag the datalinks that you created, and which you want to gather into an IPMP group, into the "Interfaces" column. Give the interfaces a name and an IP address of 0.0.0.0/24. Once you have created the individual interfaces with 0.0.0.0/24 IP addresses, click the (+) plus sign next to the "Interfaces" label at the top of the column to create the Network Interface that will be bound together with IPMP. Give the interface a name, an IP address/netmask, and click the checkbox with the label "IP MultiPathing Group" as seen in the following figure. From there you can select the interfaces you just created.

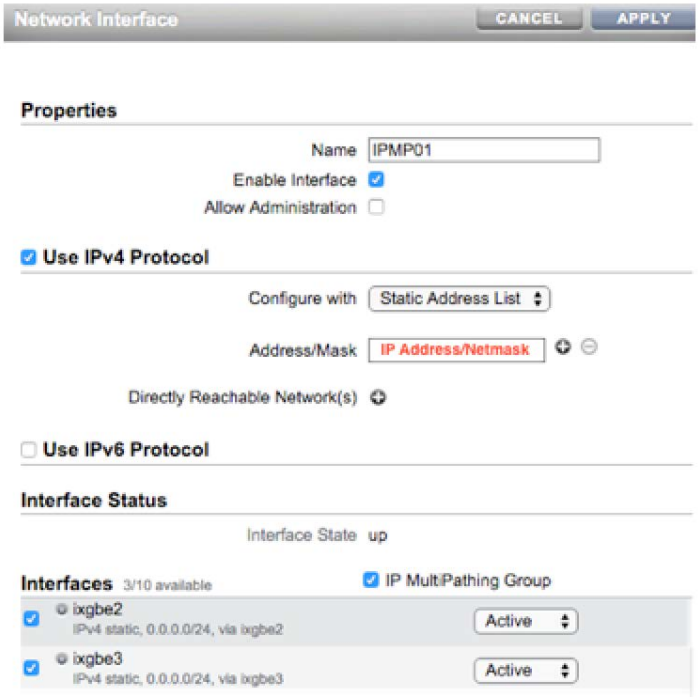

Figure 5. IPMP network configuration on Oracle ZFS Storage Appliance cluster head 2 for Oracle Private Cloud Appliance

Figure 6 shows a sample recommended IPMP network configuration for Oracle ZFS Storage Appliance (head 2) within an Oracle VM 3 environment.

Figure 6. Example IPMP network configuration for Oracle ZFS Storage Appliance cluster head 2 and Oracle VM 3

You now will have physical network interfaces that benefit from both LACP load balancing and IPMP failover protection. For more details, refer to the Oracle white paper titled "Networking Best Practices with the Oracle ZFS Storage Appliance" at: http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/networking-bestprac-zfssa-2215767.pdf

Configuring the SNMP Service on the Oracle ZFS Storage Appliance

System Network Management Protocol (SNMP) is used to allow communication between the Oracle Database software and the Oracle ZFS Storage Appliance hardware. This communication is used to determine eligibility for features such as dNFS and OISP.

To enable SNMP, follow these steps:

1. Navigate **to Configuration →Services** and click on **SNMP** near the bottom.
2. Click on **SNMP** underneath System Settings (near the bottom):
3. Use following settings:
   - Version: `v1/2c`
   - Community name: `public`
   - Authorized network/mask: `0.0.0.0 / 0`
   - Appliance contact: `Any text string`
   - Trap destinations: `127.0.0.1`

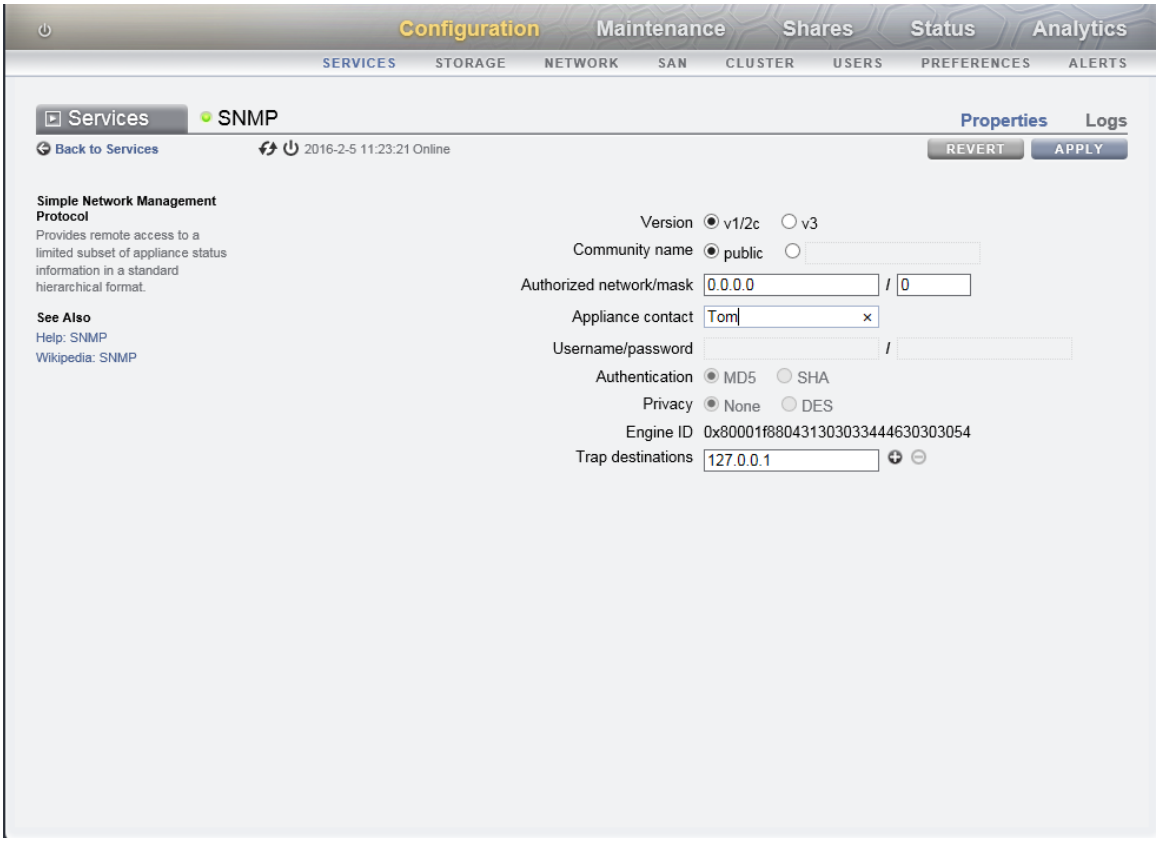

Figure 7. Configuring SNMP on the Oracle ZFS Storage Appliance BUI

Verify the SNMP configuration using the `snmpget` command:

```
snmpget -v1 -c public <IP address of ZFS Storage Appliance data path>
1.3.6.1.4.1.42.2.225.1.4.2.0
```

```
SNMPv2-SMI::enterprises.42.2.225.1.4.2.0 = STRING: "Sun ZFS Storage 7420"
```

Creating Storage Projects and Shares on the Oracle ZFS Storage Appliance

A new project for the Oracle Private Cloud Appliance should be created in the Oracle ZFS Storage Appliance BUI, as seen in the following figure. It is assumed that at least one storage pool has already been configured on the controller that will host the shares. As seen in the highlighted text labels, click **Shares**, **Projects**, and then click the plus icon (**+**). Enter the name of the project.

NOTE: The example shown uses OracleVM3 as the project name for the Oracle Private Cloud Appliance environment.

Figure 8. Configuring a new project in the Oracle ZFS Storage Appliance BUI

As a best practice for better performance and high availability of the Oracle ZFS Storage Appliance NFS shares presented to the Oracle VM environment, split the NFS shares across different cluster heads and different Oracle ZFS Storage Appliance storage pools in order to spread the load effectively. The recommended storage layout of Oracle ZFS Storage Appliance for Oracle Database shares under Oracle VM is:

- Set up three shares per Oracle Database 12c database: one for data files, one for redo logs and one for FRA (if used). The ZFS Database record size for each share should be set to 128K, and Synchronous write bias to Latency. With Oracle Database 12c, Oracle Intelligent Storage Protocol (OISP) will work in conjunction with the Oracle ZFS Storage Appliance to dynamically select the optimal Database record size, Synchronous write bias, and other attributes.
- Configure Direct NFS (dNFS) for the database, and specify NFSv4 as the access protocol.
- Configure SNMP so that OISP is able to automatically adjust record sizes and throughput/latency attributes.
- Balance NFS shares across both heads of the Oracle ZFS Storage Appliance cluster.
- For critical databases and applications which require low latency and high performance, use Oracle ZFS Storage Appliance performance disks – 10k RPM disks as opposed to capacity 7500 RPM disks.
- For best performance, ensure that 9000 Maximum Transmission Unit (MTU) has been correctly enabled.
- Do not enable data compression or encryption.
- Ensure that NFS exceptions are correctly configured, with access granted only to the Oracle VM servers.
- Disable services such as: SMB, HTTP, FTP, SFTP, and TFTP for file systems. Only NFS service should be enabled as read/write share mode.
- Disable file system properties such as: read only, update access time on read, non-blocking mandatory locking, data deduplication, and virus scan.
- Set the check sum property to Fletcher4 (standard).
- Set cache device usage set to all data and metadata.
- Set additional replication to normal.

- Enable file system properties such as prevent destruction and restrict ownership change.

Figure 9 presents the initial configuration of an Oracle ZFS Storage Appliance network file system – including project name, file system name, permission, and standard configurations such as normalization, encryption, and mount point – which will later be used as a cluster file system for the Oracle VM 3 environment.

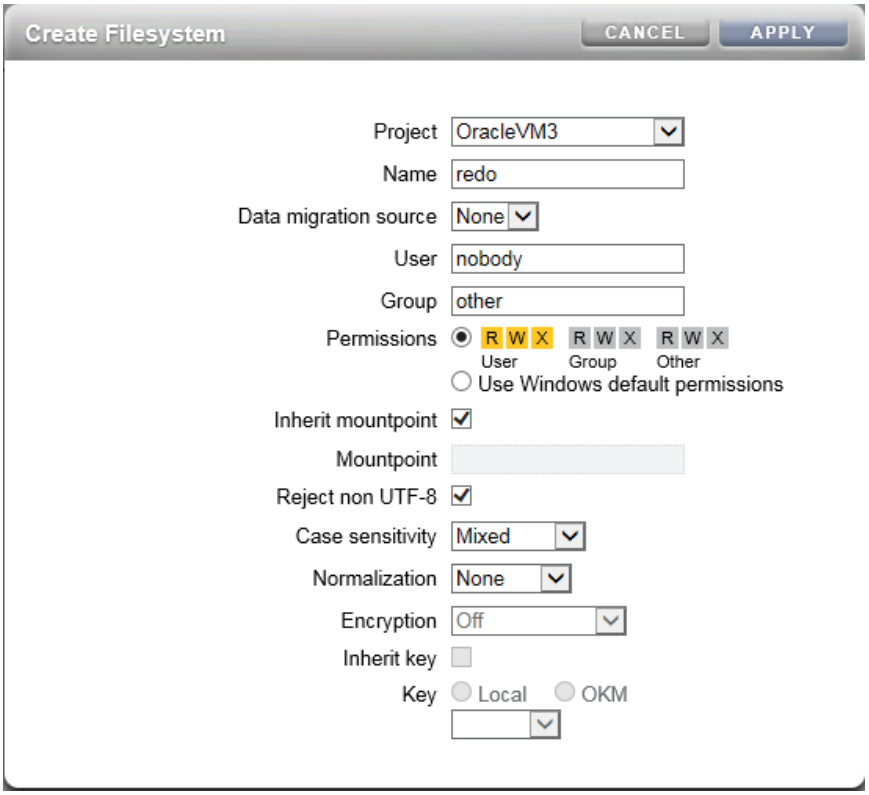

Figure 9. Configuring the network file system in the Oracle ZFS Storage Appliance BUI

Figure 10 shows the subsequent configuration of the Oracle VM 3 environment (project named `oracleVM3`).

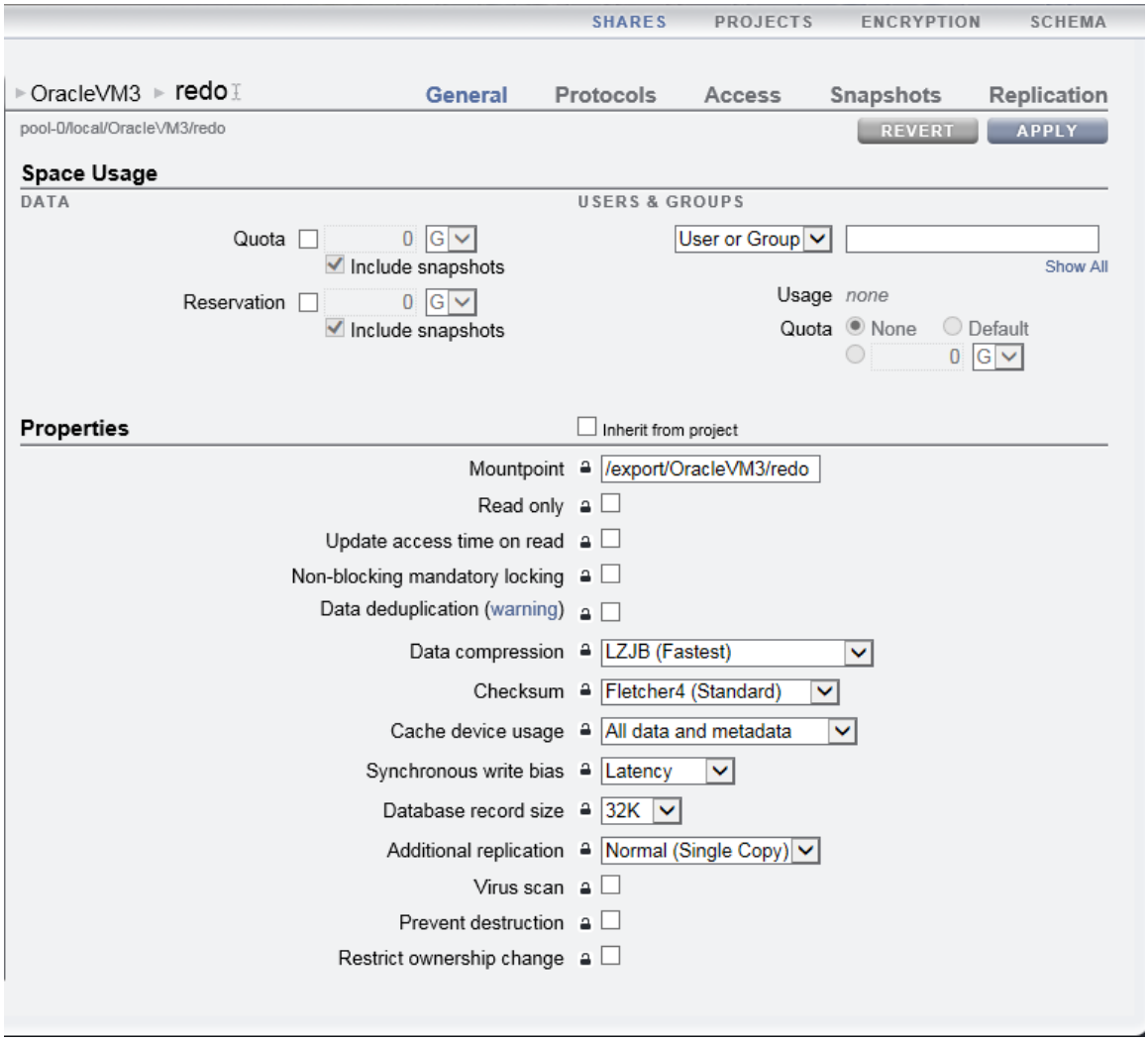

Figure 10. Configuring Oracle VM3 in the Oracle ZFS Storage Appliance BUI

Using the DeployCluster Tool for Oracle VM Templates for Oracle Database to Deploy a Virtual Machine

Initial installation and configuration of the Oracle Private Cloud Appliance itself is beyond the scope of this document. It is assumed that you have an installed and running Oracle Private Cloud Appliance environment with connections to the external, non-management data center network through which an external Oracle ZFS Storage Appliance can be accessed.

The operating system storage used for VM Guests in the Oracle Private Cloud Appliance is usually hosted from an Oracle VM Repository residing on the internal Oracle ZFS Storage ZS3-ES. While you can use the internal repository for database files, the internal ZS3-ES has limited scalability. This paper illustrates running the OS and Oracle Database binaries from an internal repository and hosting the database files on an externally attached Oracle ZFS Storage Appliance.

Oracle VM Templates are a fast and easy way of provisioning a fully featured and fully configured Oracle Database system with features like DTrace and Ksplice already installed. Templates are prebuilt images containing the OS, Grid Infrastructure, Database, and other binaries that can be easily loaded and deployed into virtual machines on the Oracle Private Cloud Appliance. The DeployCluster tool is the preferred method for deploying Oracle VM Templates beginning with Oracle VM 3. DeployCluster can be used for either a single instance or an Oracle Real Application Cluster (RAC) deployment. To deploy a pre-built template using the DeployCluster tool, access the Oracle VM Templates page on the Oracle Technology Network and select "Oracle VM Templates for Oracle Database". A list of available templates will appear along with install and configuration documents for both test and production deployments. The install instructions for the selected template will guide you through the deployment of a fully configured Linux VM running the desired version of the Oracle Database.

The DeployCluster tool documentation lists the steps necessary to deploy RAC, Single Instance/HA, and Single Instance VMs. In RAC and HA environments, ASM is used, but in this paper's Single Instance non-H/A example, we will be using NFS filesystems on the Oracle ZFS Storage Appliance to contain the database files.

This white paper is using the following Oracle VM Template obtained from: "Single Instance and Oracle RAC 12c Release 1 Enterprise Edition, including Oracle Grid Infrastructure Patch Set one Update #5 (12.1.0.2.5) and Oracle Linux 7 Update 1 - Available from My Oracle Support under Patch number 18888877 for 64-Bit Linux."

For further details on executing Steps 1-7, refer to the DeployCluster.py documentation for the specific template you are using.

1. Import the desired DB/RAC Template.
2. Adjust the template's network bridges.
3. Clone VM(s) from the DB/RAC Template.
4. Based on deployment mode, attach the (shared) disks to all VMs as needed.
5. Download the DeployCluster tool.
6. Create a `netconfig.ini` file.

   Your `netconfig.ini` file must contain at least one network interface with access to a network on which the external Oracle ZFS Storage Appliance can be reached.

7. Run the `deploycluster.py` tool.

   a. After running `deploycluster.py`, you should have an active Virtual Machine running a Linux OS with a preconfigured Oracle Database deployment. By default, a small sample database is running on the local repository storage, and an oracle user is created as well as appropriate groups.

   b. Additional configuration may be necessary to configure a tty connection to an X-Windows server or to configure the vncservers service to provide virtual network computing (VNC) support. A windowing environment is necessary to be able to run the Oracle Database Configuration Assistant (dbca) in a graphical environment.

c. For production environments, Oracle recommends to run the Ksplice tool to determine whether your kernel has vulnerabilities. Ksplice is an optional tool that examines the Linux kernel level and compares it to a database of known issues. It then suggests kernel patches to resolve the issues, and often these kernel patches can be applied while the operating system is active, avoiding reboots. You can visit www.ksplice.com/inspector , or execute this command:

```
[root@myguest ~]# (uname -s; uname -m; uname -r; uname -v) | \

 curl https://uptrack.api.ksplice.com/api/1/update-list/ \

 -L -H "Accept: text/text" --data-binary @-

Your kernel needs the following updates:
CVE-2015-5307: KVM host denial-of-service …….
…
…..Deadlock when syncing frozen filesystem.
```

8. Mount the NFS shares that reside on the Oracle ZFS Storage Appliance to the running Linux VM.
9. Configure the System and Oracle Database NFS shares by editing `/etc/fstab` and creating an `oranfstab` file, and testing SNMP functionality.
10. Create an Oracle Database on the NFS shares residing on the external Oracle ZFS Storage Appliance.

Configuring the Oracle Linux Virtual Machine and Oracle Database Software

Use the following directions to configure the Oracle Linux Virtual Machine and Oracle Database software.

Configuring NFSv4

NFSv4 implements several enhancements, such as a stronger security model, file locking managed within the core protocol, and delegations which can help improve the accuracy of client side caching. NFSv4 is required, along with a 12c version of Oracle Database, to enable Oracle Intelligent Storage Protocol, which allows dynamic storage tuning by the Oracle Database software.

NFS Mount Options

Add the Oracle Database shares to `/etc/fstab`, utilizing the following mount options:

```
aie-7320d-h1:/export/OracleVM3/datafiles /mnt/datafiles  nfs  \
rw,bg,hard,nointr,tcp,nfsvers=3,timeo=600,rsize=1048576,wsize=1048576,actimeo=0

aie-7320d-h1:/export/OracleVM3/redo      /mnt/redo       nfs  \
rw,bg,hard,nointr,tcp,nfsvers=3,timeo=600,rsize=1048576,wsize=1048576,actimeo=0

aie-7320d-h1:/export/OracleVM3/fra       /mnt/fra        nfs  \
rw,bg,hard,nointr,tcp,nfsvers=3,timeo=600,rsize=1048576,wsize=1048576,actimeo=0
```

NOTE: NFSv4 will be used by Oracle Database to access the shares, by virtue of NFSv4 being specified in the `oranfstab` file. The `fstab` should specify `nfsvers=3` as shown above.

Create mountpoint directories on the Linux system, mount the shares that you created on the external Oracle ZFS Storage Appliance, and modify the permissions and share ownership.

```
# mkdir /mnt/datafiles
# mkdir /mnt/redo
# mkdir /mnt/fra
# mount /mnt/datafiles
# mount /mnt/redo
# mount /mnt/fra
# chmod 755 /mnt/datafiles
# chmod 755 /mnt/redo
# chmod 755 /mnt/fra
# chown oracle:oinstall /mnt/datafiles/
# chown oracle:oinstall /mnt/redo
# chown oracle:oinstall /mnt/fra
#
```

Figure 11. Creating mountpoint directories on the Linux system

Creating an Oracle Database on NFS Shares

The Oracle VM Template used to load the Virtual Machine already contains the Oracle Database binaries in the usual directory structure under the `/u01` directory on storage served by the Oracle Private Cloud Appliance Oracle VM repository. An oracle user and oinstall group are preconfigured. Once the Oracle ZFS Storage Appliance shares are created and mounted, a database can be created which resides on those shares using the Database Configuration Assistant (`dbca`) tool. Follow these steps

1. Log into the running virtual machine using the oracle user. The default `oraenv` executed out of the profile will set up an executable PATH and ORACLE_HOME.
2. An X-Windows environment is necessary to execute `dbca`. Either set the DISPLAY variable to your X server environment on your workstation, or log in to a vncserver session that has an X-Windows environment.
3. Run the `dbca` command and select Advanced Mode in the Creation Mode window, as seen in the following figure.

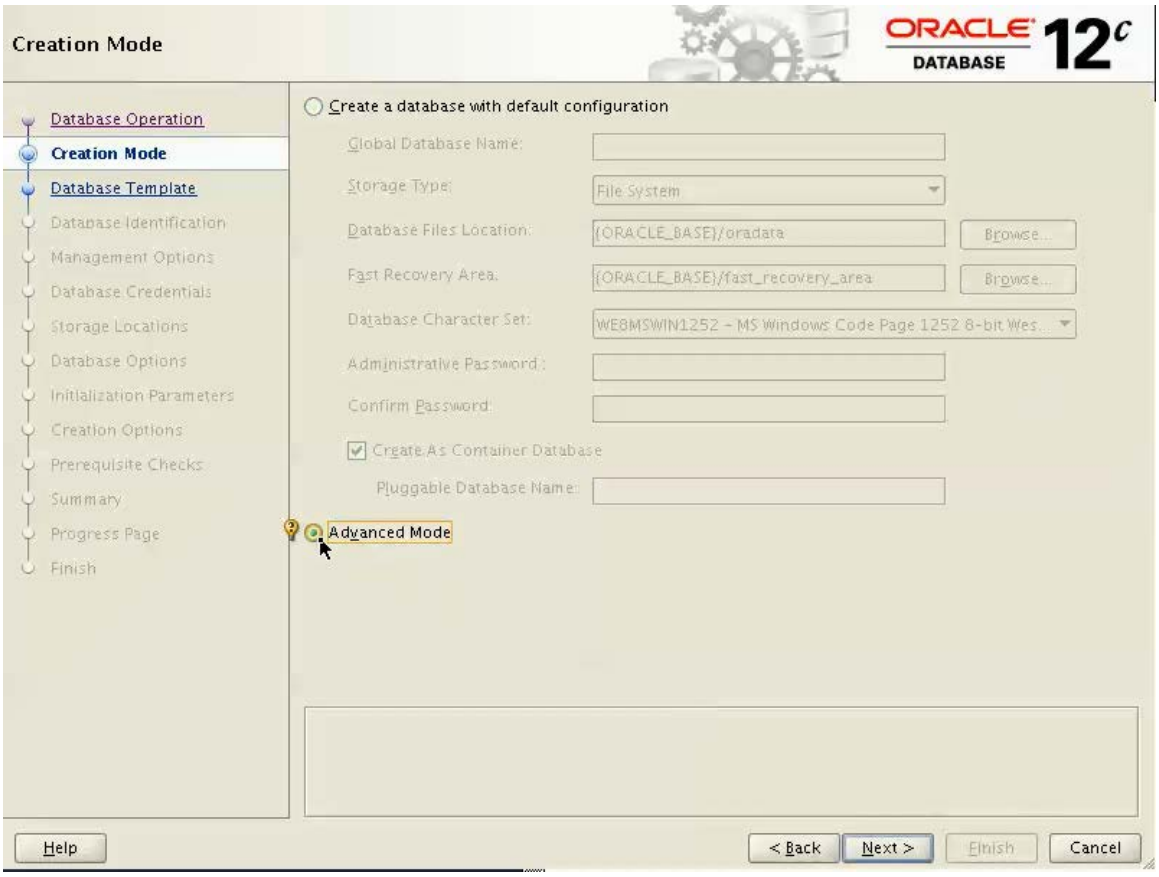

Figure 12. Configuring a database with the dbca tool in the virtual machine

4. Continue through the configuration screens until the "**Storage Locations**" screen appears. Select "**Use Common Locations for All Database Files**".

5. Specify `/mnt/datafiles` for File Location, `/mnt/fra` for the Fast Recovery Area and Archive Log locations. Click "**Next**" and continue through the configuration screens.

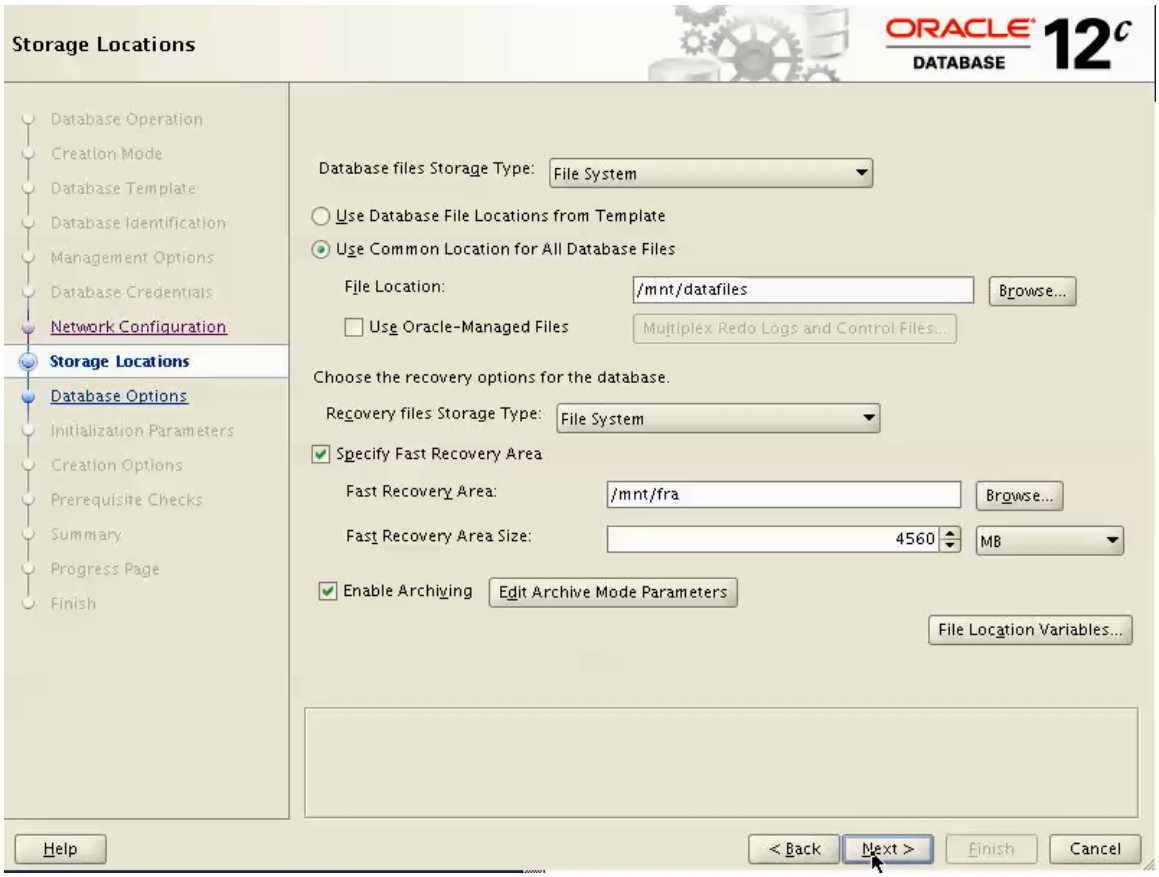

Figure 13. Specifying storage location details in the virtual machine for the database instance

6. On the "**Creation Options**" screen, click "Customize Storage Locations." and edit each redo log member to place it on the `/mnt/redo` share.

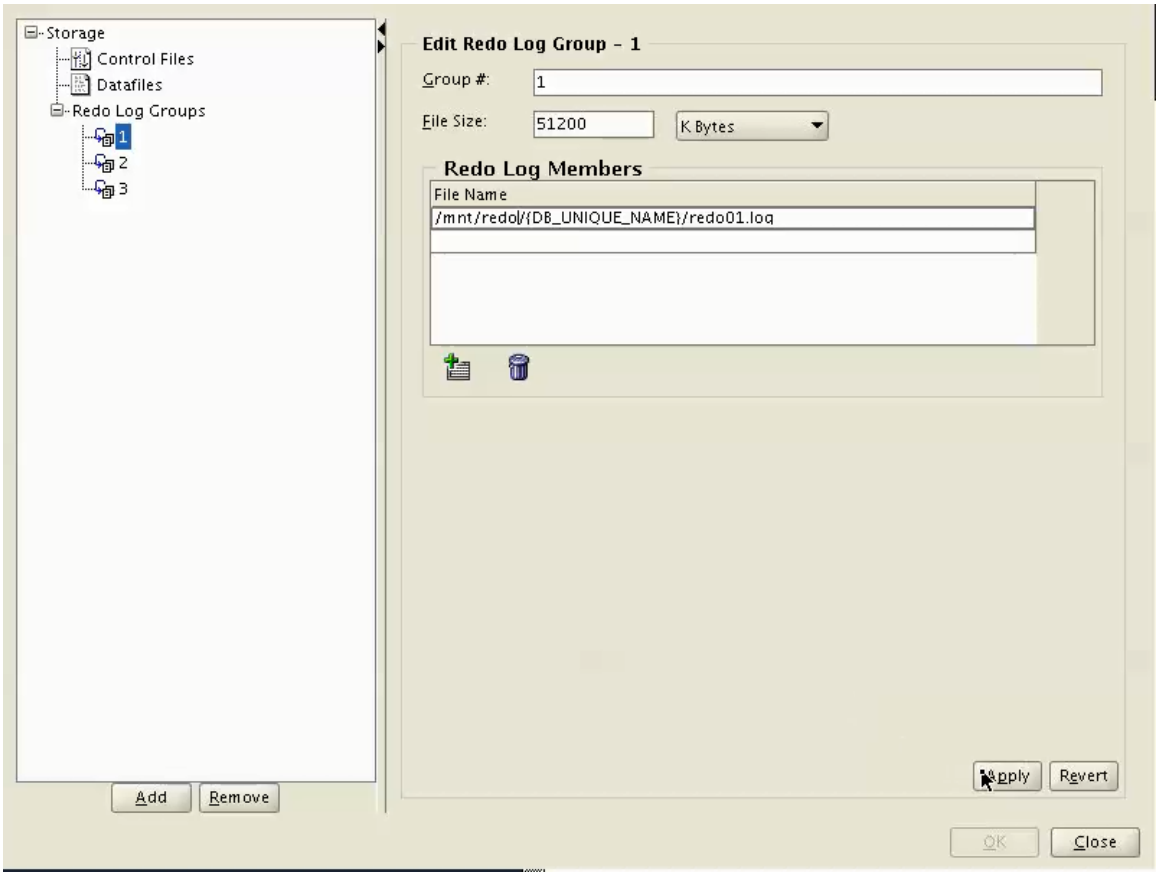

Figure 14. Editing redo log groups in the virtual machine running the dbca tool

7. Finish specifying options, and then create the database by clicking "**Finish**".

Configuring Oracle Direct NFS

By default, Direct NFS Client is installed in a disabled state in single instance Oracle Database installations. To enable the Direct NFS Client, complete the following steps:

1. Change the directory to `$ORACLE_HOME/rdbms/lib`.
2. Enter the following command:

   ```
   make -f ins_rdbms.mk dnfs_on
   ```

dNFS may be disabled with the following command:

```
$ make -f $ORACLE_HOME/rdbms/lib/ins_rdbms.mk dnfs_off
```

Creating an `oranfstab` File

The `oranfstab` file is required to control the use of dNFS with an Oracle Database. The file is created in `$ORACLE_HOME/dbs/oranfstab` and applies to all database instances that share the Oracle home.

The `oranfstab` file allows load spreading of dNFS connections over multiple addresses on the Oracle ZFS Storage Appliance (represented by "path") and/or multiple addresses on the Virtual Machine (represented by "local"). In a production environment, load balancing over multiple interfaces will reduce or eliminate two possible system bottlenecks: network interface bandwidth and TCP/IP buffering.

Specifying NFSV4 in the `oranfstab` is recommended to allow OISP to function.

Here is an example of a simple one-path `oranfstab`:

```
server: abc-zfssa-h1

path: 192.168.44.112

nfs_version: nfsv4

export: /export/OracleVM3/datafiles mount: /mnt/datafiles

export: /export/OracleVM3/fra mount: /mnt/fra

export: /export/OracleVM3/redo mount: /mnt/redo
```

The database must be shut down and restarted in order to activate dNFS after the `oranfstab` is configured. Confirm that dNFS is enabled by checking the database alert log for an Oracle ODM message after database startup:

```
Oracle instance running with ODM: Oracle Direct NFS ODM Library Version 3.0
```

dNFS activity can also be confirmed by connecting to the database and using SQL queries:

```
SQL> select * from v$dnfs_servers;
SQL> select * from v$dnfs_files;
SQL> select * from v$dnfs_stats;
SQL> select * from v$dnfs_channels;
```

For more information about dNFS, refer to the *Oracle Database Installation Guide for Linux*. at http://docs.oracle.com/database/121/LADBI/toc.htm.

Conclusion

The Oracle ZFS Storage Appliance and the Oracle Private Cloud Appliance combine to provide an excellent solution for Oracle Databases or any other virtualized applications. When the total cost of ownership, flexibility, scalability and performance are considered, the Oracle ZFS Storage Appliance is an obvious choice for the Oracle Private Cloud Appliance expansion storage.

References

See the following resources for additional information relating to the products covered in this document.

- Oracle ZFS Storage Appliance White Papers and Subject-Specific Resources
  http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/index.html
  including:
  - " Expanding Oracle Private Cloud Appliance Using ZFS Storage Appliance "http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/expand-opca-using-ozfssa-2613723.pdf
  - "Networking Best Practices with the Oracle ZFS Storage Appliance" http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/networking-bestprac-zfssa-2215767.pdf
- Oracle ZFS Storage Appliance Product Information
  https://www.oracle.com/storage/nas/index.html
- Oracle ZFS Storage Appliance Documentation Library, including Installation, Analytics, Customer Service, and Administration guides:
  http://www.oracle.com/technetwork/documentation/oracle-unified-ss-193371.html
- The *Oracle ZFS Storage Appliance Administration Guide* is also available through the Oracle ZFS Storage Appliance help context.
  The Help function in Oracle ZFS Storage Appliance can be accessed through the browser user interface.
- "Oracle VM Windows Paravirtual (PD) Drivers" – information from Oracle Technology Network's Oracle Virtualization information site
  http://www.oracle.com/us/technologies/virtualization/virtualilzation-066470.html

ORACLE

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Oracle ZFS Storage Appliance with Oracle Private Cloud Appliance: Configuration Best Practices for Supporting Oracle Database
April 2016
Author: Scott Ledbetter