Exadata X8M and ZFS Storage 25Gb Backup Solution

Configuration of Oracle Exadata X8M with Oracle ZFS Storage Appliance ZS7-2 using dedicated backup switches

May 2021 | Version 2.0

Public

DISCLAIMER

Table of Contents

EXECUTIVE OVERVIEW

Protecting the mission-critical data that resides on Oracle Exadata Database Machine (Oracle Exadata) is a top priority. The Oracle ZFS Storage Appliance is ideally suited for this task due to superior performance, enhanced reliability, extreme network bandwidth, powerful features, simplified management, and cost-efficient configurations.

The Oracle Exadata X8M product line introduces a new and improved networking infrastructure built on high-speed, low latency network fabric optimized for RDMA over Converged Ethernet (RoCE). This paper focuses on network attachment and configuration best practices when using an Oracle ZS7 Storage Appliance with its optional Top-of-Rack (ToR) switches to provide data protection services for a RoCE based Oracle Exadata Database Machine. More comprehensive general-purpose guidelines for using ZFS Storage in an Exadata environment are provided in My Oracle Support document 2087231.1 "Guidelines when Using ZFS Storage in an Exadata Environment."

There are two options when choosing how to back up an Exadata X8M to an Oracle ZFS Storage Appliance: use a ToR switch based backup network or directly connect to the X8M's internal 100Gb RoCE network fabric.

» ZS7-2 ToR switches connected to 25GbE backup network ports on the Exadata X8M database nodes
  - Provides a cost efficient, flexible backup solution with great performance
  - Provides an isolated backup network allowing the Oracle ZFS Storage Appliance to be a shared backup target for multiple Exadata and other Oracle systems in larger configurations
  - Two dual-port 25GbE cards per ZS7 controller are connected to the ToR switches
  - Two 25GbE ports on each Exadata database server are connected to the ToR switches
  - LACP (Link Aggregation Control Protocol) and jumbo frames are used to increase network throughput and provide high availability

» ZS7-2 direct connection to the Exadata X8M RoCE network fabric over 100 GbE
  - ZS7-2 is the only storage platform to offer a 100Gb backup solution
  - Dual-port 100GbE adapters in each ZS7-2 controller are directly connected to the Oracle Exadata X8M internal leaf switches
  - Bypasses the traditional backup network and provides a local backup option that is optimized for performance and simplicity
  - Eliminates the need to acquire and manage dedicated switches
  - RoCE interfaces use active path management along with IPMP (IP network Multi-Pathing) on the ZS7 to provide high availability
  - This solution can serve as a 100Gb backup target for a single-rack or multi-rack X8M
  - For details about configuring this solution, see the Oracle solution brief 100Gb Backup Solution

Using the ZS7-2 ToR switches has the benefits of isolating the backup network and providing a greater level of control. This solution provides more flexibility for a single Oracle ZS7 Storage Appliance to serve as a shared backup target for multiple Exadata environments that are configured with isolated RoCE networks. These dedicated switches can support higher levels of backup throughput for large configurations due to the ability to use jumbo frames. Top of rack switches are not managed as part of the Oracle Exadata HW stack and will require occasional software upgrades.

The following graph shows maximum sustainable backup and restore performance for a X8M full rack and ZS7-2 ToR switch configuration:

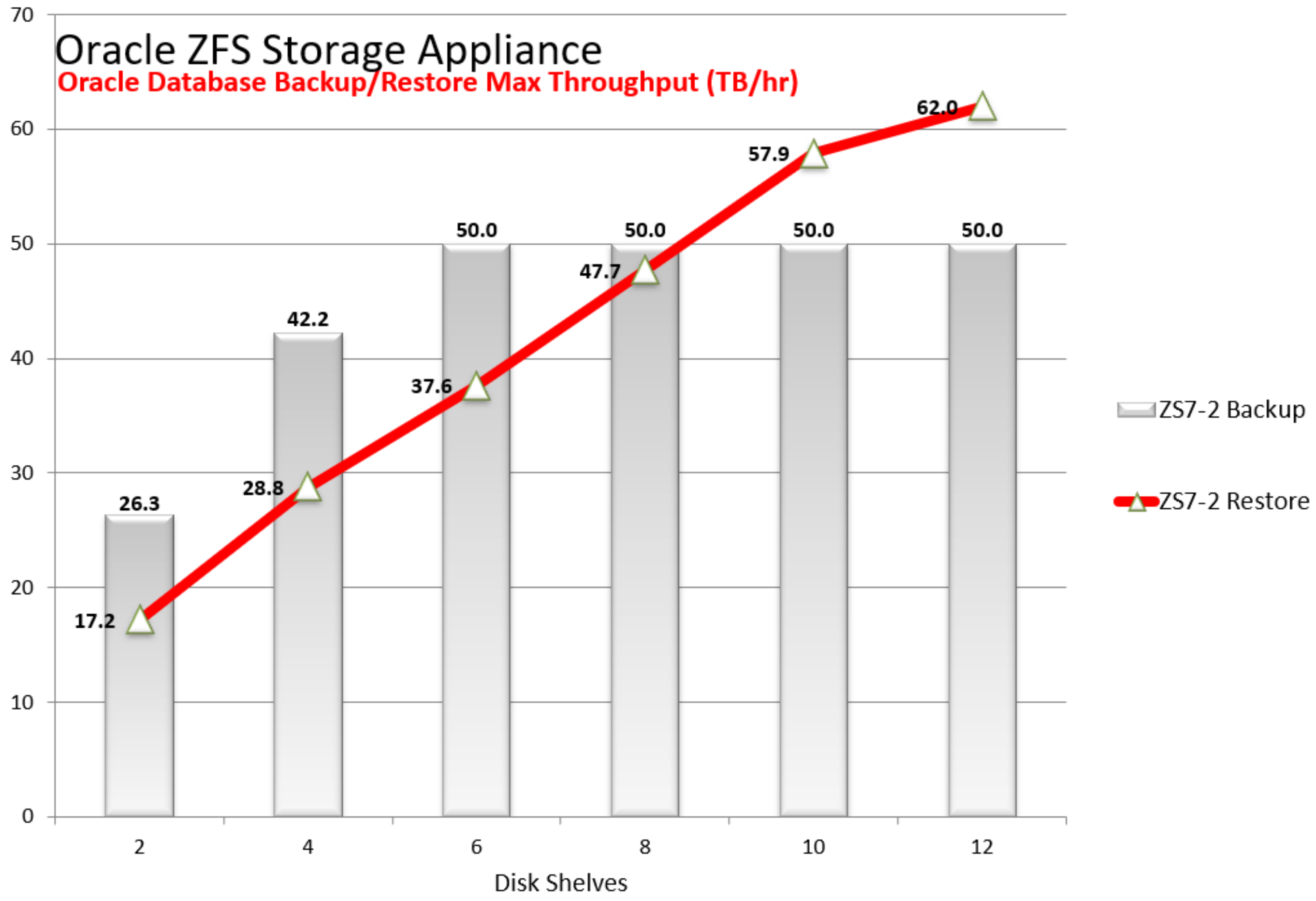


Figure 1. Maximum sustainable backup and restore throughput of Oracle ZFS Storage ZS7-2 with ToR switches when connected to Oracle Exadata X8M full rack

The rates depicted in the above chart represent the maximum sustainable backup and restore throughput that Oracle ZFS Storage Appliance ZS7-2 configurations ranging from 2 to 12 disk shelves are capable of supporting when connected to a full rack Exadata X8M. For general-purpose performance characterization of RMAN workloads across a complete range of ZS7-2 configurations, please reference My Oracle Support document 2087231.1 "Guidelines when Using ZFS Storage in an Exadata Environment."

These are complete, real-world results using Oracle Database 19c and an online transactional processing (OLTP) database that was populated with sample customer data in a sales order-entry schema. Advanced Row Compression, an Oracle Database feature, was used at the database level to align with best-practice recommendations for customers that are running OLTP workloads. These throughput rates were not obtained using a database or input/output (I/O) generator test tool, which can be misleading. Also, they were not projected based on low-level system benchmarks. The backup and restore performance data collected for this document was measured using level 0 backup and restore operations for an otherwise idle Oracle Database. When accounting for database-level compression or incremental backup strategies, effective backup rates that are much higher than the physical rates documented here are routinely observed.

The rates shown in figure 1 represent the maximum database backup/restore throughput that an Oracle ZS7-2 with one to six high-capacity disk shelves is capable of sustaining. Achieving these rates is contingent on having an Oracle Exadata X8M configuration large enough and with enough available resources to support this performance. Small configurations and configurations with a heavy concurrent workload will have difficulty achieving maximum throughput rates that the storage can support.

High performance is an important consideration when choosing a solution to protect Oracle Exadata. The following technologies make it possible for Oracle ZFS Storage Appliance systems to achieve these backup and restore rates:

- **Oracle Recovery Manager Integration** – Oracle Recovery Manager (Oracle RMAN) is a highly parallelized application that resides within Oracle Database and optimizes backup and recovery operations. Oracle ZFS Storage Appliance systems are designed to integrate with Oracle RMAN by utilizing up to 3,000 concurrent threads that distribute I/O across many channels spread across multiple controllers. This improves performance dramatically with sequential large block streaming I/O workloads that are typical for most backup and restore situations.
- **Oracle Database's Direct NFS Client feature** – The optimized Direct NFS Client feature is an aggressive implementation that allocates individual TCP connections for each Oracle Database process, in addition to reducing CPU and memory overhead, by bypassing the operating system and writing buffers directly to user space.
- **1 MB Record Sizes** – Oracle ZFS Storage Appliance systems enable larger 1 MB record sizes. This reduces the number of input/output operations per second (IOPS) that are required to disk, preserves the I/O size from Oracle RMAN buffers to storage, and improves performance of large-block sequential operations.
- **Hybrid Storage Pools** – Oracle ZFS Storage Appliance systems have an innovative Hybrid Storage Pool (HSP) architecture that utilizes dynamic storage tiers across memory, flash, and disk. The effective use of dynamic random-access memory (DRAM) and enterprise-class software specifically engineered for multilevel storage is a key component that facilitates the superior performance of Oracle ZFS Storage Appliance systems.

The performance benefits of the Oracle ZFS Storage Appliance are well documented and independently verified. Oracle periodically publishes Storage Performance Council's SPC-1 and SPC-2 benchmark results, as well as Standard Performance Evaluation Corporation's SPEC SFS benchmark results to demonstrate performance results for the Oracle ZFS Storage Appliance. Combine this with the powerful features, simplified management, and Oracle-on-Oracle integrations, and it is easy to see why these systems are a compelling solution for protecting mission-critical data on Oracle Exadata.

INTRODUCTION

Database, system, and storage administrators are faced with a common dilemma when it comes to backup and recovery of Oracle Database instances—how to back up more data, more often, in less time, and within the same budget. Moreover, practical challenges associated with real-world outages mandate that data protection systems be simple and reliable to ensure smooth operation under compromised conditions. The Oracle ZFS Storage Appliance helps administrators meet these challenges by providing cost-effective and high-bandwidth storage systems that combine the simplicity of the NFS protocol with ZFS-enhanced disk reliability. Through Oracle ZFS Storage Appliance technology, administrators can reduce the capital and operational costs associated with data protection while maintaining strict service-level agreements with end customers.

Oracle ZFS Storage Appliance systems are easy-to-deploy unified storage systems uniquely suited for protecting data contained in Oracle Exadata. With native 25 GbE, 40 GbE, and 100GbE connectivity, they are an ideal match for Oracle Exadata X8M. These high-bandwidth interconnects reduce backup and recovery time, as well as reduce backup application costs and support fees, compared to traditional NAS storage systems. With support for both traditional tiered and incrementally updated backup strategies, Oracle ZFS Storage Appliance systems deliver enhanced storage efficiency that can further reduce recovery time and simplify system administration.

Deploying Oracle ZFS Storage Appliance systems for protecting the mission-critical data that resides on Oracle Database on Oracle Exadata requires that backup window and recovery time objectives (RTOs) be met to ensure timely recovery in the event of a disaster.

This document focuses on networking guidelines for setting up an Oracle ZFS Storage Appliance ZS7-2 for optimal backup and recovery of Oracle Databases running on an Oracle Exadata X8M Database Machine. Selecting the right backup strategy, understanding encryption options and when to deploy a deduplication solution are all important considerations. Guidelines to help understand and make these deployment decisions are provided in the My Oracle Support document 2087231.1 "Guidelines when Using ZFS Storage in an Exadata Environment." This solution brief illustrates the networking guidelines by utilizing a reference architecture with a standard RMAN backup strategy with no encryption or deduplication requirements.

This paper addresses the following topics:

» Configuring the Network
» Configuring the Oracle ZFS Storage Appliance
» Configuring the Oracle Exadata X8M

CONFIGURING THE NETWORK

This section provides the details of deploying the backup network. It documents specific cabling guidelines and networking best practices to support a high performance and highly available data protection solution. The backup network is comprised of bondeth1 (25Gb) interfaces on the Exadata compute nodes and 25GbE data path interfaces on the ZS7-2 controllers.

Oracle ZS7-2 Top of Rack (ToR) Switches

The backup network is built using 25Gb Ethernet cards in the X8M database nodes and in the ZS7 controllers. The recommended configuration is two cards installed in each ZS7-2 controller.

25Gb connections use QSFP28 to 4x SFP28 breakout cables to form four 25Gb links on a 100Gb port. The “Configuring IO Cards” section of this document provides more details on the installation and setup of network cards on the storage.

Figure 2 depicts the network cabling diagram when the backup network is configured with ZS7-2 top of rack switches and 25GbE connections.

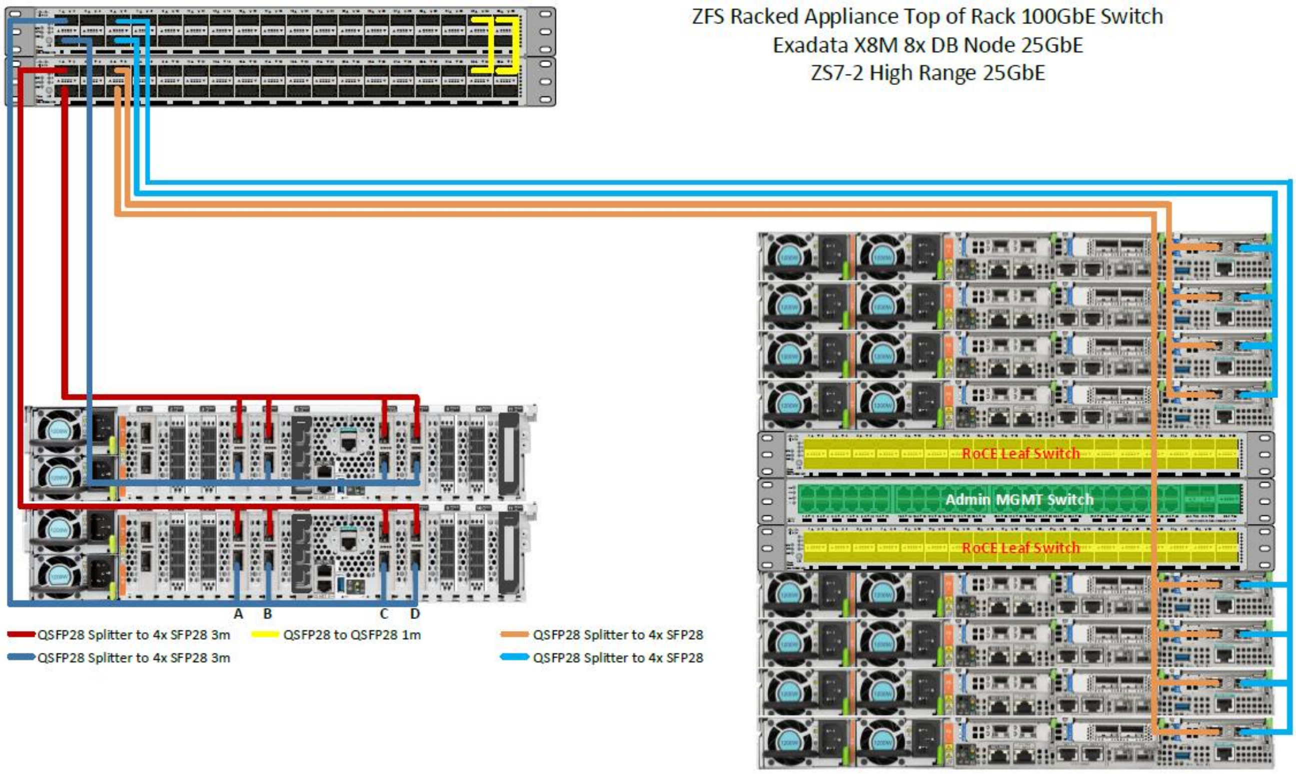


Figure 2. Network diagram with Oracle ZS7-2 top of rack switches

Switch Configuration Best Practices

The following section is a high-level view of the primary configuration components. For specific configuration details, see My Oracle Support Document 2635423.1 “Set Up and Configure Exadata X8M Backup with ZFS Storage ZS7-2.”

LACP and vPC Configuration

For optimal performance and availability, Link Aggregation Control Protocol (LACP) groups are implemented that spans both switches.

A port-channel configuration on the switch provides the basis to aggregate multiple interfaces together and then perform load balancing. However, it is limited to ports on the same switch. For optimal availability it is recommended to have connections to redundant switches.

The Cisco switches have a feature called Virtual Port-Channel (vPC) that provides the ability to configure a port-channel that spans multiple switches. vPC requires two 100 Gb peer links that are configured on ports 33 and 35.

vPC also requires a keepalive connection which should be configured on the management port (mgtm0). vPC is used to configure LACP groups that span both switches for the four ports on each ZS7-2 controller and the bondeth1 backup network ports on the database nodes.

The following table shows the vPC configuration for a backup network deployed on ZS7 ToR switches in a full rack X8M configuration.

Table 1. Virtual port channel (vPC) configuration (ZS7 ToR switches – X8M Full rack)

Port-Channel	Switch 1	Switch 2	Comments
Po1(SU)	Eth1/33(P), Eth1/35(P)	Eth1/33(P), Eth1/35(P)	vPC peer link (LACP)
Po10(SU)	Eth1/1/1(P), Eth1/1/2(P) Eth1/1/3(P), Eth1/1/4(P)	Eth1/1/1(P), Eth1/1/2(P) Eth1/1/3(P), Eth1/1/4(P)	ZS7-2 Controller 1 (LACP)
Po11(SU)	Eth1/2/1(P), Eth1/2/2(P) Eth1/2/3(P), Eth1/2/4(P)	Eth1/2/1(P), Eth1/2/2(P) Eth1/2/3(P), Eth1/2/4(P)	ZS7-2 Controller 2 (LACP)
Po12(SU)	Eth1/5/1(P)	Eth1/5/1(P)	db01 backup network (LACP)
Po13(SU)	Eth1/5/2(P)	Eth1/5/2(P)	db02 backup network (LACP)
Po14(SU)	Eth1/5/3(P)	Eth1/5/3(P)	db03 backup network (LACP)
Po15(SU)	Eth1/5/4(P)	Eth1/5/4(P)	db04 backup network (LACP)
Po16(SU)	Eth1/6/1(P)	Eth1/6/1(P)	db05 backup network (LACP)
Po17(SU)	Eth1/6/2(P)	Eth1/6/2(P)	db06 backup network (LACP)
Po18(SU)	Eth1/6/3(P)	Eth1/6/3(P)	db07 backup network (LACP)
Po19(SU)	Eth1/6/4(P)	Eth1/6/4(P)	db08 backup network (LACP)

MTU

The Maximum Transmission Unit (MTU) setting helps determine the largest packet size that can be communicated over the network. Jumbo frames (9000 bytes) reduce the level of IP fragmentation and allow for more efficient processing of large streaming workloads with lower CPU overhead.

When using ZS7-2 top of rack (ToR) switches for the backup network, configure the switch ports for jumbo frames and set the MTU to 9126 on the switch for the ZS7-2 and the database nodes.

IP networks perform path MTU discovery and auto-negotiate the MTU size between the client, switches and server. Client interfaces using a smaller MTU can still communicate with switches and servers that support jumbo frames.

Configuration Steps

Configure and optimize the switches to use vPC and LACP for maximum performance and redundancy. There are three steps in the configuration process. First, enabling the switches for vPC and establishing peer links. Second, configuring LACP and port-channel groups for the storage controllers. Third, configuring LACP and port-channel groups for the database nodes. See specific commands in My Oracle Support Document 2635423.1 "Set Up and Configure Exadata X8M Backup with ZFS Storage ZS7-2".

CONFIGURING THE ORACLE ZFS STORAGE APPLIANCE

This section provides best practices for optimizing an Oracle ZFS Storage Appliance system to provide Oracle Database protection in an Oracle Exadata X8M environment with a standard RMAN backup strategy.

Choosing a Controller

Oracle ZFS Storage Appliance systems are currently available in two models: Oracle ZFS Storage ZS7-2 mid-range and Oracle ZFS Storage ZS7-2 High-End. The following table provides details for each model.

Table 2. Oracle ZFS Storage Appliance Details

Features	Oracle ZFS Storage ZS7-2 Mid-Range	Oracle ZFS Storage ZS7-2 High-End
CPU Cores	72	96
DRAM	1 or 2 TB	3 TB
Maximum Read-Optimized Flash	Up to 737 TB	Up to 1.4 PB
Maximum Write-Optimized Flash	Up to 18.6 TB	Up to 37.5 TB
Raw Storage Capacity	24 TB to 8 PB Scalability	24 TB to 16 PB Scalability
High-Availability Cluster Option	Yes	Yes
Focus	Mid-range	Scalability

Please refer to oracle.com/storage/nas/index.html for the latest Oracle ZFS Storage Appliance model specifications.

Oracle ZFS Storage ZS7-2 High-End is a flagship product that offers maximum levels of scalability, CPU, and DRAM. This is a highly scalable platform that can support up to 16 PB of raw storage capacity.

Oracle ZFS Storage ZS7-2 Mid-Range is a cost-efficient model that can still achieve high levels of throughput and redundancy but does not provide the same level of scalability that the high-end model does.

A ZS7-2 high-end model is recommended for Exadata backup.

Choosing the Correct Disk Shelves

Oracle ZFS Storage Appliance systems include a configurable number of disk shelves. A standard RMAN backup use case should use high-capacity (DE3-24C) disk shelves. Oracle Storage Drive Enclosure DE3-24C features high-capacity 14 TB disks. Each disk shelf contains 24 disks and can be configured with optional write-optimized flash. (Up to four disks per disk shelf can be replaced with solid-state drive [SSD] write-flash accelerators.) Oracle ZFS Storage ZS7-2 can be customized based on disk shelf and write-optimized flash requirements. A minimum of 3 disk shelves should be included in the configuration to provide full redundancy. This will allow for a No Single Point of Failure (NSPF) storage configuration that can tolerate the loss of a disk shelf.

For standard RMAN backup use cases it is recommended to include write-flash accelerators at a minimum ratio of 4 per every 4 disk shelves. For example, an 8 disk shelf configuration would include 6 shelves of 24 14 TB drives and 2 shelves of 20 14 TB drives and 4 write-flash accelerators. If standard RMAN backup is the primary use case but the ZS7-2 will also be handling other use cases that generate a significant number of synchronous writes, then it is recommended to include write-flash accelerators at a minimum ratio of 4 per every 2 disk shelves.

Choosing a Storage Profile

When a storage profile is selected to protect Oracle Exadata, mirrored, single-parity, and double-parity profiles are all worthy of consideration. The following table provides a comparison of the storage profiles. For a standard RMAN backup use case it is recommended to use a double parity storage profile.

Table 3. Storage Profile Comparison [1]

	Usable Capacity	Advantages	Negatives
Mirrored	42.2%	Restore performance Maximum protection Maximum flexibility	Costly
Single Parity	69.3%	Backup performance Moderate flexibility	Limited redundancy
Double Parity	76.7%	Streaming performance Most efficient	Limited IOPS

Double Parity

Double parity provides the best usable capacity and performs as well as single parity for large streaming I/O, which is typical for standard Oracle RMAN workloads. It accomplishes this 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, but it 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 profiles. The ability to handle IOPS-intensive workloads is severely diminished.

Double parity is recommended when Oracle ZFS Storage Appliance systems are 100 percent dedicated to large sequential workloads, such as traditional Oracle RMAN backup and restore workloads. It is not advisable for use cases such as cloning for DevTest provisioning or utilizing an incrementally updated backup strategy. Mirrored or single-parity profiles are more flexible for handling additional use cases that might result in heavier disk IOPS with lower latencies. Figure 3 reflects raw disk capacity distribution for different storage profiles.

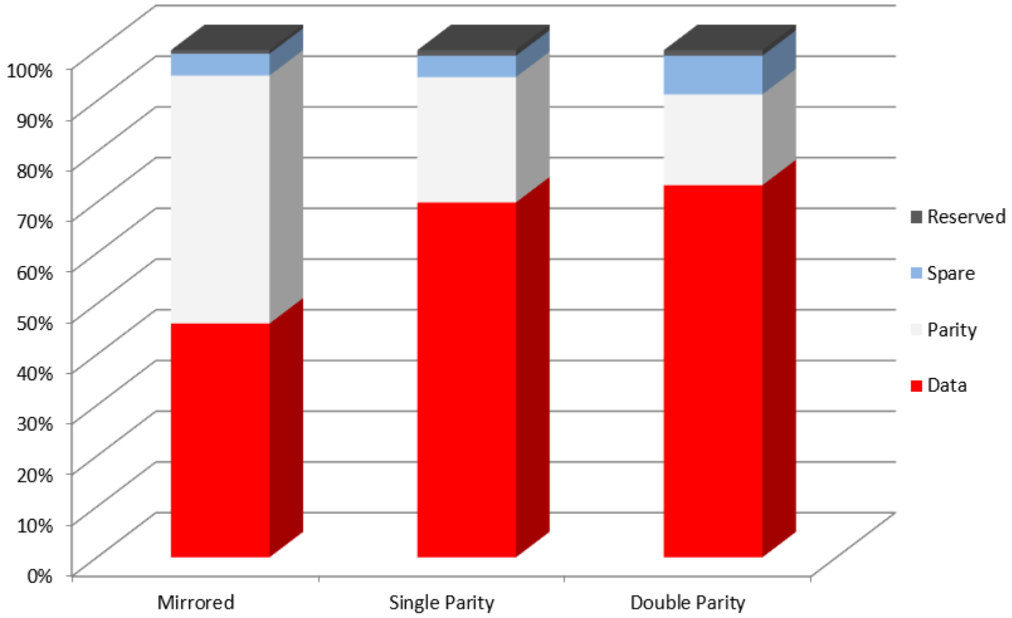


Figure 3. Raw disk capacity distribution

Configuring the Storage Pools

In most situations it is recommended to configure two storage pools with one primary on each controller. This allows for an active/active backup environment that leverages the CPU, memory and networking resources of both ZS7 controllers. This provides better performance and ROI.

[1] **Note**: Useable capacity accounts for raw capacity lost due to parity, spares, and file system overhead, as well as small amounts of space lost on each disk due to operating system (OS) overhead, drive manufacturer overhead, and scratch space reservations. This will vary slightly depending on the size of the storage pool; this example assumes a configuration with four disk shelves.

In unusual situations with strict performance SLAs that must be maintained even during failure scenarios it may be recommended to configure a single storage pool and enable an active/standby backup environment where one of the ZS7 controllers will sit idle waiting to take over for the active controller.

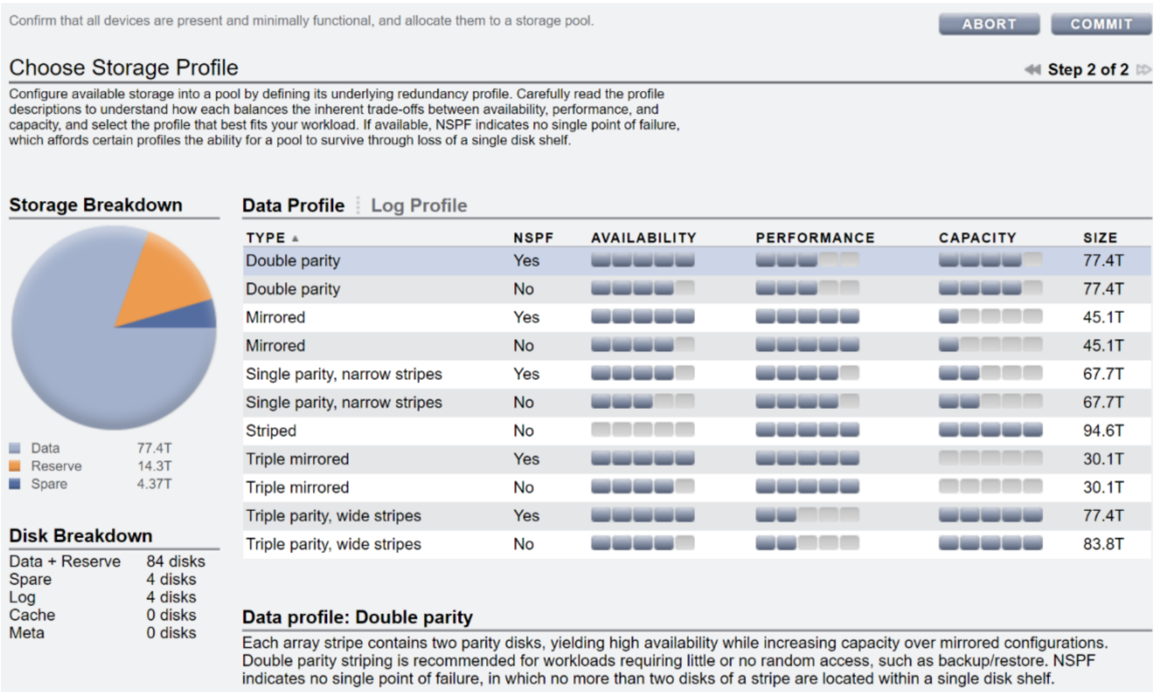
In an active/active backup environment each storage pool should be configured with half of the data disk drives in each Oracle Storage Drive Enclosure DE3-24C disk shelf. Each storage pool should also include half of the write-flash accelerators in each DE3-24C disk shelf. This allows for maximum performance and redundancy.

It is recommended to select the No Single Point of Failure (NSPF) option when configuring the storage pool. This ensures that the loss of an entire disk shelf does not compromise the availability of data. To enable NSPF a minimum of three disk shelves is needed for double-parity profiles.

Follow these steps to configure storage pools for an active/active standard backup deployment using the ZFS Storage Appliance's Browser User Interface (BUI).

Build a storage pool that will be primary on controller 1:

1. Access the BUI on controller 1, select **Configuration > Storage** and **configure a storage pool** by clicking the plus icon next to Available Pools.
2. **Provide a descriptive name for the new storage pool**. Controller 1 will have primary ownership for this storage pool. In this example the storage pool is named "BACKUP1". **Click apply** to move to the device allocation screen.
3. Select half of the data devices and half of the log devices from **each disk shelf** to be included in the storage pool. This method of device allocation will provide optimal performance and redundancy. **Select commit** to move to the storage profile configuration screen.
4. **Choose a data profile of double parity with NSPF**. For maximum redundancy it is recommended to select a log profile of mirrored. **Select commit** to complete the configuration and build the new storage pool.



TYPE	NSPF	AVAILABILITY	PERFORMANCE	CAPACITY	SIZE
Double parity	Yes				77.4T
Double parity	No				77.4T
Mirrored	Yes				45.1T
Mirrored	No				45.1T
Single parity, narrow stripes	Yes				67.7T
Single parity, narrow stripes	No				67.7T
Striped	No				94.6T
Triple mirrored	Yes				30.1T
Triple mirrored	No				30.1T
Triple parity, wide stripes	Yes				77.4T
Triple parity, wide stripes	No				83.8T

Figure 4. Choosing a storage profile

Build a storage pool that will be primary on controller 2:

1. Access the BUI on **controller 2**, select **Configuration > Storage** and **configure a storage pool** by clicking the plus icon next to Available Pools.
2. **Provide a descriptive name** for the new storage pool. Controller 2 will have primary ownership for this storage pool. In this example the storage pool is named "BACKUP2". **Click apply** to move to the device allocation screen.
3. Select the remaining data and log devices **from each disk shelf** to be included in the storage pool. **Select commit** to move to the storage profile configuration screen.

4. **Choose a data profile of double parity with NSPF**. For maximum redundancy it is recommended to select a log profile of mirrored. **Select commit** to complete the configuration and build the new storage pool.

Configuring the Projects and Shares

A project provides a single access and control point for managing filesystems (shares). Projects can be used for grouping logically related shares together. Shares within a project typically share common settings. Quotas can be enforced at the project level in addition to the share level.

It is recommended to use a project for managing the Oracle Exadata backup solution. Projects for different storage pools can share the same name.

Create a project for the BACKUP1 storage pool:

1. Access the BUI on **controller 1**, select **Shares > Projects** and **create a new project** by clicking the plus icon next to Projects.
2. **Provide a descriptive name** for the new project such as "bkup_x8m"**.**
3. **Click apply** to create the new project.

Repeat this process to create a project on the BACKUP2 storage pool by accessing the BUI on **controller 2**.

Project settings can be configured so that any filesystem (share) created within the project with inherit settings from the project.

Optimize the new project settings for a standard RMAN backup use case:

1. Access the BUI on **controller 1**, select **Shares > Projects > "bkup_x8m" (project name) > General**.
2. **Configure the mountpoint**. A mountpoint of "/export/zs1" was used in this example.
3. **Select LZ4 compression**. LZ4 compression should be enabled on the storage. It provides additional benefit when combined with Oracle Database compression by reducing the bandwidth to back-end disk with only minor impact to Oracle ZFS Storage Appliance ZS7-2 CPU utilization.

   Physical network throughput is typically increased when using LZ4 since SAS bandwidth and disk utilization are often limiting factors.

   With Advanced Row Compression enabled for an OLTP Oracle Database, LZ4 typically provides additional space savings in the range of 1.8x to 2.4x.
4. **Set cache device usage to "do not use cache devices"**. Read-optimized flash should not be used for caching standard Oracle RMAN workloads because there is little benefit from storing Oracle RMAN backup sets in cache. Moreover, the level 2 ARC is not intended for streaming workloads.
5. **Set synchronous write bias to throughput**. This is a share setting that controls behavior for servicing synchronous writes. It can be optimized for latency or throughput.

   All writes are initially written to the ZFS adaptive replacement cache (ARC), regardless of whether they are asynchronous, synchronous, latency-optimized, or throughput-optimized. Also, all writes are copied from the ARC to the storage pool.

   An asynchronous write returns an acknowledgement to the client after the write to ARC is complete. When synchronous writes are optimized for throughput, an acknowledgement is not returned until the write is copied to the storage pool.

   When synchronous writes are optimized for latency, an additional copy is written to persistent storage so that acknowledgements can be returned to the client faster. When write-optimized flash is configured in the storage pool, it is used as the persistent storage for latency-sensitive synchronous writes.

   Standard RMAN backup use cases generate bandwidth-sensitive workloads and writes are mostly asynchronous.
6. **Configure the record size to 1 MB**. The record size influences the size of back-end disk I/O.

   Optimal settings depend on the network I/O sizes used by the application—in this case, Oracle RMAN. Standard Oracle RMAN workloads with Direct NFS Client generate large 1 MB writes and reads at the network layer. In this case, a 1 MB record-size setting should be used.

   The ability to use large record sizes has significant advantages, such as increased throughput performance, which is critical for bandwidth-intensive workloads. Other benefits include reduced utilization of controller CPU resources.

In recent years, HDD capacities have grown as quickly as ever, yet the IOPS these disks can deliver has leveled off. Oracle RMAN workloads often generate datasets on the TB scale, with only a small frequency of read-backs. As such, caching is not an optimal solution for handling IOPS. Maximizing the throughput and limiting the IOPS to disk are important factors for achieving the best performance from the backup solution. Oracle RMAN standard backup strategies enable this by delivering large, multichannel network I/O that greatly benefits from large record sizes on the filesystem.

7. **Configure permissions for the oracle user (1001) and dba group (1002) with file access of 750**. Reference figure 5 for an example.
8. **Click apply** to configure the default settings for all shares in this project.

Repeat this process to optimize settings for the other storage pool's project by accessing the BUI on **controller 2**.

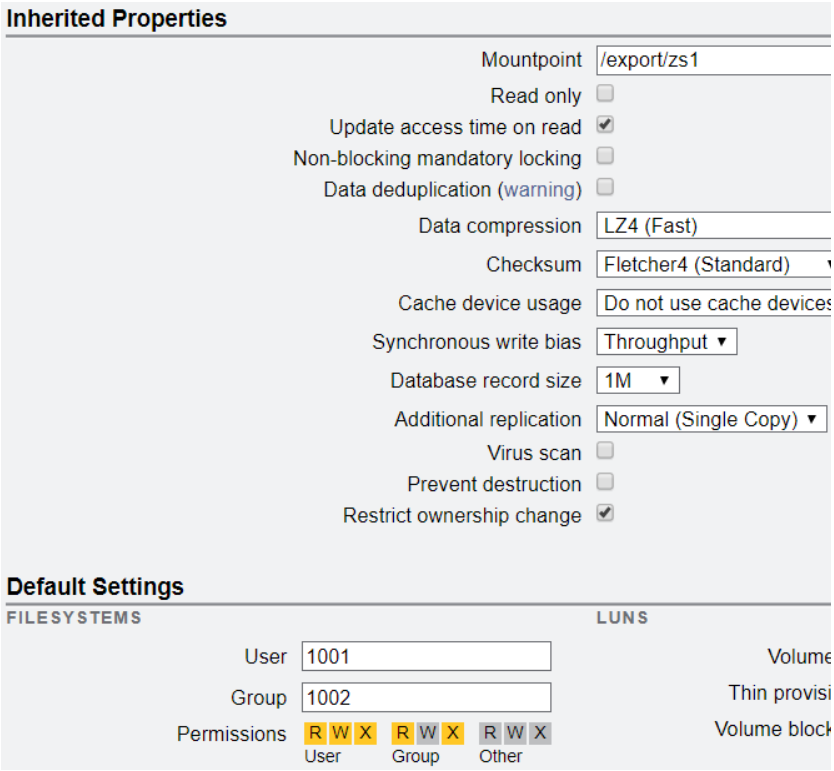


Figure 5. Setting default share settings for each project

Configure the NFS share mode:

1. Navigate to **Shares > Projects > "bkup_x8m" (project name) > Protocols** and set the NFS share mode to **read/write**.
2. Ensure that the NFS protocols settings match figure 6 and then **click apply**.
3. Repeat this step for the other project.

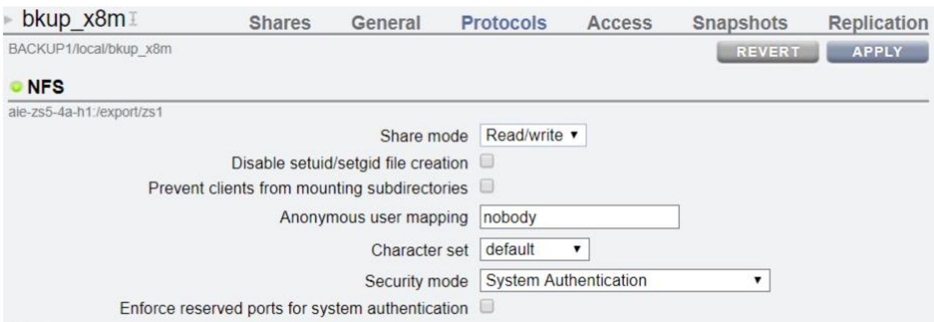


Figure 6. NFS protocols settings for "bkup_x8m" project

The next step is to create the shares that will be mounted and accessed by the NFS clients.

Create a filesystem (share) for the BACKUP1 storage pool:

1. Access the BUI on **controller 1** and select **Shares > Projects > "bkup_x8m" (project name) > Shares**.
2. **Create a filesystem** by clicking on the plus icon next to Filesystems.
3. **Assign the filesystem a unique name.** This example creates a filesystem named "bkup_x8m_1". The filesystem will inherit the properties that were configured for the "bkup_x8m" project. No further configuration is required.
4. **Click apply** to create the filesystem.

Create a filesystem (share) for the BACKUP2 storage pool:

1. Access the BUI on **controller 2** and select **Shares > Projects > "bkup_x8m" (project name) > Shares**.
2. **Create a filesystem** by clicking on the plus icon next to Filesystems.
3. **Assign the filesystem a unique name**. This example creates a filesystem named "bkup_x8m_2".
4. **Click apply** to create the filesystem.

Two filesystems have been created, one on each storage pool.

Configuring the NFS Server

Start the NFS service and optimize settings:

1. Access the BUI and navigate to **Configuration > Services** and **select NFS**
2. **Ensure that the service is enabled and online**.
3. **Set the maximum supported version to NFSv4.1**.
4. **Configure the maximum # of server threads to 3000**.
5. **Click apply** to commit the changes.

It is only required to perform this step on one controller.

Configuring IO Cards

It is recommended that the Oracle ZFS Storage Appliance ZS7-2 high-end system be configured with four SAS-3 cards (slots 2, 3, 9 and 10) and two dual port 25Gb Ethernet cards (slots 4 and 8) in each controller.

If the ZS7-2 will be providing data protection services for multiple Exadatas, then two additional 25Gb Ethernet cards can be installed in slots 5 and 7 to allow greater bandwidth from the ZS7 to the ToR switches. This is usually only required in larger installations. The network analytics available in the Oracle ZFS Storage Appliance can help in determining the need for additional bandwidth to the switch.

Figure 7 provides an example of the recommended PCIe configuration.

Disk Slot CPU DIMM Fan PSU SP

ID	MANUFACTURER	MODEL
PCIe 1	-	-
PCIe 2	PMC	Oracle Storage 12 Gb SAS-3 HBA PCIe, low profile: 16 Port
PCIe 3	PMC	Oracle Storage 12 Gb SAS-3 HBA PCIe, low profile: 16 Port
PCIe 4	Oracle	Oracle Dual Port 25Gb Ethernet Adapter
PCIe 5	-	-
PCIe 6	Oracle	Fishworks CLUSTRON 300
PCIe 7	-	-
PCIe 8	Oracle	Oracle Dual Port 25Gb Ethernet Adapter
PCIe 9	PMC	Oracle Storage 12 Gb SAS-3 HBA PCIe, low profile: 16 Port
PCIe 10	PMC	Oracle Storage 12 Gb SAS-3 HBA PCIe, low profile: 16 Port
PCIe 11	LSI Logic	ORCL-INT-SAS3

Figure 7. Oracle ZFS Storage Appliance ZS7-2 high-end PCIe configuration

Configuring Data Path Network

The data path network is represented by the blue 25Gb SFP28 ports in figure 8.

These interfaces are connected to 100Gb switches using 4x25Gb breakout cables. Port 1 of each card is connected to switch 1 and port 2 of each card is connected to switch 2. Reference the network configuration chapter for detailed port mapping.

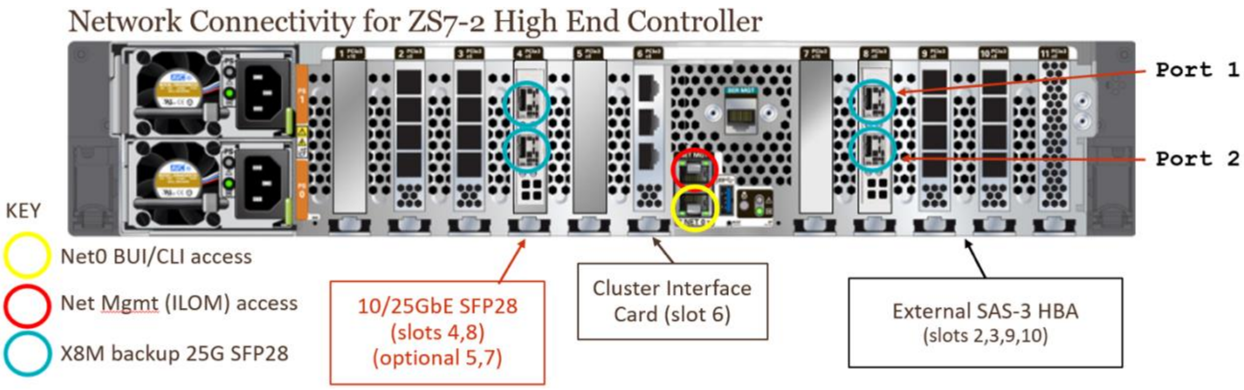


Figure 8. Backup network connectivity for ZS7-2 high-end

A Link Aggregation Control Protocol (LACP) group is used to provide full HA redundancy. The Oracle Direct NFS Client can provide a level of HA, but currently relies on the kernel NFS mount for opening or creating files. LACP is required to provide full HA in all situations.

A Virtual Network Interface Card (VNIC) is defined on the Oracle ZFS Storage Appliance to provide secondary access to the physical link aggregation group. The physical datalink will provide data path services for one of the storage pools while the VNIC will provide data path services for the other storage pool.

Storage pools and the data path network are clustered resources and will transfer ownership during a controller failover.

Create an LACP aggregation of four 25GbE ports:

1. Access the BUI on **controller 1**, select **Configuration > Network > Configuration** and **build a new datalink** by clicking the plus icon next to the Datalinks column.
2. **Provide a descriptive name** for the new datalink. This datalink will provide data path for the BACKUP1 storage pool. Controller 1 has primary ownership for this storage pool. In this example the datalink's name is 4p-lacp-1.
3. **Select the Maximum Transmission Unit** (MTU). The MTU setting helps determine the maximum packet size that can be communicated over the network.

   Jumbo frames (9000 bytes) reduce the level of IP fragmentation and allow for more efficient processing of large streaming workloads with lower CPU overhead. Larger MTUs improve performance for backup networks.

   When using Oracle ToR switches for the backup network, the ports will use an MTU of 9000.

   IP networks will always perform path MTU discovery and auto-negotiate the MTU size between the client, switches and server. Client interfaces using a smaller MTU can still communicate with switches and servers that support jumbo frames.
4. **Select the LACP Aggregation checkbox and add all four 25GbE devices**. Reference figure 9 for an example.
5. **Configure the policy to L3 and L4**. This manages load balancing across the underlying network devices by hashing based on IP as well as (TCP/UDP) port numbers.
6. **Set the timer to short**. This manages the frequency of LACP control messages. A short setting will match the configuration on the Oracle Exadata compute nodes.
7. **Click apply** to create the LACP aggregation datalink.

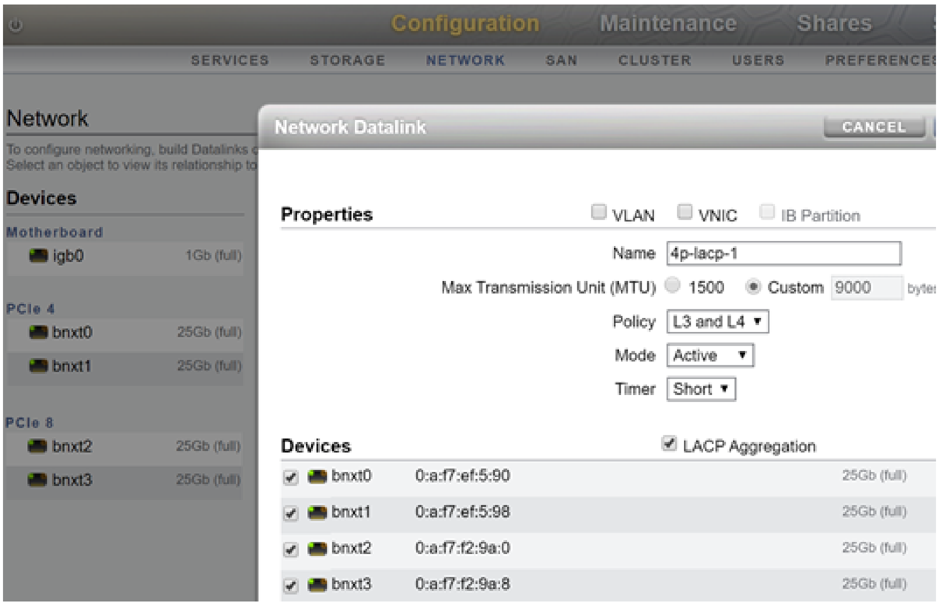


Figure 9. Create 25Gb 4-port LACP datalink

Create a VNIC of the physical link aggregation group:

1. Access the BUI on **controller 2**, select **Configuration > Network > Configuration** and **build a new datalink** by clicking the plus icon next to the Datalinks column.
2. **Provide a descriptive name** for the new datalink. This datalink will provide data path for the BACKUP2 storage pool. Controller 2 has primary ownership for this storage pool. In this example the datalink's name is 4p-lacp-2.
3. **Configure the MTU**. Match the MTU with the previously created LACP aggregation datalink.
4. **Select the VNIC checkbox** under properties and **add the previously created LACP aggregation datalink**. Reference figure 10 for an example.
5. **Click apply** to create the VNIC datalink.

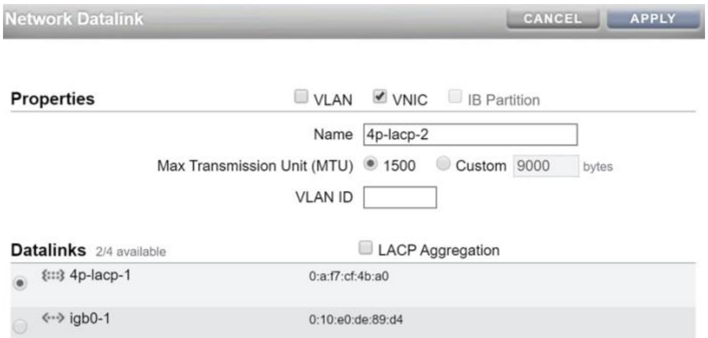


Figure 10. Create a VNIC for data path

Create a network interface for the first storage pool:

1. Access the BUI on **controller 1**, select **Configuration > Network > Configuration** and **build a new interface** by clicking the plus icon next to the Interfaces column.
2. **Provide a descriptive name** for the new interface. This interface will provide data path for the BACKUP1 storage pool. Controller 1 has primary ownership for this storage pool. In this example the interface's name is 4p-lacp-1.
3. **Assign a unique IP address to this interface**. This will be the address used by the Oracle Exadata compute nodes to access the first storage pool.
4. **Select the physical LACP group (4p-lacp-1) as the datalink associated with this interface**. Reference figure 11 as an example.
5. **Click apply** to create the interface.

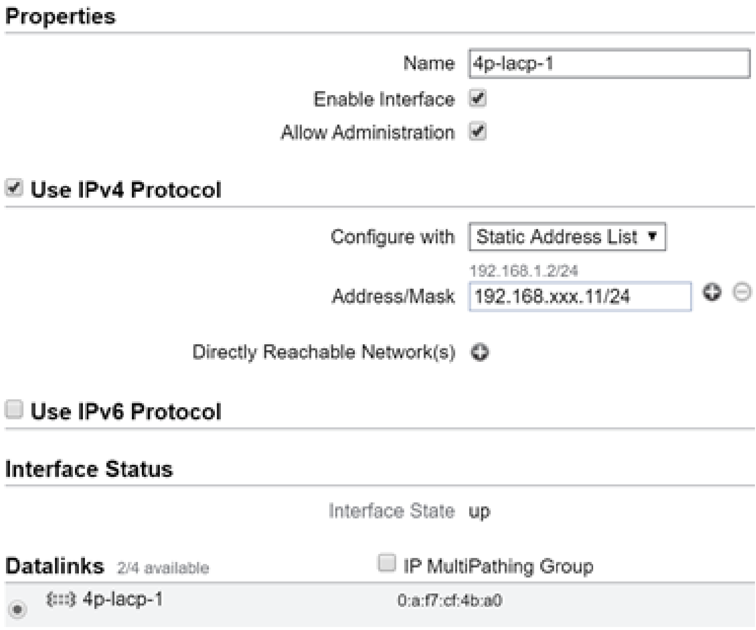


Figure 11. Create an interface for data path

Create a network interface for the second storage pool:

1. Access the BUI on **controller 2**, select **Configuration > Network > Configuration** and **build a new interface** by clicking the plus icon next to the Interfaces column.
2. **Provide a descriptive name** for the new interface. This interface will provide data path for the BACKUP2 storage pool. Controller 2 has primary ownership for this storage pool. In this example the interface's name is 4p-lacp-2.
3. **Assign a unique IP address to this interface**.
4. **Select the VNIC LACP group (4p-lacp-2)** as the datalink associated with this interface.
5. **Click apply** to create the interface.

Enable adaptive routing for the multihoming policy when using LACP to ensure that outbound traffic from an Oracle ZFS Storage Appliance system is balanced over the network links and IP addresses. Access the BUI, select Configuration > Network > Routing, and then select option multihoming=adaptive.

Choosing Direct NFS Client

The Direct NFS Client feature of Oracle Database is highly recommended for all Oracle RMAN workloads between Oracle Exadata and Oracle ZFS Storage Appliance systems, and it is required to achieve optimal performance.

Direct NFS Client is a custom NFS client that resides within the Oracle Database kernel and provides several key advantages:

» Significantly reduces system CPU utilization by bypassing the 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 TCP connection for each Oracle 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 advantages enable increased bandwidth and reduced CPU overhead.
» No additional steps are required on Oracle ZFS Storage Appliance systems to enable Direct NFS Client.

Oracle Intelligent Storage Protocol

Oracle Intelligent Storage Protocol, a feature of Oracle ZFS Storage Appliance systems, was introduced to interact with Direct NFS Client in Oracle Database 12c. It enables Oracle Database–aware storage by dynamically tuning record size and synchronous write bias on Oracle ZFS Storage Appliance systems. This simplifies the configuration process and reduces the performance impact of configuration errors. Hints are passed from the Oracle Database kernel to the Oracle ZFS Storage Appliance system. These hints are interpreted to construct a workload profile to dynamically optimize storage settings.

Oracle Intelligent Storage Protocol is an optional protocol that requires NFSv4 and SNMP. In the current implementation, a properly configured environment that adheres to the best practices in this document performs equally well without Oracle Intelligent Storage Protocol. For instructions on how to enable Oracle Intelligent Storage Protocol, see My Oracle Support document 1943618.1 "Oracle ZFS Storage Appliance: How to Enable Oracle Intelligent Storage Protocol (OISP)."

CONFIGURING THE ORACLE EXADATA DATABASE MACHINE

This section provides best practices for configuring an Oracle Exadata X8M when using an Oracle ZFS Storage Appliance ZS7-2 high end system to perform Oracle Database protection with a standard RMAN backup strategy.

Configuring Backup Network

The backup network is represented by the blue 25Gb SFP28 ports in figure 12 and figure 13.

These interfaces are connected to 100Gb switches using 4x25Gb breakout cables. Port 1 is connected to the top switch and port 2 is connected to the bottom switch. Reference the network configuration chapter for detailed port mapping.

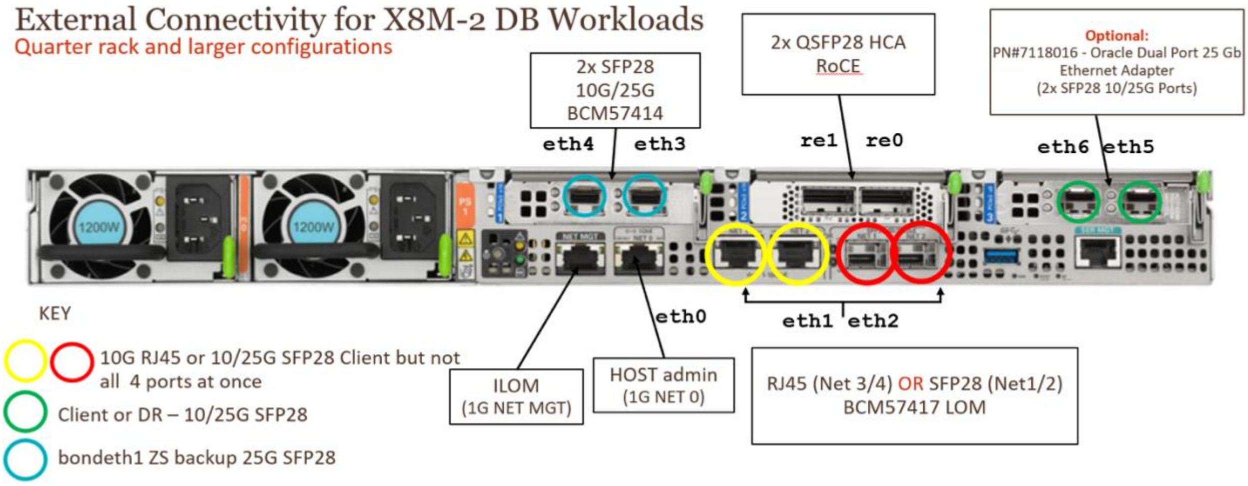


Figure 12. Network connectivity for X8M-2 quarter rack or larger database nodes

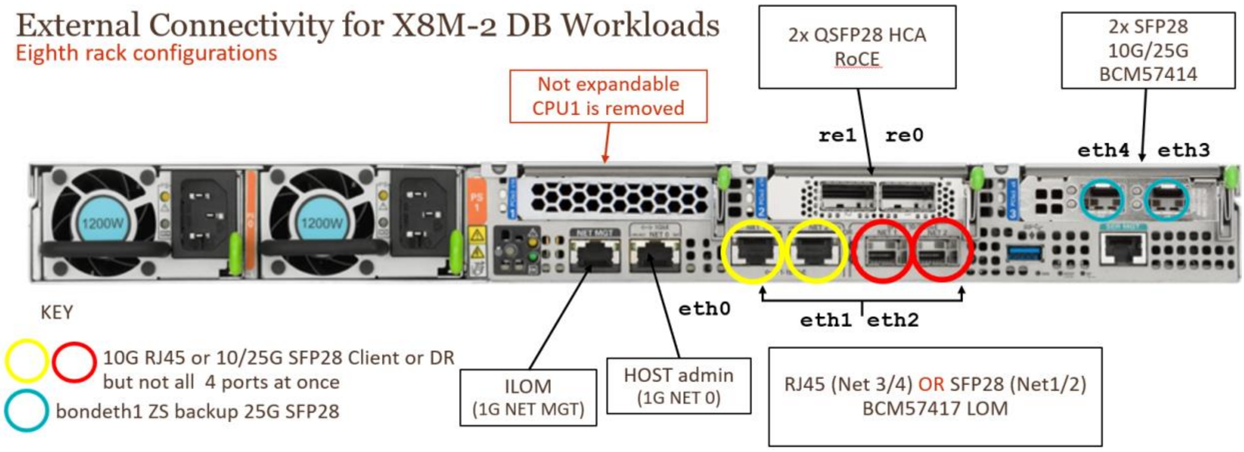


Figure 13. Network connectivity for X8M-2 eighth rack database nodes

The backup network is configured as bondeth1 interface on each database node. This interface is comprised of eth3 and eth4 and can be configured using either Linux bonding in active-backup (mode 1) or Link Aggregation Control Protocol (LACP). It is recommended to use LACP to load balance across both underlying interfaces. The configuration steps in this solution brief assume the use of LACP. Oracle Exadata Deployment Assistant (OEDA) should be used to assist with the configuration of the backup network.

LACP should be configured to use a policy of layer3+4 and a fast polling rate. The default MTU is 1500 bytes. Jumbo frames (9000 bytes) reduce the level of IP fragmentation and allow for more efficient processing of large streaming workloads with lower CPU overhead.

Larger MTUs improve performance for backup networks. When using ZS7-2 top of rack (ToR) switches for the backup network it is recommended to configure the switch ports for jumbo frames.

Optimal settings for bondeth1 are shown in this example:

```
$ cat /proc/net/bonding/bondeth1
Bonding Mode: IEEE 802.3ad Dynamic link aggregation
Transmit Hash Policy: layer3+4 (1)
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 200
Down Delay (ms): 200

802.3ad info
LACP rate: fast
Min links: 0
Aggregator selection policy (ad_select): stable

Slave Interface: eth3
MII Status: up
Speed: 25000 Mbps
Duplex: full
<…>
Slave Interface: eth4
MII Status: up
Speed: 25000 Mbps
Duplex: full
<…>
```

Configure Backup Shares

The following mount options are recommended for shares dedicated to standard RMAN backup use cases:

```
rw,_netdev,hard,rsize=1048576,wsize=1048576,tcp,nfsvers=4,timeo=600 0 0
```

Direct NFS Client does not utilize NFS mount options. However, setting the proper mount options is recommended to follow Oracle Database requirements and to improve performance and functionality if Direct NFS Client is not available and the system reverts to NFS.

Backup shares should be mounted on all Oracle Database nodes.

Mount the Oracle ZFS Storage Appliance backup shares:

1. Edit the **/etc/fstab** file and add entries like the following three-line example

   ```
   ######## lines added for zs7-2 with a backup workload ########
   192.168.x.11:/export/zs1/bkup_x8m_1 /zfssa/zs1/bkup_x8m_1 nfs
   rw,_netdev,hard,rsize=1048576,wsize=1048576,tcp,nfsvers=4,timeo=600 0 0
   192.168.x.12:/export/zs1/bkup_x8m_2 /zfssa/zs1/bkup_x8m_2 nfs
   rw,_netdev,hard,rsize=1048576,wsize=1048576,tcp,nfsvers=4,timeo=600 0 0
   ```

   In this example, 192.168.xxx.11 is the data path IP address for the “BACKUP1” storage pool and 192.168.xxx.12 is the data path IP address for the “BACKUP2” storage pool. The export path of the filesystems are /export/zs1/bkup_x8m_1 and /export/zs1/bkup_x8m_2. The mount points on the compute nodes are /zfssa/zs1/bkup_x8m_1 and /zfssa/zs1/bkup_x8m_2.
2. **Copy these changes to the other compute nodes**.

3. **Create the /zfssa base directory** on all compute nodes (mkdir /zfssa) and change ownership to the oracle user and dba group (`chown oracle:dba /zfssa`).
4. As the oracle user, **create the mount points** used in the fstab on all compute nodes.

   ```
   $ mkdir -p /zfssa/zs1/bkup_x8m_1
   $ mkdir -p /zfssa/zs1/bkup_x8m_2
   ```
5. **Mount the backup shares** on all compute nodes.

   ```
   $ mount /zfssa/zs1/bkup_x8m_1; mount /zfssa/zs1/bkup_x8m_2
   ```

Configuring Direct NFS Client

In Oracle Database 12c and later, Direct NFS Client is enabled by default.

Confirm that Direct NFS Client is enabled by checking the Oracle Database alert log for an Oracle Disk Manager (ODM) message after Oracle Database startup:

```
Oracle instance running with ODM: Oracle Direct NFS ODM Library Version 6.0
```

Direct NFS Client activity can also be confirmed by SQL query:

```
SQL> select * from v$dnfs_servers;
```

Database v$ dnfs views will only be populated if Direct NFS Client IO has occurred since the last database start.

For a complete list of recommended patches, see My Oracle Support document 1495104.1 "Recommended Patches for Direct NFS Client."

Optional oranfstab

The oranfstab file is an optional configuration file that is required to use advanced features of Direct NFS. The Direct NFS client will still function without an oranfstab file and will use NFSv3 from a single source IP address to a single destination IP address. It is not required for this solution.

The oranfstab file configures load spreading of Direct NFS Client connections over multiple addresses on an Oracle ZFS Storage Appliance system (represented by "path") or multiple addresses on Oracle Exadata for Oracle Database (represented by "local"). This is unnecessary if LACP groups are in use since a single source and destination address will be used and load balancing across multiple underlying interfaces will be handled by the Link Aggregation Control Protocol.

The file is created in `$ORACLE_HOME/dbs/oranfstab` and applies to all Oracle Database instances that share `ORACLE_HOME`. When changes are made to the oranfstab file, the database should be bounced for the changes to take effect. NFSv4 is required to take advantage of Oracle Intelligent Storage Protocol (OISP). To enable NFSv4 for the Direct NFS, configure an oranfstab like the following example.

```
server: <STORAGE_NODE1>
local: 192.168.x.101 path: 192.168.x.11
nfs_version: nfsv4
export: /export/zs1/bkup_x8m_1 mount: /zfssa/zs1/bkup_x8m_1
server: <STORAGE_NODE2>
local: 192.168.x.102 path: 192.168.x.12
nfs_version: nfsv4
export: /export/bkup_x8m_2 mount: /zfssa/zs1/bkup_x8m_2
```

Configuring Oracle RMAN Backup Services

Oracle RMAN backup services should be created and used to balance Oracle RMAN workloads across all Oracle Exadata compute nodes. Spreading a backup across multiple Oracle Real Application Clusters (Oracle RAC) nodes improves performance, increases parallel tasks, and reduces utilization load on any single component. Oracle

RMAN backup services are automatically migrated to other Oracle Exadata servers for Oracle Database in the Oracle RAC cluster when the preferred instance is unavailable.

The configuration steps below assume an admin-managed RAC database on an Exadata X8M-2 half rack with four Oracle RAC nodes. Hulk is the database name.

The syntax is: srvctl add service -d <db_name> -r <preferred instance> -a <alternate instance(s)> -s <name for newly created service>

```
[oracle@<HOSTNAME> $ srvctl add service -d hulk -r hulk1 -a hulk2,hulk3,hulk4 -s
hulk_bkup1
[oracle@<HOSTNAME> $ srvctl add service -d hulk -r hulk2 -a hulk1,hulk3,hulk4 -s
hulk_bkup2
[oracle@<HOSTNAME> $ srvctl add service -d hulk -r hulk3 -a hulk1,hulk2,hulk4 -s
hulk_bkup3
[oracle@<HOSTNAME> $ srvctl add service -d hulk -r hulk4 -a hulk1,hulk2,hulk3 -s
hulk_bkup4

[oracle@<HOSTNAME> $ srvctl start service -d hulk -s hulk_bkup1
[oracle@<HOSTNAME> $ srvctl start service -d hulk -s hulk_bkup2
[oracle@<HOSTNAME> $ srvctl start service -d hulk -s hulk_bkup3
[oracle@<HOSTNAME> $ srvctl start service -d hulk -s hulk_bkup4

[oracle@<HOSTNAME> $ srvctl status service -d hulk
Service hulk_bkup1 is running on instance(s) hulk1
Service hulk_bkup2 is running on instance(s) hulk2
Service hulk_bkup3 is running on instance(s) hulk3
Service hulk_bkup4 is running on instance(s) hulk4
```

When the database is restarted the RMAN backup services should be rebalanced. This can be accomplished with:

```
[oracle@<HOSTNAME> $ srvctl stop service -d hulk; srvctl start service -d hulk
```

Preparing Oracle Database for Backup

Archivelog Mode

Archiving of the online redo logs is enabled when Oracle Database is configured to operate in “archivelog” mode. Benefits of using archivelog mode include

» Protection is provided in the event of media failure.
» Oracle Database transactions that occurred after the most recent backup can be recovered.
» Backups can be performed while Oracle Database is open and active.
» Inconsistent backups can be used to restore Oracle Database.

It is recommended that Oracle Database run in archivelog mode. It is also recommended that Data Guard is used to provide site recovery services. If Data Guard is not in use the archivelogs should be duplexed or multiplexed with the primary copy on Oracle Exadata storage and an optional copy on the Oracle ZFS Storage Appliance system.

Block-Change Tracking

Block-change tracking is an Oracle RMAN feature that records changed blocks within a datafile. The level 0 backup scans the entire datafile, but subsequent incremental backups rely on the block-change tracking file to scan just the blocks that have been marked as changed since the last backup.

It is recommended to enable block-change tracking to improve performance for incremental backups. If the chosen backup strategy includes only full or level 0 backups, block-change tracking should not be enabled.

Oracle RMAN Configuration

Compression

Oracle RMAN compression is suitable with Oracle ZFS Storage Appliance systems only when network bandwidth is limited. RMAN compression should not be used with local backups. It is far more efficient to rely on compression at the database level and at the storage level using LZ4. RMAN compression is recommended for remote backups where network bandwidth is a limiting factor.

Oracle RMAN compression should not be used with Transparent Data Encryption (TDE) because it causes the data to be unencrypted, compressed, and re-encrypted during the backup session. This creates a major bottleneck for backup performance and places an enormous strain on Oracle Database CPU resources. Do not use RMAN encryption. TDE tablespace encryption is the recommended method.

Optimizing Channels

Determining the number of Oracle RMAN channels to use is an important aspect of tuning a backup solution. When Oracle RMAN opens a new channel, it allocates a new set of input and output buffers. Each channel can take a datafile or a section of a datafile and process the backup or restore job in parallel to work being done by other channels. Channels can be assigned to different nodes in the Oracle RAC cluster, and can have different backup destinations, with shares potentially owned by different Oracle ZFS Storage Appliance controllers.

Additional channels increase scalability and can provide significantly improved performance, more-efficient resource utilization, load balancing across Oracle Database nodes, a more robust HA architecture, and workload spreading between storage controllers.

As hardware limits are approached, allocating additional Oracle RMAN channels provides diminishing returns. It is not recommended to over-allocate channels because there is no performance gain, despite additional memory and CPU resources allocated for more Oracle RMAN buffers and added complexity in the form of more backup pieces being created.

Determining the recommended number of channels for a configuration depends on the hardware factor that will limit overall performance in an optimally configured solution. Performance limiting components could be many things, including Oracle Exadata or the network, HDD, CPU, or SAS resources. Thorough testing is always recommended when implementing major changes in a production environment. However, the following table provides guidance for how many Oracle RMAN channels to configure in a traditional Oracle RMAN backup strategy for each hardware configuration.

This table assumes an Oracle Exadata X8M-2 and an Oracle ZFS Storage ZS7-2 with storage balanced across both controllers. It assumes that network and SAS bandwidth are not limiting factors, that the best practices in this document are implemented, and that there are no other significant concurrent workloads during the backup window.

When following these sizing guidelines, CPU resources consumed on the X8M database nodes for RMAN workloads will always be less than 5%.

Table 4. Suggested RMAN channels per configuration for a Standard backup strategy

	Channels per Oracle Exadata Eighth Rack	Channels per Oracle Exadata Quarter Rack	Channels per Oracle Exadata Half Rack	Channels per Oracle Exadata Full Rack
1 Disk Shelf	8	8	8	8
2 Disk Shelves	8	12	12	12
3–4 Disk Shelves	8	16	16	16
5–6 Disk Shelves	8	16	24	24
7–8 Disk Shelves	12	20	32	32
9+ Disk Shelves	12	20	32	40

When Oracle RMAN channels are configured or allocated, they should be alternated across the Oracle RAC nodes and storage shares.

Section Size

Enabling highly parallel Oracle RMAN workloads is critical for achieving optimal performance and resource utilization from the backup solution. One challenge is when a very large datafile is encountered. If it is processed by a single Oracle RMAN channel, throughput slows significantly, and other hardware resources in the environment sit idle while waiting for the outlier datafile processing to be completed.

Oracle RMAN's solution to this problem lies in its ability to break up large files into smaller pieces that can be processed in parallel by multiple channels. This is called multisection support and is determined by the section size parameter. It is recommended to set the section size to 100 gigabytes (100G).

Filesperset Parameter

The filesperset parameter determines how many datafiles or sections of datafiles are included in each backup set. When multiple input files are read to create a single backup set, it can improve performance, particularly when the read or copy phases are limiting factors. The default filesperset setting is 64; however, this is detrimental for single-file or partial Oracle Database restore operations because the entire backup set will be read back, even though only a small section is used. Also, an excessively large filesperset setting can impact the load balancing and performance scaling properties of Oracle RMAN. The objective is to have all Oracle RMAN channels effectively utilized throughout the backup. If there is a limited number of datafiles or data sections, it might not be possible to create full backup sets on every channel.

As a general practice, it is recommended to set the filesperset parameter to 1. Testing has shown that this provides excellent performance while load balancing across all channels. If deduplication is enabled, it is a requirement to set filesperset to 1. Including multiple files or sections in the same backup piece diminishes deduplication benefits.

Snapshot Control File

The snapshot control file is not part of the backup. It is a temporary file used by the Oracle RMAN process. It should be placed on local shared Exadata storage. Reference the sample run block for an example.

Sample Run Block

Here is a sample run block for a weekly level 0 backup that can be included as part of an incremental backup strategy. This example assumes an Oracle Exadata X8M-2 half rack backing up to both controllers of an Oracle ZFS Storage Appliance configured with four disk shelves. Hulk is the name of the database used in this example. RMAN backup services are used to evenly spread channels across all four RAC nodes. Channels are alternated between the two storage shares with one owned by each controller. The first file handles persistent configurations and only needs to be run initially and again after any changes to RMAN settings. The second file would be run during every backup cycle.

```
RUN
{
CONFIGURE CHANNEL 1 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup1' FORMAT
'/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 2 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup2' FORMAT
'/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 3 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup3' FORMAT
'/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 4 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup4' FORMAT
'/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 5 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup1' FORMAT
'/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 6 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup2' FORMAT
'/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 7 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup3' FORMAT
'/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 8 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup4' FORMAT
'/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 9 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup1' FORMAT
'/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 10 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup2'
FORMAT '/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 11 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup3'
FORMAT '/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 12 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup4'
FORMAT '/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 13 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup1'
FORMAT '/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 14 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup2'
FORMAT '/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE CHANNEL 15 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup3'
FORMAT '/zfssa/zs1/bkup_x8m_1/backup_%d_%U';
CONFIGURE CHANNEL 16 DEVICE TYPE DISK CONNECT 'sys/passwd@<SCAN_ADDRESS>/hulk_bkup4'
FORMAT '/zfssa/zs1/bkup_x8m_2/backup_%d_%U';
CONFIGURE DEVICE TYPE DISK PARALLELISM 16;
CONFIGURE CONTROLFILE AUTOBACKUP ON;
CONFIGURE CONTROLFILE AUTOBACKUP FORMAT FOR DEVICE TYPE DISK TO
'/zfssa/zs1/bkup_x8m_1/ctrl_%F';
CONFIGURE SNAPSHOT CONTROLFILE NAME TO '+RECOC1/hulk/snapcf_hulk.f';
CONFIGURE ARCHIVELOG DELETION POLICY TO BACKED UP 1 TIMES TO DISK;
}

RUN
{
BACKUP AS BACKUPSET
  SECTION SIZE 100G
  INCREMENTAL LEVEL 0
  DATABASE
  FILESPERSET 1
  TAG 'bkup_weekly_L0';
}
```

CONCLUSION

Selecting the right backup solution for Oracle Exadata is an important decision. Costly third-party alternatives provide poor ROI and cannot support high-performance environments. Competitive offerings are inflexible and do not address all the customer's needs.

Oracle Exadata X8M leverages a new and improved networking infrastructure built on high-speed, low latency 100Gb Ethernet. The solution in this paper details the Oracle ZS7-2 and dedicated top of rack switches to provide a robust, fault-tolerant, and high-performance data protection architecture.

Oracle ZFS Storage Appliance systems have proven to be an ideal solution for protecting the mission-critical data that resides on Oracle Exadata. Powerful features combined with custom Oracle-on-Oracle integrations enable a wide range of Oracle RMAN backup strategies. These provide outstanding performance and flexibility unmatched by third-party solutions.

Extreme restore throughput helps satisfy even the most stringent RTOs. Archive log multiplexing delivers recovery points of 20 minutes or less. Oracle Intelligent Storage Protocol, HCC, LZ4 storage compression, large 1 MB record sizes, data deduplication, and Direct NFS Client provide unique advantages when protecting Oracle Database.

In addition to data protection benefits, an Oracle Exadata backup solution using Oracle ZFS Storage Appliance systems provides many other advantages, such as low-cost, high-performance storage for unstructured data that resides outside of Oracle Database. It is easy to see why Oracle ZFS Storage Appliance systems offer an excellent solution for protecting Oracle Exadata.

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

Exadata X8M and ZFS Storage 25Gb Backup Solution
May 2021
Author: Greg Drobish
Updated by: Greg Drobish, Brock Frank, Joe Hartley

Contributing Authors: Oracle ZFS Storage Team, Oracle Exadata Team, Oracle MAA Team