An Oracle Technical White Paper
May 2013

Configuring a Sun ZFS Backup Appliance with Oracle SPARC SuperCluster

Executive Overview

The SPARC SuperCluster T4-4 is one of the latest innovations from Oracle. It incorporates technologies from other Oracle engineered systems such as Oracle Exadata, Oracle Exalogic, and Oracle's Sun ZFS Storage Appliance. As with Oracle Exadata, the SPARC SuperCluster T4-4 requires a fast backup and recovery solution to protect its Oracle databases.

The Sun ZFS Backup Appliance is a high-performance backup and recovery system that is optimized for Oracle's engineered systems. It is able to connect directly into the SPARC SuperCluster's high-speed InfiniBand infrastructure and operates with Oracle Database's direct NFS (dNFS) feature.

This paper provides instructions, recommendations, and examples for how to configure and tune the Sun ZFS Backup Appliance and SPARC SuperCluster, including:

- Instructions for physically connecting the Sun ZFS Backup Appliance to the SPARC SuperCluster InfiniBand infrastructure
- Instructions for configuring Sun ZFS Backup Appliance storage and networking
- Recommendations for tuning the Oracle Solaris 11 operating system on the SPARC SuperCluster
- Lastly, recommendations and examples on how to configure Oracle Recovery Manager (RMAN) for best performance

Introduction

The configuration of the Sun ZFS Backup Appliance with the SPARC SuperCluster T4-4 is tightly coupled. Only certain ports of the SPARC SuperCluster InfiniBand infrastructure are available for connectivity. Once the connections are made, the Sun ZFS Backup Appliance InfiniBand ports must be activated and then configured on the InfiniBand switches.

There are multiple steps required to properly configure the Sun ZFS Backup Appliance, including setting up InfiniBand networking, creating storage pools, creating RMAN projects and shares, and setting up Sun ZFS Backup Appliance DTrace Analytics for performance monitoring. All of these steps are similar to procedures for configuring a Sun ZFS Backup Appliance for an Oracle Exadata system, and they adhere to current best practices.

The remaining steps are all performed on the SPARC SuperCluster's Oracle Database Logical Domains (LDOMs). These include creating the Sun ZFS Backup Appliance mount points, tuning the Oracle Solaris 11 operating system and networking parameters, configuring dNFS for the database, tuning RMAN parameters, and creating the final RMAN scripts.

Connecting the Sun ZFS Backup Appliance to InfiniBand Switches

Connection of the Sun ZFS Backup Appliance to the SPARC SuperCluster InfiniBand network is dependent on the configured environment. The Sun ZFS Backup Appliance can be:

- Connected directly to SPARC SuperCluster InfiniBand leaf switches – This option is used if the Sun ZFS Backup Appliance is the only other appliance or device (other than SPARC SuperCluster expansion) that will be connected to the infrastructure.
- Connected to external InfiniBand leaf switches – This option is used if more appliances or devices will be connected. In this case, two additional leaf switches are introduced into the InfiniBand network, but are not located within a SPARC SuperCluster or a SPARC SuperCluster expansion rack. A typical scenario for this configuration is when both a Sun ZFS Backup Appliance and Backup Application Media Servers (connected to tape drives) are needed.

Connecting a Sun ZFS Backup Appliance Directly to the SPARC SuperCluster InfiniBand Switches

There are four ports available on each of the InfiniBand (listed as IB in the following directions) leaf switches that can be used to connect to the Sun ZFS Backup Appliance. Connect the cables to these ports as follows:

- Sun ZFS Backup Appliance Head 1
  - PCIe 3 Port 1 to Upper IB Leaf Switch (U24) Port 2A
  - PCIe 3 Port 2 to Lower IB Leaf Switch (U18) Port 2B
  - PCIe 6 Port 1 to Upper IB Leaf Switch (U24) Port 7B
  - PCIe 6 Port 2 to Lower IB Leaf Switch (U18) Port 12A
- Sun ZFS Backup Appliance Head 2
  - PCIe 3 Port 1 to Lower IB Leaf Switch (U18) Port 2A
  - PCIe 3 Port 2 to Upper IB Leaf Switch (U24) Port 2B
  - PCIe 6 Port 1 to Lower IB Leaf Switch (U18) Port 7B
  - PCIe 6 Port 2 to Upper IB Leaf Switch (U24) Port 12A

Connecting a Sun ZFS Backup Appliance to External InfiniBand Leaf Switches

When external InfiniBand switches are used, the eight ports that were used for directly connecting a Sun ZFS Backup Appliance to the SPARC SuperCluster infrastructure are instead used for connecting to the external switches. The external switches are connected as shown in Figure 1. The numbers corresponding with the interconnections represent the number of connections between the switches.

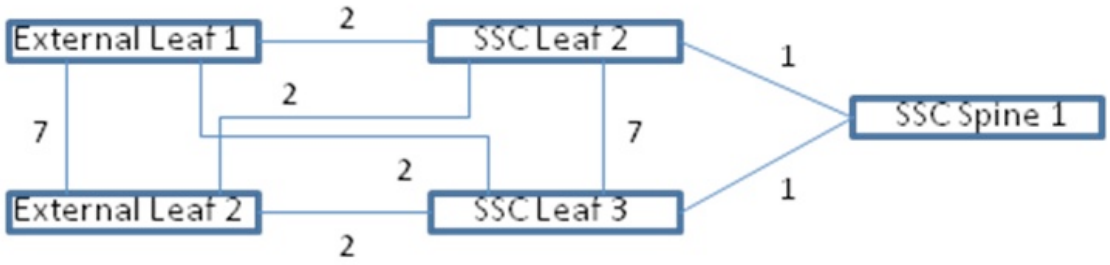

Figure 1. External InfiniBand leaf switches connected to the SPARC SuperCluster (SSC) InfiniBand leaf switches

The port connections on the external leaf switches are not tightly controlled, as are the ports on the SPARC SuperCluster InfiniBand switches. Thus, specific port selection on the external leaf switches is not critical.

Connection of the Sun ZFS Backup Appliance (listed as ZBA in the following connection instructions) to the external leaf switches should follow these basic guidelines:

- Connect the two Port 1 InfiniBand HBA ports on ZBA Head 1 to two ports on External Leaf 1
- Connect the two Port 2 InfiniBand HBA ports on ZBA Head 1 to two ports on External Leaf 2
- Connect the two Port 1 InfiniBand HBA ports on ZBA Head 2 to two ports on External Leaf 2
- Connect the two Port 2 InfiniBand HBA ports on ZBA Head 2 to two ports on External Leaf 1

These connections provide for the maximum throughput and availability.

Configuring InfiniBand Datalinks on the Sun ZFS Backup Appliance

Each of the Sun ZFS Backup Appliance InfiniBand connections must be configured before use.

1. Log on to the ZBA BUI of Head 1 and navigate to **Configuration** -> **Network**.
2. Click on the plus (+) icon next to Datalinks to open the Network Datalink dialog.
3. Complete the Network Datalink dialog as follows (refer to Figure 2):
    a. Check the `IB Partition` box.
    b. Provide a meaningful name for the datalink name.
    c. Set the `Partition Key` to **`8503`**.
    d. Select **`Connected Mode`** for the `Link Mode`.
    e. Do not check the `LACP Aggregation` box.

f. Select `Partition Device` **`ibp0`**.

g. Record the GUID number (for example, `21280001ef43bb`) and click **APPLY**.

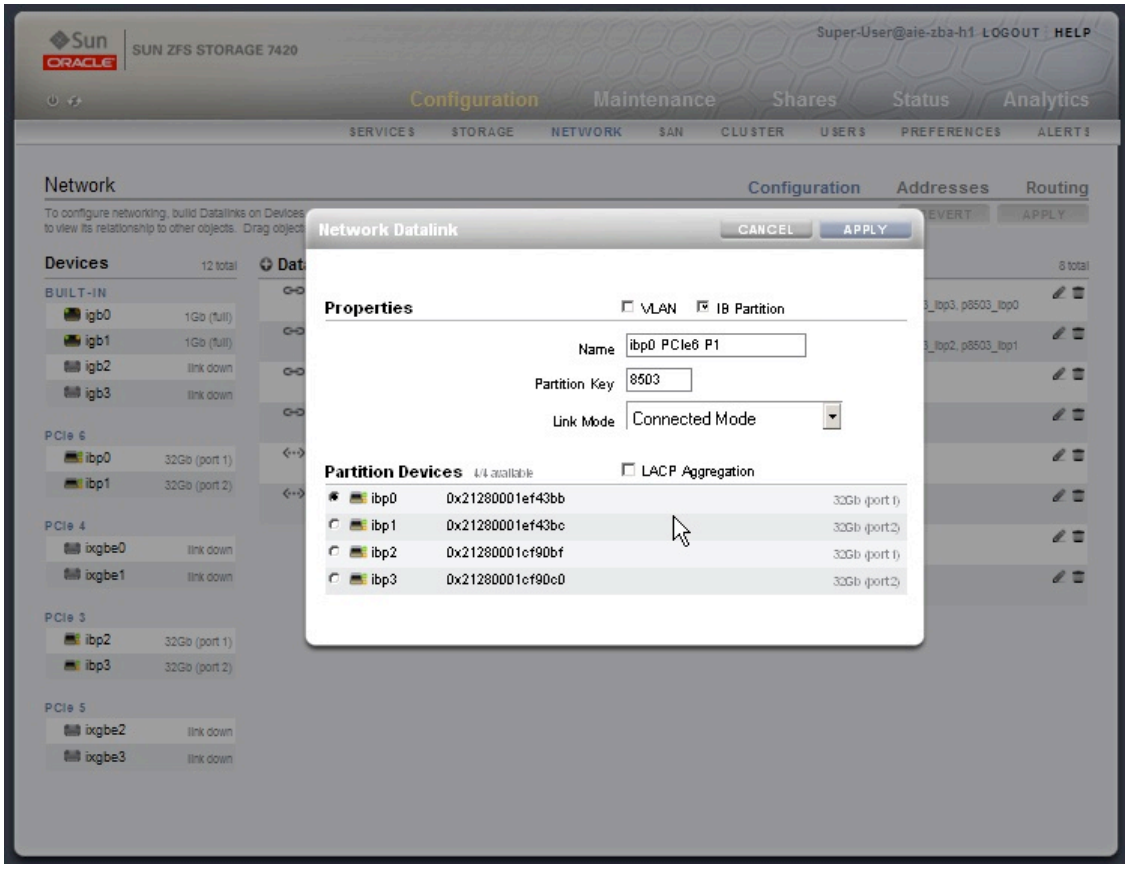

Figure 2. Network Datalink dialog box

4. Repeat Steps 2 and 3 for each remaining InfiniBand interface (`ibp1`, `ibp2`, and `ibp3`).

5. Repeat Steps 1 through 4 for ZBA Head 2.

Reconfiguring the SPARC SuperCluster InfiniBand Switches

The GUIDs (globally unique ID numbers) of the Sun ZFS Backup Appliance InfiniBand HBA ports must be added to the existing SPARC SuperCluster InfiniBand configuration with the partition key of `8503` in order for the components to communicate with each other. Use the following steps to do so.

1. In the CLI, log on to the SPARC SuperCluster InfiniBand spine switch as `root` and verify that it is running subnet manager by executing the `enablesm` command.

```
login as: root
root@aiesscsw-ib1's password:
Last login: Tue Sep 25 08:19:01 2012 from dhcp-brm-b15-204-3e-east-10-135-75-
254.usdhcp.oraclecorp.com
```

```
[root@aiesscsw-ib1 ~]# enablesm
opensm (pid  15906) is already running...
Starting partitiond daemon
/usr/local/util/partitiond is already running
(You may also perform a 'restart' if wanted)
[root@aiesscsw-ib1 ~]#
```

By default, the spine switch is given a hostname of <*sscid*>sw-ib1, where <*sscid*> is the prefix name given to the entire SPARC SuperCluster system. In these examples, the <*sscid* > is aiessc.

The enablesm command will report that the subnet manager partitiond is already running, and will attempt to start it if it is not.

2. Enter the command getmaster to verify that this is the master switch of the configuration. If the master switch is not running on the spine switch, log out and log in to the designated master switch for the remainder of this procedure.

```
[root@aiesscsw-ib1 ~]# getmaster
Local SM enabled and running
20120913 10:16:51 Master SubnetManager on sm lid 13 sm guid 0x2128e8ac27a0a0 :
SUN DCS 36P QDR aiesscsw-ib1.us.oracle.com
[root@aiesscsw-ib1 ~]#
```

3. Back up the switch configuration according to the documented backup procedures. See the documentation for the Sun Datacenter InfiniBand Switch 36 Firmware Version 2.0 at: http://docs.oracle.com/cd/E26698_01/index.html. The doc *Oracle Integrated Lights Out Manager (ILOM) Supplement* at http://docs.oracle.com/cd/E26698_01//pdf/E26432.pdf provides specific instruction for both CLI- and Web-based methods.

4. Enter the command smpartition list active to verify that partition key 0x0503 is assigned to partition name **sto**.

NOTE: The partition key was set to 8503 on the Sun ZFS Backup Appliance datalinks, but the InfiniBand switch reports 0503. This is intentional because the InfiniBand protocol reserves the most significant bit (0x8000) of the hexadecimal partition key (pkey) for its own use. Therefore, pkeys of 0x8503 and 0x0503 are the same.

```
[root@aiesscsw-ib1 ~]# smpartition list active
# Sun DCS IB partition config file
# This file is generated, do not edit
#! version_number : 11
Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=full;
SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
ic1s10 = 0x0501,ipoib,defmember=full:
0x0021280001ef30f7,
0x0021280001ef33bf,
0x0021280001ef30b7,
0x0021280001ef314b;
ic2s10 = 0x0502,ipoib,defmember=full:
0x0021280001ef30f8,
0x0021280001ef33c0,
0x0021280001ef30b8,
0x0021280001ef314c;
sto = 0x0503,ipoib,defmember=full:
0x0021280001ef43f8,
0x0021280001ef43b7,
0x0021280001cf90c0,
```

```
0x0021280001ef43bb,
… more …
```

5. Add the Sun ZFS Backup Appliance to the InfiniBand configuration.

   a. Enter the command `smpartition start` to start a reconfiguration session.

   ```
   [root@aiesscsw-ib1 ~]# smpartition start
   [root@aiesscsw-ib1 ~]#
   ```

   b. Enter the command `smpartition add` to add the eight new GUIDs to the configuration.

   ```
   [root@aiesscsw-ib1 ~]# smpartition add -n sto -port 21280001ef43bb
   21280001ef43bc 21280001cf90bf 21280001cf90c0 21280001ef43f7 21280001ef43f8
   21280001ef43b7 21280001ef43b8
   [root@aiesscsw-ib1 ~]#
   ```

   c. Enter the command `smpartition list modified` to verify the new GUIDs have been added correctly.

   ```
   [root@aiesscsw-ib1 ~]# smpartition list modified
   # Sun DCS IB partition config file
   # This file is generated, do not edit
   #! version_number : 11
   Default=0x7fff, ipoib : ALL_CAS=full, ALL_SWITCHES=full, SELF=full;
   SUN_DCS=0x0001, ipoib : ALL_SWITCHES=full;
   ic1s10 = 0x0501,ipoib,defmember=full:
   0x0021280001ef30f7,
   0x0021280001ef33bf,
   0x0021280001ef30b7,
   0x0021280001ef314b;
   ic2s10 = 0x0502,ipoib,defmember=full:
   0x0021280001ef30f8,
   0x0021280001ef33c0,
   0x0021280001ef30b8,
   0x0021280001ef314c;
   sto = 0x0503,ipoib,defmember=full:
   0x0021280001ef43f8,
   0x0021280001ef43b7,
   0x0021280001cf90c0,
   0x0021280001ef43bb,
   0x0021280001ef43bc,
   0x0021280001cf90bf,
   0x0021280001ef43b8,
   0x0021280001ef43f7,
   0x0021280001ef3048,
   0x0021280001ef30af,
   0x0021280001ef30f8,
   0x0021280001ef30f7,
   0x0021280001ef33c0,
   0x0021280001ef33bf,
   0x0021280001ef30cc,
   0x0021280001ef342b,
   0x0021280001ef30b8,
   0x0021280001ef30b7,
   0x0021280001ef314c,
   0x0021280001ef314b,
   0x0021280001efec65,
   0x0021280001efec66,
   0x0021280001efecb1,
   0x0021280001efecb2;
   ```

   d. Enter the command `smpartition commit` to apply the new configuration and propagate configuration changes to all InfiniBand switches in the configuration.

```
[root@aiesscsw-ib1 ~]# smpartition commit
[root@aiesscsw-ib1 ~]#
```

6. Log off of the InfiniBand switch.
7. Back up the switch configuration according to the documented backup procedures previously noted. (See the documentation for the Sun Datacenter InfiniBand Switch 36 Firmware Version 2.0 at: http://docs.oracle.com/cd/E26698_01/index.html).

Configuring the Sun ZFS Backup Appliance Networking

The InfiniBand ports on the Sun ZFS Backup Appliance must be configured for IP multipathing. Since the interfaces will be running in an active-active configuration, four IP addresses will be required for each Sun ZFS Backup Appliance head (therefore, eight addresses total). These IP addresses must be on the same private storage subnet as the <*HOST*>-stor host addresses (see the SPARC SuperCluster host table [`/etc/inet/hosts`] on one of the SPARC SuperCluster nodes.

To configure each InfiniBand datalink as its own network interface:

1. Log on to the Head 1 BUI and navigate to **Configuration -> Network**.
2. Click on the plus (+) icon next to **Interfaces** to open the Network Interface dialog box.
3. Complete the dialog box as follows (refer to Figure 3):
   a. Enter a meaningful name for the network interface.
   b. Verify that `Enable Interface` is checked.
   c. Verify that `Allow Administration` is checked.
   d. Verify that `Use IPv4 Protocol` is checked.
   e. Verify that the `Configure with` menu selection is `Static Address List`.
   f. In the box below that, enter **`0.0.0.0/8`**.
   g. Verify that `Use IPv6 Protocol` is not checked.
   h. Select the datalink for `ibp0` and click **APPLY**.
4. Repeat steps 2 and 3 for the remaining datalinks (`ibp1`, `ibp2`, and `ibp3`).
5. Repeat steps 1 through 4 for the datalinks on Head 2.

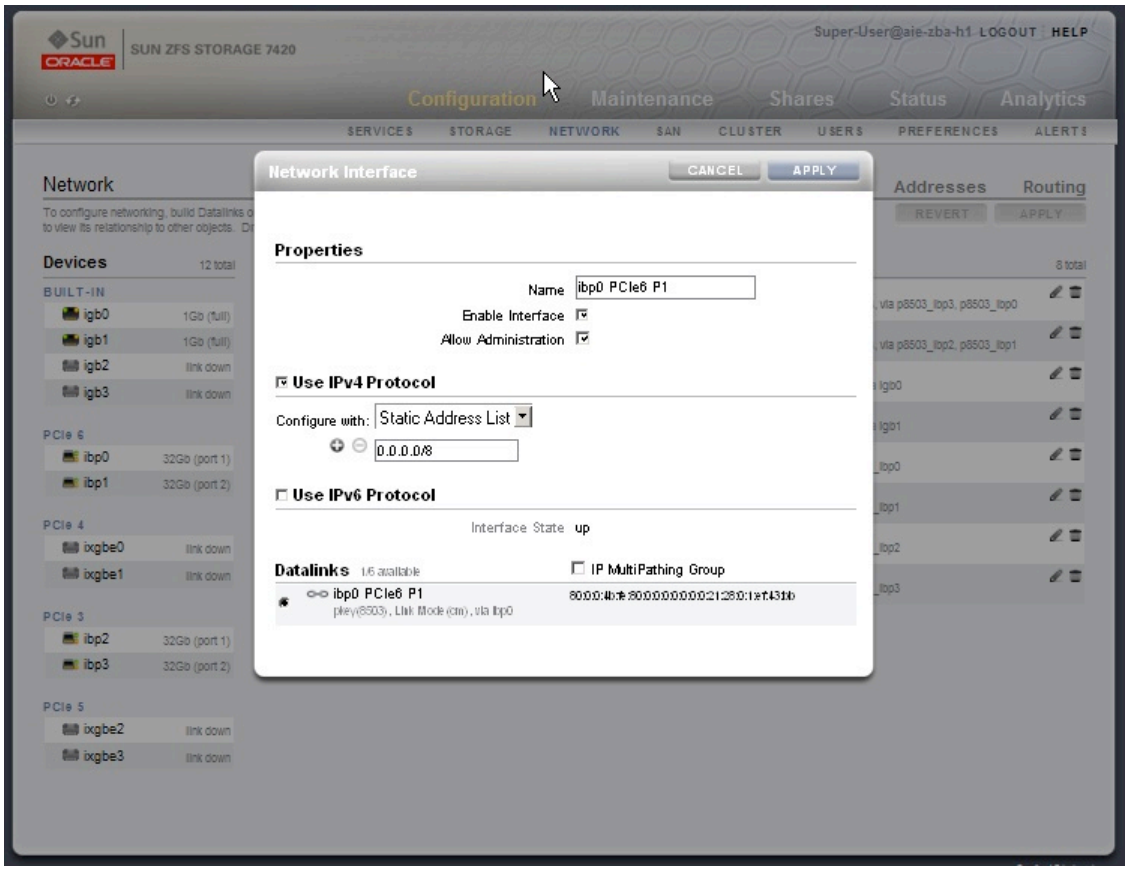

Figure 3. Creating network interfaces for each InfiniBand datalink

To configure the IPMP interface on the Sun ZFS Backup Appliance Head 1:

1. Using the Sun ZFS Backup Appliance BUI of Head 1, navigate to **Configuration -> Network**.
2. Click on the plus (+) icon next to Interfaces to open the Network Interface dialog.
3. Complete the dialog box as follows (refer to Figure 4):
    a. Enter a meaningful name for the IPMP network interface.
    b. Verify that `Enable Interface` is checked.
    c. Verify that `Allow Administration` is checked.
    d. Verify that `Use IPv4 Protocol` is checked.
    e. Verify that the `Configure with` menu selection is `Static Address List.`
    f. Click the plus (+) sign next to the empty box three times, so that a total of four empty boxes are displayed.

g. In each empty box, enter one of the IP addresses reserved for the InfiniBand connections with its respective /24 netmask designation. As a best practice, do not use consecutive IP addresses from the block, but rather every other one (that is, all odd or all even).

h. Verify that `Use IPv6 Protocol` is not checked.

i. Check the `IP MultiPathing Group` box.

j. Check the boxes next to the interfaces corresponding with datalinks **`ibp0`** and **`ibp3`**.

k. Verify that each of the two interfaces are set to **`Active`** and click **APPLY**.

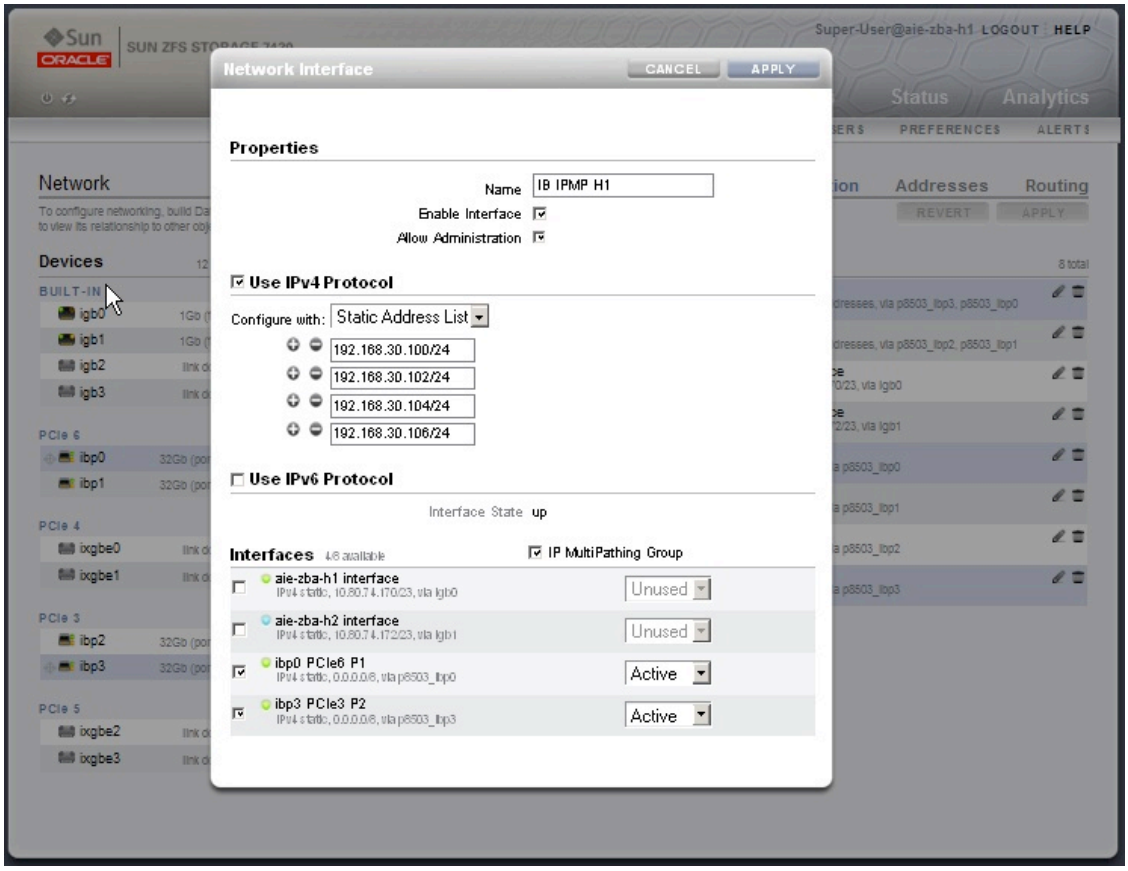

Figure 4. IPMP InfiniBand group for Sun ZFS Backup Appliance Head 1

4. From **Configuration** -> **Network**, click **Routing**.

5. Click on the `Multihoming model` corresponding with `Adaptive` (refer to Figure 5).

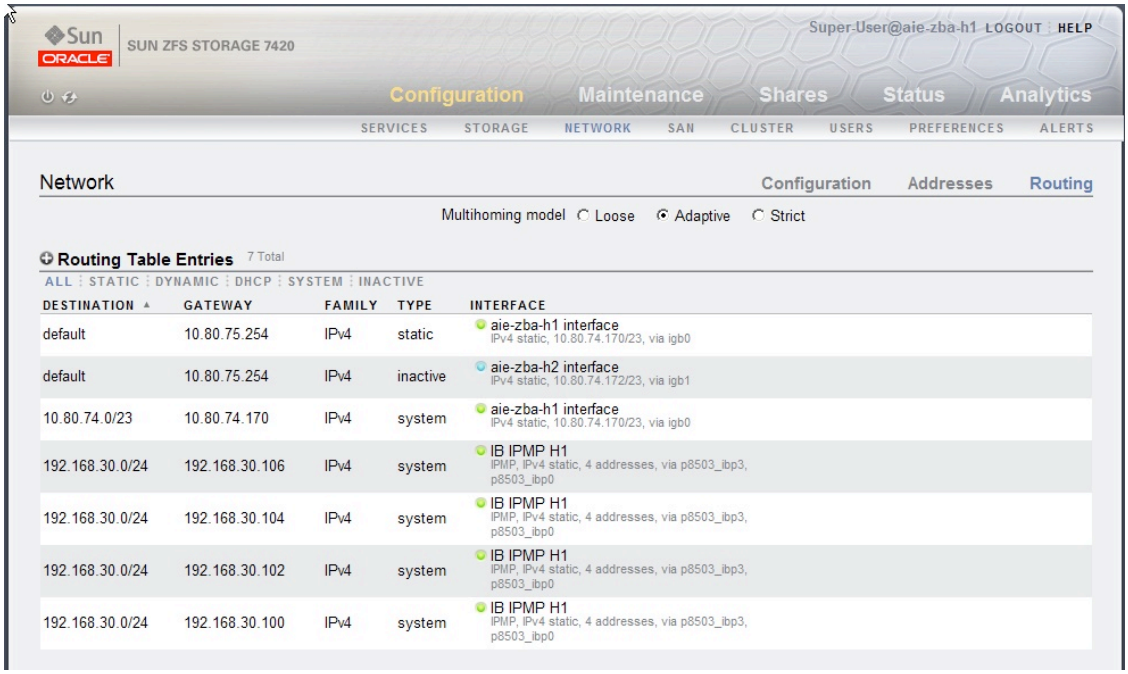

Figure 5. Selecting “Adaptive” for Multihoming model

Next, configure the IPMP interface on Head 2:

1. Log on to the BUI of Head 2 and navigate to **Configuration -> Network**.
2. Click the plus (+) icon next to **Interfaces** to open the Network Interface dialog box.
3. Complete the dialog box as follows (refer to Figure 6):
    a. Enter a meaningful name for the IPMP network interface.
    b. Verify that `Enable Interface` is checked.
    c. Verify that `Allow Administration` is checked.
    d. Verify that `Use IPv4 Protocol` is checked.
    e. Verify that the `Configure with` menu selection is `Static Address List`.
    f. Click the plus (+) sign next to the empty box three times so that a total of four empty boxes are displayed.
    g. In each empty box, enter one of the remaining four IP addresses reserved for the InfiniBand connections with their respective /24 netmask designation. These IP addresses should be the ones not used on Head 1.
    h. Verify that `Use IPv6 Protocol` is not checked.
    i. Check the `IP MultiPathing Group` box.
    j. Check the boxes next to the interfaces corresponding with datalinks **`ibp1`** and **`ibp2`**.

k. Verify that each of the two interfaces is set to **`Active`** and click **APPLY**.

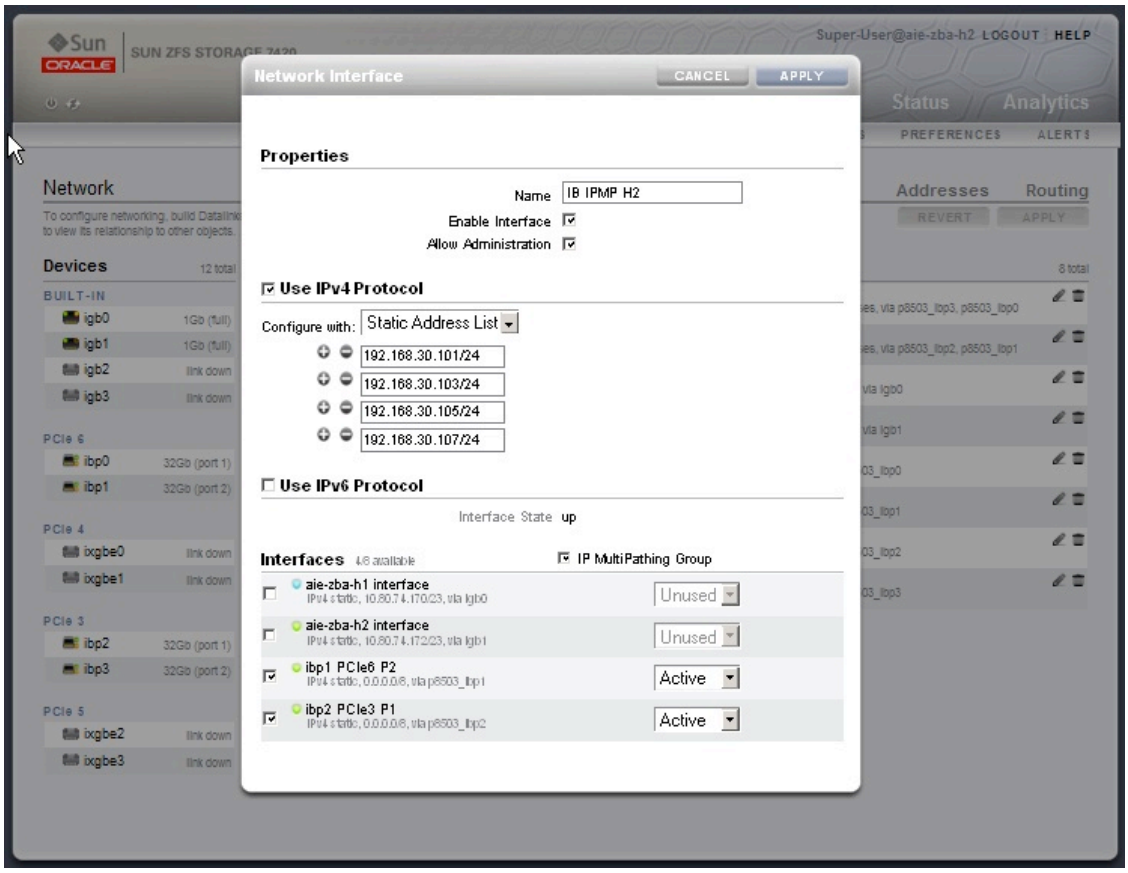

Figure 6. IPMP InfiniBand group for Sun ZFS Backup Appliance Head 2

4. From **Configuration** -> **Network**, click **Routing**.

5. Click on the `Multihoming model` corresponding with **`Adaptive`**.

Finally, verify connectivity with the SPARC SuperCluster's Oracle Database LDOMs.

Verify that each of the SPARC SuperCluster's Oracle Database LDOMs can ping each of the eight addresses used in the IPMP groups on the Sun ZFS Backup Appliance. Add these IP addresses to the `/etc/inet/hosts` table of each SPARC SuperCluster Oracle Database LDOM.

Configuring the Sun ZFS Backup Appliance Storage Pools

Pool configuration assigns physical disk drive resources to logical storage pools for backup data storage. To maximize system throughput, configure two equally sized storage pools by assigning half of the physical drives in each drive tray to each storage pool as shown in Figure 7.

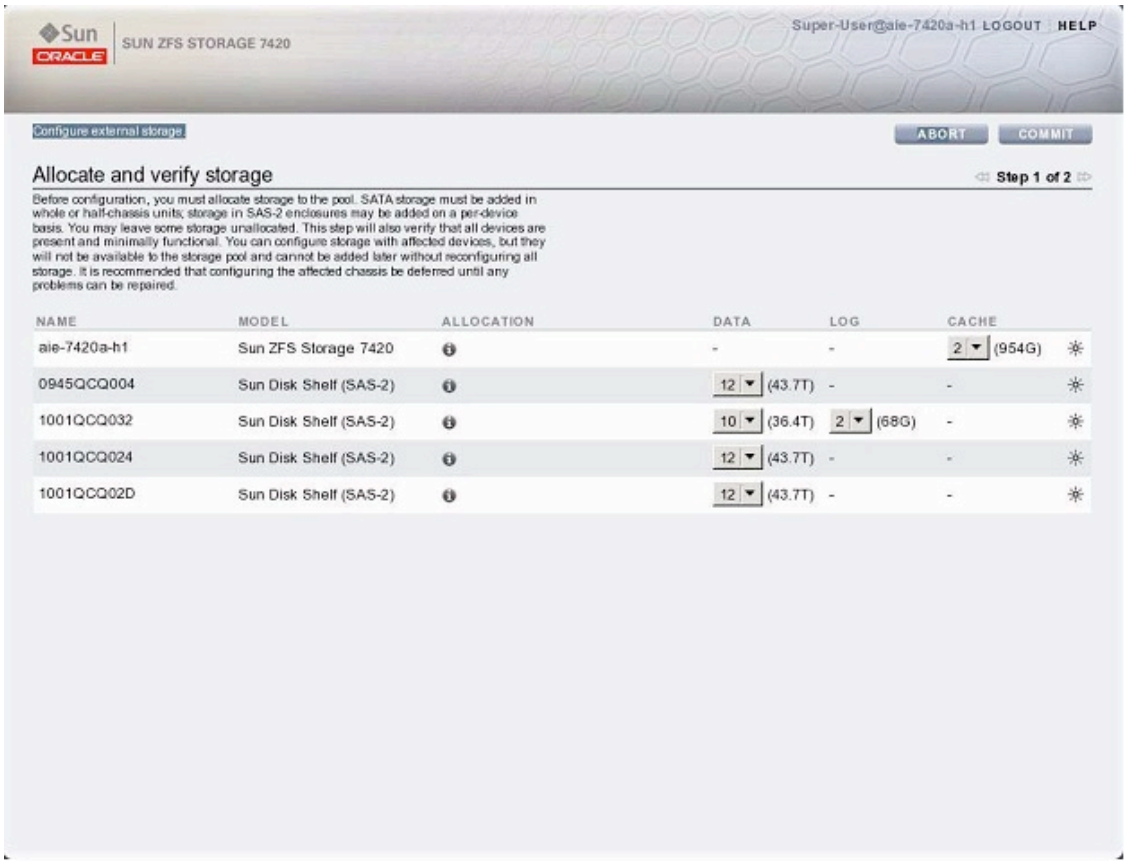

Figure 7. Storage pool configured based on half of the drives in each tray

Configuring the Sun ZFS Backup Appliance RMAN Projects and Shares

Share configuration is the process of setting up and running NFS mount points for client access. Two projects should be created for the SPARC SuperCluster's Oracle Database LDOMs configuration; one project per pool. A project is an entity that provides a higher-level management interface point for a collection of shares. To optimize share management, update the default mount point for shares contained in the project to reference the database name, such as `/export/dbname`. For a performance-optimized system, create four shares for each project in each pool, for a total of eight shares (four on each head). To configure a project, perform the following:

1. Log on to the BUI of Head 1 and navigate to **Shares -> Projects**.
2. Click on the plus (+) icon next to **Projects**, enter a meaningful name for the project, and click on **APPLY**. Since a similar project will be created on the other head, uniquely name the project for Head 1, such as `H1-dbname`.
3. Click the Pencil icon next to the new Project Name to edit the project.

4. Click **General** and complete the properties as follows (refer to Figure 8):
   a. Change the `Mountpoint` to include the database name (for example, `/export/dbname`).
   b. Change `Synchronous write bias` from `Latency` to **`Throughput`** and click **APPLY**.

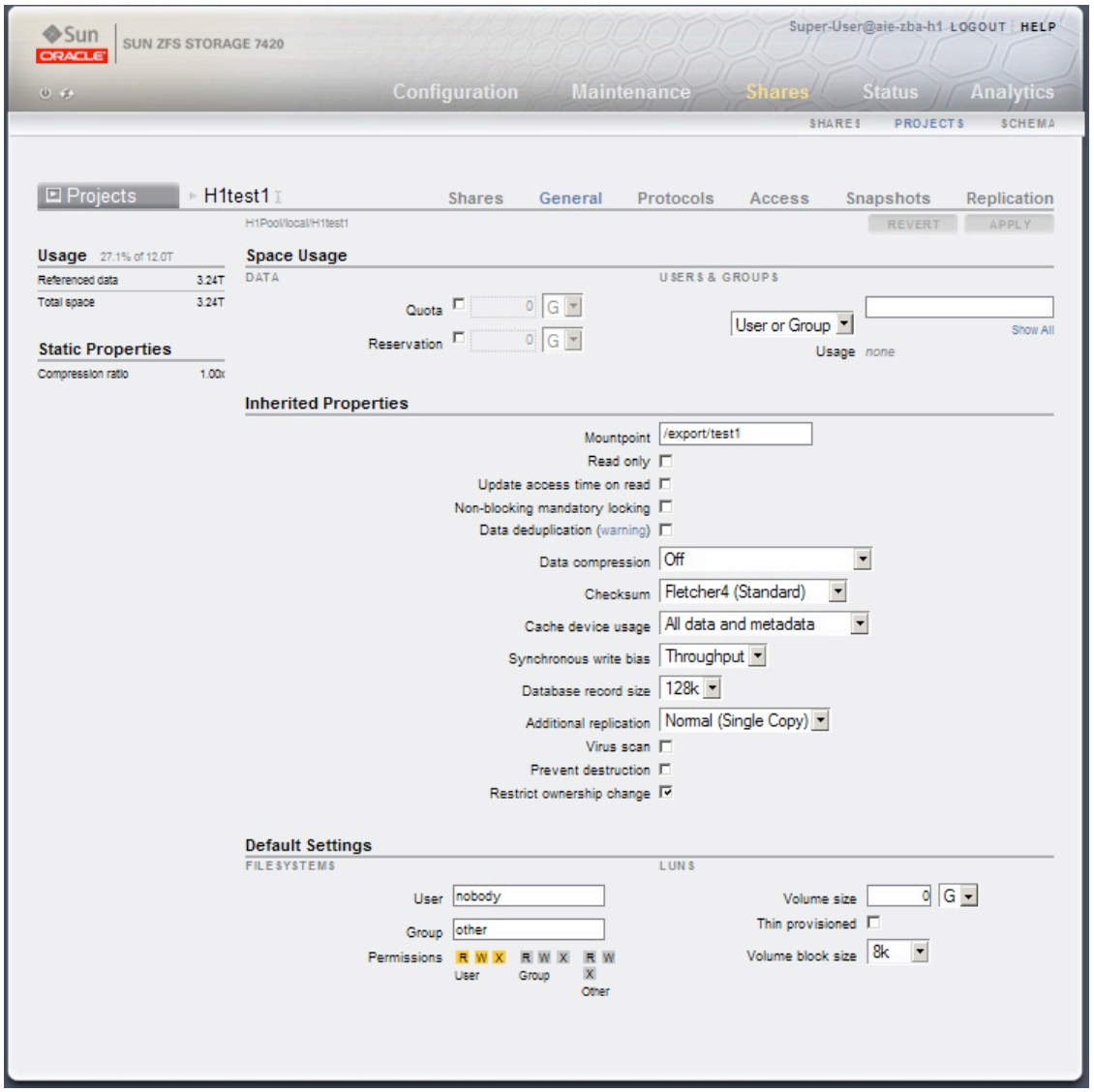

Figure 8. Project general parameter settings

5. Click Protocols and add an NFS Exception as follows (refer to Figure 9):
   a. Click on the plus (+) icon next to **NFS Exceptions**.
   b. Change `Type` to **`Network`**.

c. Enter the subnet and netmask (for example, `192.168.30.0/24`) of the InfiniBand network.

d. Change `Access Mode` to **`Read/write`**.

e. Verify that `Charset` is set to **`default`**.

f. Check the `Root Access` box and click **APPLY**.

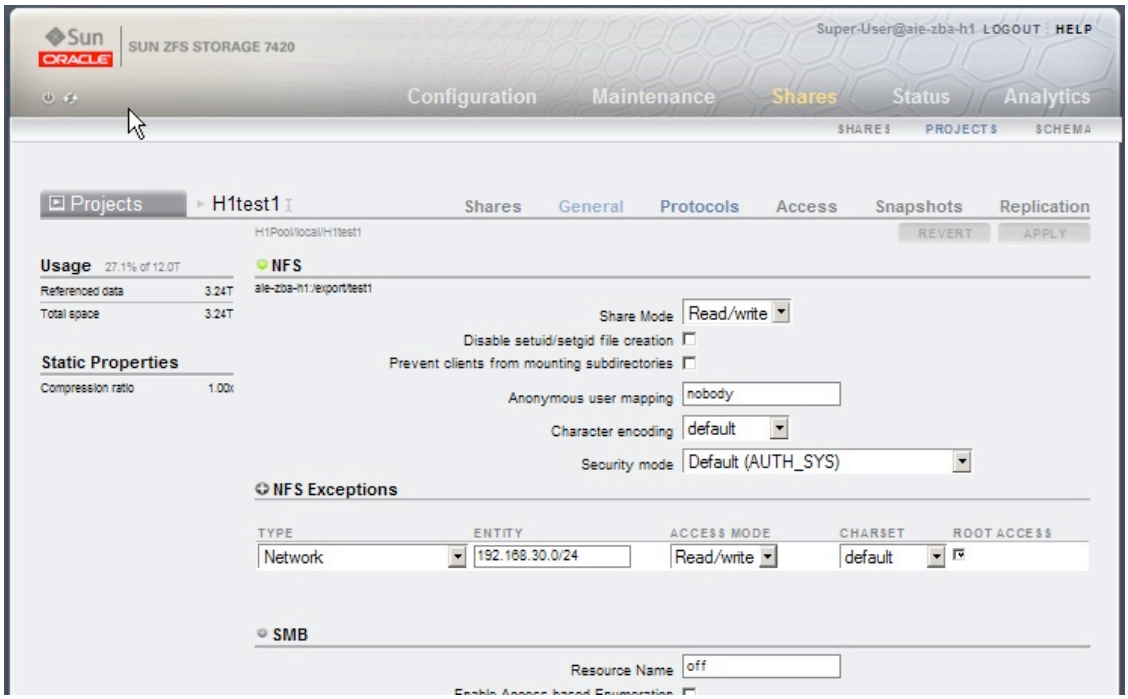

Figure 9. Setting up project NFS exceptions

6. Next to **General**, click **Shares**.

7. Create four filesystems for Head 1 and uniquely name them so they will be different from the names for Head 2. The names of the filesystems on Head 1 should be backup1, backup3, backup5, and backup7. The names of the filesystems on Head 2 will be backup2, backup4, backup6, and backup8. This is done for interleaving of the backup streams to distribute the data across the two heads, thereby providing better performance.

To create the filesystems, click the plus (+) icon next to **Filesystems**, enter the name of the filesystem (`backup1`), and click **APPLY**. Repeat this step to create the remaining three filesystems (`backup3`, `backup5`, and `backup7`). The filesystem listing should be similar to Figure 10.

Figure 10. Filesystem listing for Head 1

8. Repeat steps 1 through 7 using Head 2. Remember to use a unique project name (for example, `H2-dbname`) for the project, and specify the even-numbered backup IDs (`backup2`, `backup4`, `backup6`, and `backup8`) for the filesystem names. The filesystem listing should be similar to the example in figure 11.

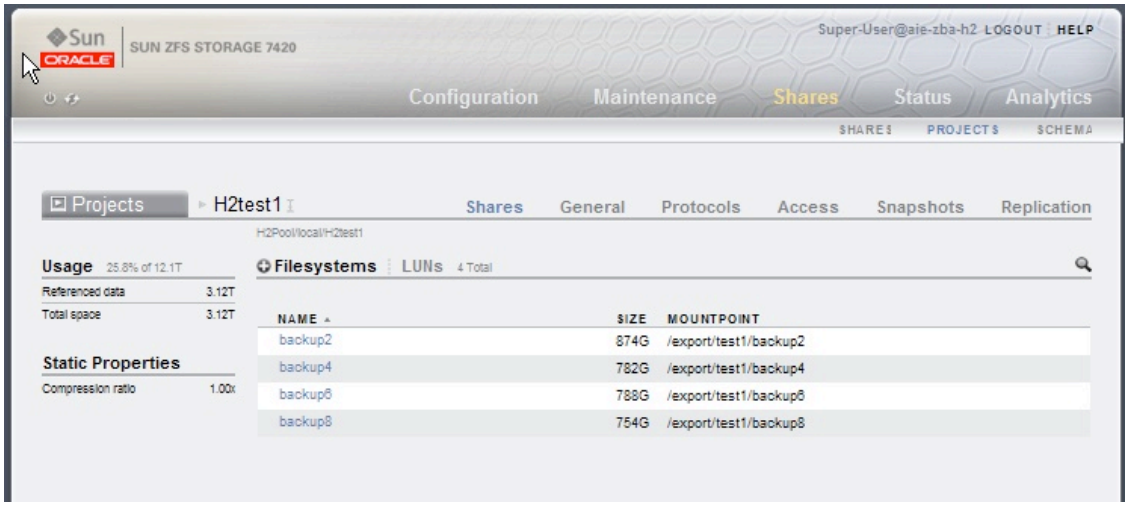

Figure 11. Filesystem listing for Head 2

Configuring DTrace Analytics in the Sun ZFS Backup Appliance

The Sun ZFS Backup Appliance includes a comprehensive performance analysis tool call DTrace Analytics. DTrace Analytics is a framework that monitors important subsystem performance accounting statistics. A subnet of the available accounting statistics should be monitored to provide comprehensive data on the effectiveness and performance of Oracle RMAN backup and restore workloads. The following analytics are configured on the Sun ZFS Backup Appliance (**Configuration -> Preferences -> Enable Advanced Analytics**):

- CPU: Percent utilization broken down by CPU mode
- Disk: Average number of I/O operations broken down by state of operation
- Disk: I/O bytes per second broken down by type of operation
- Disk: I/O operations per second broken down by latency
- Disk: Disks with utilization of at least 95 percent broken down by disk
- Network: Interface bytes per second broken down by direction
- Network: Interface bytes per second broken down by interface
- Protocol: NFSv3 operations per second broken down by size
- Protocol: NFSv3 operations per second broken down by type of operation
- Protocol: NFSv3 operations per second of type read broken down by latency
- Protocol: NFSv3 operations per second of type write broken down by latency
- Protocol: NFSv3 operations per second of type read broken down by size
- Protocol: NFSv3 operations per second of type write broken down by size

Implementing these accounting statistics helps end users gain a quantitative understanding of the instantaneous and historical resource consumption used in Quality of Service (QoS) Management for their specific implementation.

Configuring the Client NFS Mount

When configuring the Sun ZFS Backup Appliance, any server that accesses the appliance, including Oracle SPARC SuperCluster nodes, is considered a client. Configuring the client NFS mount includes creating the target directory structure for access to the Sun ZFS Backup Appliance as well as the specific NFS mount options necessary for optimal system performance. Mount options for Oracle Solaris clients are:

```
rw,bg,hard,nointr,rsize=1048576,wsize=1048576,proto=tcp,vers=3,forcedirectio
```

The mount points of the directories created on the Sun ZFS Backup Appliance should be created on each of the SPARC SuperCluster's Oracle Database LDOMs and added to their `/etc/inet/hosts` table.

Tuning the Oracle Solaris 11 Network and Kernel

The following entries should be added to the `/etc/system` file of each of the SPARC SuperCluster's Oracle Database LDOMs:

```
set rpcmod:clnt_max_conns = 8
set nfs:nfs3_bsize = 131072
```

Additionally, the following commands will need to be run on each of the SPARC SuperCluster's Oracle Database LDOMs every time the LDOMs are rebooted:

```
/usr/sbin/ndd -set /dev/tcp tcp_max_buf 2097152
/usr/sbin/ndd -set /dev/tcp tcp_xmit_hiwat 1048576
/usr/sbin/ndd -set /dev/tcp tcp_recv_hiwat 1048576
```

Additional tuning may be necessary to achieve optimal performance. Refer to the SPARC SuperCluster Tunables Document 1474401.1 on My Oracle Support (MOS) at https://support.oracle.com for the latest information.

Starting with the January 2013 Quarterly Full Stack Download Patch (QFSDP) Update, the SPARC SuperCluster release added an "ssctuner" tool which automatically sets tunables. Refer to the latest SPARC SuperCluster release notes, also available on MOS, for additional information.

Configuring Oracle Direct NFS (dNFS)

On each of the SPARC SuperCluster's Oracle Database LDOMs, configure dNFS as follows:

1. Shut down the running instance of the Oracle Database software.
2. Change directory to `$ORACLE_HOME/rdbms/lib`.
3. Enable dNFS:

   ```
   make -f $ORACLE_HOME/rdbms/lib/ins_rdbms.mk dnfs_on
   ```
4. Update the `oranfstab` (located in `$ORACLE_HOME/dbs`) with the server, path, and export names specific to the configuration.
   a. The `server` parameter refers to the local name of the Sun ZFS Backup Appliance head on the InfiniBand network.
   b. The `path` parameters should reflect the addresses for that head specified in the IPMP groups.
   c. The `export` parameters should reflect the mount points similar to the entries created in the `/etc/vfstab`.

The entries should look similar to these:

```
server: aie-zba-h1-stor
path: 192.168.30.100
path: 192.168.30.102
path: 192.168.30.104
path: 192.168.30.106
export: /export/test1/backup1 mount: /zba/test1/backup1
export: /export/test1/backup3 mount: /zba/test1/backup3
export: /export/test1/backup5 mount: /zba/test1/backup5
export: /export/test1/backup7 mount: /zba/test1/backup7
server: aie-zba-h2-stor
path: 192.168.30.101
path: 192.168.30.103
path: 192.168.30.105
path: 192.168.30.107
export: /export/test1/backup2 mount: /zba/test1/backup2
export: /export/test1/backup4 mount: /zba/test1/backup4
export: /export/test1/backup6 mount: /zba/test1/backup6
```

```
export: /export/test1/backup8 mount: /zba/test1/backup8
```

5. Restart the Oracle Database software instance.

Tuning the Oracle Database Instance for Oracle RMAN

Optimizing high-bandwidth backup and restore operations using Oracle RMAN and the Sun ZFS Storage Appliance requires adjusting the instance parameters that control I/O buffering. For information about how to tune these parameters, see Article ID 1072545.1: *RMAN Performance Tuning Using Buffer Memory Parameters* at http://support.oracle.com.

For Oracle Database 11*g*R2 on SPARC SuperCluster, tuning the following four parameters should be considered:

- `_backup_disk_bufcnt` – Number of buffers used to process backup sets
- `_backup_disk_bufsz` – Size of the buffers used to process backup sets
- `_backup_file_bufcnt` – Number of buffers used to process image copies
- `_backup_file_bufsz` – Size of the buffers used to process image copies

For backup and restore operations on backup sets and image copies, set the number of buffers to 64 and the buffer size to 1MB:

```
SQL> alter system set "_backup_disk_bufcnt"=64;
SQL> alter system set "_backup_file_bufcnt"=64;
SQL> alter system set "_backup_disk_bufsz"=1048576;
SQL> alter system set "_backup_file_bufsz"=1048576;
```

These commands may be configured persistently by adding them to the SPFILE, or they may be set dynamically in the Oracle RMAN run block used to execute the backup or restore operations.

The following code fragments show how to dynamically tune the buffer sizes and counts for backup and restore operations.

- Backup set backup:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64';
        sql 'alter system set "_backup_disk_bufsz"=1048576';
        allocate channel...
        ...
        backup as backupset database;
}
```

- Backup set restore:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64';
        sql 'alter system set "_backup_disk_bufsz"=1048576';
        allocate channel...
        ...
        restore database;
}
```

- Image copy backup:

```
run
```

```
{
        sql 'alter system set "_backup_file_bufcnt"=64';
        sql 'alter system set "_backup_file_bufsz"=1048576';
        allocate channel...
        ...
        backup as copy database;
}
```

- Image copy restore:

```
run
{
        sql 'alter system set "_backup_file_bufcnt"=64';
        sql 'alter system set "_backup_file_bufsz"=1048576';
        allocate channel...
        ...
        restore database;
}
```

Performing an incrementally applied backup requires reading an incremental backup set and writing to an image copy. To tune buffers for incrementally applied backups, run the following:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64';
        sql 'alter system set "_backup_disk_bufsz"=1048576';
        sql 'alter system set "_backup_file_bufcnt"=64';
        sql 'alter system set "_backup_file_bufsz"=1048576';
        allocate channel...
        ...
        recover copy of database;
}
```

Configuring Oracle RMAN

Configuring Oracle RMAN channel and parallelism includes specifying the filesystem targets for the Oracle RMAN backup channels and the total number of channels used for backup and restore operations. Performance benefits can be realized by configuring 16 or 32 Oracle RMAN channels such that they are evenly distributed over the Oracle Database instances and nodes in the RAC cluster and evenly distributed over the shares exported from the Sun ZFS Backup Appliance.

The following code fragments show sample Oracle RMAN run blocks for performing backup and restore operations for backup sets and image copies as well as applying incremental merges to image copies. The sample code is based on the following database configuration:

- Database name: `mydb`
- SYSDBA login: `sys/welcome`
- Scan address: `myssc-scan`
- Service names for the backup: `mydb_bkup[1-2]`

The Oracle RMAN run blocks for backup and restore using backup sets and image copies are shown in the examples in the following sections. In these examples, the mount points for the eight shares are accessed as `/zba/mydb/backup1` through `/zba/mydb/backup8`. Sixteen channels are configured, two per filesystem.

Backup set level 0 backup:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
        allocate channel ch01 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup1/%U';
        allocate channel ch02 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup2/%U';
        allocate channel ch03 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup3/%U';
        allocate channel ch04 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup4/%U';
        allocate channel ch05 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup5/%U';
        allocate channel ch06 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup6/%U';
        allocate channel ch07 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup7/%U';
        allocate channel ch08 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup8/%U';
        allocate channel ch09 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup1/%U';
        allocate channel ch10 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup2/%U';
        allocate channel ch11 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup3/%U';
        allocate channel ch12 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup4/%U';
        allocate channel ch13 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup5/%U';
        allocate channel ch14 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup6/%U';
        allocate channel ch15 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup7/%U';
        allocate channel ch16 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup8/%U';
        configure snapshot controlfile name to '/zba/mydb/backup1/snapcf_dbname.f';
        backup as backupset incremental level 0 section size 32g database
        tag 'FULLBACKUPSET_L0' plus archivelog tag 'FULLBACKUPSET_L0';
}
```

Backup set level 1 backup:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
        allocate channel ch01 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup1/%U';
        allocate channel ch02 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup2/%U';
        allocate channel ch03 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup3/%U';
        allocate channel ch04 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup4/%U';
        allocate channel ch05 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup5/%U';
        allocate channel ch06 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup6/%U';
        allocate channel ch07 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1' format '/zba/mydb/backup7/%U';
        allocate channel ch08 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2' format '/zba/mydb/backup8/%U';
        allocate channel ch09 device type disk connect 'sys/welcome@myssc-
```

```
        scan/dbname_bkup1’ format ’/zba/mydb/backup1/%U’;
        allocate channel ch10 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup2/%U’;
        allocate channel ch11 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup3/%U’;
        allocate channel ch12 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup4/%U’;
        allocate channel ch13 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup5/%U’;
        allocate channel ch14 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup6/%U’;
        allocate channel ch15 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup7/%U’;
        allocate channel ch16 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup8/%U’;
        configure snapshot controlfile name to
        ’/zba/mydb/backup1/snapcf_dbname.f’;
        backup as backupset incremental level 1 database tag
        ’FULLBACKUPSET_L1’ plus archivelog tag ’FULLBACKUPSET_L1’;
}
```

Image copy backup:

```
run
{
        sql ’alter system set "_backup_file_bufcnt"=64 scope=memory’;
        sql ’alter system set "_backup_file_bufsz"=1048576 scope=memory’;
        allocate channel ch01 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup1/%U’;
        allocate channel ch02 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup2/%U’;
        allocate channel ch03 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup3/%U’;
        allocate channel ch04 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup4/%U’;
        allocate channel ch05 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup5/%U’;
        allocate channel ch06 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup6/%U’;
        allocate channel ch07 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup7/%U’;
        allocate channel ch08 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup8/%U’;
        allocate channel ch09 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup1/%U’;
        allocate channel ch10 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup2/%U’;
        allocate channel ch11 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup3/%U’;
        allocate channel ch12 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup4/%U’;
        allocate channel ch13 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup5/%U’;
        allocate channel ch14 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup6/%U’;
        allocate channel ch15 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup1’ format ’/zba/mydb/backup7/%U’;
        allocate channel ch16 device type disk connect ’sys/welcome@myssc-
        scan/dbname_bkup2’ format ’/zba/mydb/backup8/%U’;
        configure snapshot controlfile name to
        ’/zba/mydb/backup1/snapcf_dbname.f’;
        backup incremental level 1 for recover of copy with tag ’IMAGECOPY’
        database;
}
```

Incremental merge to image copy:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
        sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';
        allocate channel ch01 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch02 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch03 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch04 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch05 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch06 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch07 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch08 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch09 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch10 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch11 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch12 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch13 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch14 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        allocate channel ch15 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup1';
        allocate channel ch16 device type disk connect 'sys/welcome@ad01-
        scan/dbname_bkup2';
        configure snapshot controlfile name to
        '/zfssa/dbname/backup1/snapcf_dbname.f';
        recover copy of database with tag 'IMAGECOPY';
}
```

Restore validate:

```
run
{
        sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
        sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
        sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';
        allocate channel ch01 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch02 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch03 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch04 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch05 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch06 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch07 device type disk connect 'sys/welcome@myssc-
```

```
        scan/dbname_bkup1';
        allocate channel ch08 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch09 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch10 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch11 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch12 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch13 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch14 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        allocate channel ch15 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup1';
        allocate channel ch16 device type disk connect 'sys/welcome@myssc-
        scan/dbname_bkup2';
        configure snapshot controlfile name to
        '/zba/mydb/backup1/snapcf_dbname.f';
        restore validate database;
}
```

Conclusion

The Sun ZFS Backup Appliance is a high-performance backup and recovery system that is optimized for Oracle's engineered systems. In the case of the Oracle SPARC SuperCluster T4-4, it easily integrates into the InfiniBand environment and enables high-speed backup and restore for Oracle databases capitalizing on the dNFS features.

References

- Sun ZFS Storage Appliance Administration Guide
  http://download.oracle.com/docs/cd/E22471_01/index.html
- Sun Datacenter InfiniBand Switch 36 Firmware Version 2.0
  http://docs.oracle.com/cd/E26698_01/index.html
- My Oracle Support (MOS) for latest support information, including updates, patches, troubleshooting notes
  https://support.oracle.com

Configuring a Sun ZFS Backup Appliance with Oracle SPARC SuperCluster
May 2013, Version 1.0
Author: Eric Polednik

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com

Oracle is committed to developing practices and products that help protect the environment

Hardware and Software, Engineered to Work Together