



An Oracle White Paper
April 2010

How to Manage Systems with Oracle[®] Solaris ZFS in Oracle Solaris Containers

Introduction	1
Oracle Solaris ZFS and Containers: Overview	2
Oracle Solaris ZFS and Containers: An Example	3
Creating a Zpool	4
Creating a Zone	4
Allocating an Oracle Solaris ZFS File System to a Zone.....	6
Creating New File Systems	7
Applying Quotas to the File Systems.....	8
Changing the Mountpoint of a File System.....	9
Setting the Compression Property.....	9
Taking a Snapshot.....	10
Conclusion	11
For More Information	11

Introduction

This guide is intended to show a new user the capabilities of the Oracle[®] Solaris ZFS File System when coupled with Oracle Solaris Containers. It describes the assignment of an Oracle Solaris ZFS file system to a zone, and some of the administrative tasks possible in such a configuration. Assigning a file system to a zone gives a zone administrator full delegated control of the Oracle Solaris ZFS file system, enabling them to take snapshots, create sub file systems and perform other tasks without requiring the global zone administrator's intervention.

After reading this guide, the user will have an understanding of how to delegate a Oracle Solaris ZFS file system to a container and to manage it from within the container.

Oracle Solaris ZFS and Containers: Overview

Oracle Solaris ZFS is a file system that is provided in the Oracle Solaris 10 operating system. It provides very high levels of data integrity and performance, and improves the ease of file system management by eliminating the need for a volume manager. Also, because it is 128-bit based, it opens the door to virtually unlimited data scalability.

Traditional file systems have a model of a single device or volume for each file system. This makes administration complex and potentially prone to administrative errors that could result in the destruction of data or lack of access to the data. Oracle Solaris ZFS uses a pooled storage model in which storage capacity is held within an expandable pool called a zpool. Multiple file systems can exist within a zpool and they can share the capacity and I/O resources of the pool. This new architectural approach, the storage version of virtual memory, permits easier administration of file systems and enables performance breakthroughs.

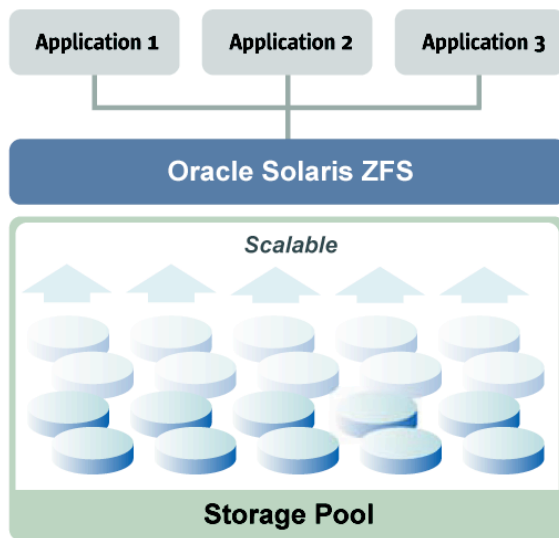


Figure 1. Oracle Solaris ZFS automatically grows shared storage

Oracle Solaris Containers is Oracle's operating system virtualization technique. It consists of several technologies that work together to foster improved resource management and isolate the environment from the underlying OS. With Oracle Solaris Containers, you can determine how to allocate resources among applications and services, and ensure they do not interfere with one another.

Oracle Solaris ZFS and Oracle Solaris Containers are tightly integrated to allow the zone administrator delegated rights to control the Oracle Solaris ZFS file systems. By using Oracle Solaris ZFS and Oracle Solaris Containers in combination, you can assign a portion of the storage pool to a container. This means that the administrator for that container has the authority to manage that portion of the storage

pool, but it is isolated from the rest of the system (Figure 2). This lets you benefit from both the data integrity of Oracle Solaris ZFS and the high availability and resource control of containers.

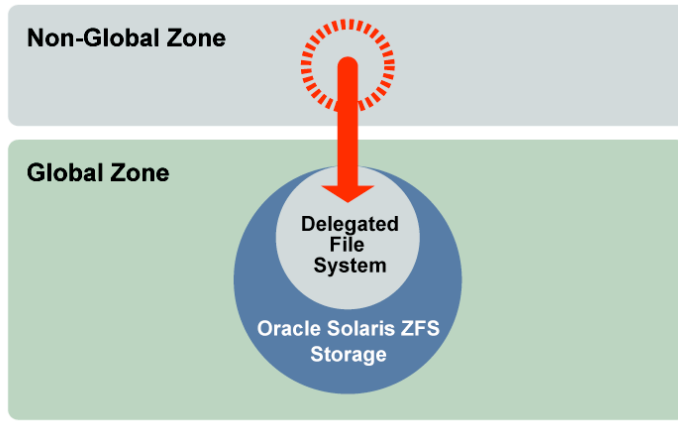


Figure 2. Containers let you isolate Oracle Solaris ZFS file systems

Note that while Oracle Solaris Containers technology has several aspects, the example in this guide will focus solely on the zones aspect.

Oracle Solaris ZFS and Containers: An Example

This section is a step-by-step guide that shows how to perform certain Oracle Solaris ZFS file system tasks inside of Oracle Solaris Containers; for example, taking snapshots and managing data compression. It does this by going through the following steps.

- Creating a zpool
- Creating a zone
- Allocating a Oracle Solaris ZFS file system to a zone
- Creating new file systems
- Applying quotas to the file systems
- Changing the mount point of a file system
- Setting the compression property
- Taking a snapshot

Each of these steps is described in detail below.

Creating a Zpool

Oracle Solaris ZFS uses device names or partition names when dealing with pools and devices. For a device, this will be something like `c1t0d0` (for a SCSI device) or `c1d1` (for an IDE device). For a partition, it will be something like `c1t0d0s0` (for a SCSI device) or `c1d1s0` (for an IDE device). This example creates a pool that is mirrored using two disks.

1. To create a zpool in the global zone, use the `zpool create` command. Typically, you use two devices to provide redundancy.

```
Global# zpool create mypool mirror c2t5d0 c2t6d0
```

Note that the `zpool create` command may fail if the devices are in use or contain some types of existing data (e.g. UFS file system). If they are in use, you will need to unmount them or otherwise stop using them. If they contain existing data you can use the `-f` (force) flag to override the safety check, but be sure that you are not destroying any data you want to retain.

2. Examine the pool properties using the `zpool list` command. This shows you that there is one zpool, named `mypool`, with a capacity of 199 GB.

```
Global# zpool list
```

NAME	SIZE	USED	AVAIL	CAP	HEALTH	ALTROOT
mypool	199G	164K	199G	0%	ONLINE	—

Creating a Zone

To show Oracle Solaris ZFS working in an environment that is isolated from other applications on the system, you need to create a zone. To create a zone:

1. Create a directory where the zone file system will reside, using the `mkdir` command. Be sure to choose a location where the file system has at least 80 MB of available space.

```
Global# mkdir /zones
```

Note that in this example, for the sake of brevity, the root file system of the zone is a UFS file system.

2. Configure the zone (myzone), using the zonecfg command, and specify the location of the zone's files (/zones/myzone). Use the following series of commands.

```
Global# zonecfg -z myzone
myzone: No such zone configured
    Use 'create' to begin configuring a new zone
zonecfg:myzone> create
zonecfg:myzone> set zonepath=/zones/myzone
zonecfg:myzone> verify
zonecfg:myzone> commit
zonecfg:myzone> exit
```

Again, for the purpose of streamlining, this example uses a very minimal zone. For more details on creating zones see the Oracle Solaris Containers How To Do Stuff guide at: https://www.sun.com/offers/details/solaris_containers.xml

3. Install the zone by using the zoneadm command.

```
Global# zoneadm -z myzone install
Preparing to install zone <myzone>
    [Output from zoneadm, this may take a few mins]
```

4. Boot the zone to complete the installation, using the zoneadm command.

```
Global# zoneadm -z myzone boot
```

5. Use the zlogin command to connect to the zone console.

```
Global# zlogin -C myzone
[Connected to zone 'myzone' console]
    [Initial zone boot output, service descriptions are loaded etc.]
```

It may take a short while for the first boot to configure everything, load all the service descriptors, and so on. You will need to answer the system configuration details. Some suggestions are:

```
Terminal=(12)X Terminal Emulator (xterms)
Not Networked
No Kerberos
Name service = None
Time Zone = your-time-zone
root passwd = (Your choice - remember it though!)
```

The zone will reboot after you have provided the configuration information.

6. Before you can proceed to the next stage, the configured zone needs to be shutdown (configuration changes are only applied when the zone boots).

```
Global# zlogin myzone init 5
```

Allocating an Oracle Solaris ZFS File System to a Zone

Now that you have a zpool (mypool) and a zone (myzone) you are ready to allocate a Oracle Solaris ZFS file system to the zone.

1. To create a Oracle Solaris ZFS file system, use the Oracle Solaris ZFS create command.

```
Global# Oracle Solaris ZFS create mypool/myzonefs
```

2. To apply a quota to the file system, use the Oracle Solaris ZFS set quota command.

```
Global# Oracle Solaris ZFS set quota=5G mypool/myzonefs
```

The file system and all of its child file systems can be no larger than the designated quota. Note that both these steps must be performed in the global zone. Also notice that creating the file system in Oracle Solaris ZFS is much simpler than with a traditional file system/volume manager combination.

To illustrate the isolation/security properties of containers with Oracle Solaris ZFS, this example now creates a Oracle Solaris ZFS file system that will remain outside the container. There is no need to apply a quota to this outside file system.

3. To create this other file system, again use the Oracle Solaris ZFS create command.

```
Global# Oracle Solaris ZFS create mypool/myfs
```

4. To show the properties of all the pool and the file systems, use the Oracle Solaris ZFS list command.

```
Global# Oracle Solaris ZFS list
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
mypool	396G	197G	99.5K	/mypool
mypool/myfs	98.5K	197G	98.5K	/mypool/myfs
mypool/myzonefs	98.5K	5G	98.5K	/mypool/myzonefs

To make the file system (myzonefs) available in the zone (myzone), the zone configuration needs to be updated.

- To update the zone configuration, use the `zonecfg` command.

```
Global# zonecfg -z myzone
zonecfg:myzone> add dataset
zonecfg:myzone:dataset> set name=mypool/myzonefs
zonecfg:myzone:dataset> end
zonecfg:myzone> commit
zonecfg:myzone> exit
```

The `mypool/myzonefs` file system is now added to the zone configuration. Note that these steps must be performed with the zone shut down, otherwise the zone configuration changes would not be visible until the next reboot. To check that the zone is shut down try logging into it using `zlogin myzone`. If the zone is shut down the login will fail; if the zone is running you will see a login prompt—login as root and shut the zone down with `init 5`. These steps are performed in the global zone.

- Now boot the zone.

```
Global# zoneadm -z myzone boot
```

- Log in to the zone. (Leave a few seconds for the zone to boot.)

```
Global# zlogin -C myzone
[Connected to zone 'myzone' pts/3]
[Usual Solaris login sequence]
```

- List the Oracle Solaris ZFS file systems in the zone

NAME	USED	AVAIL	REFER	MOUNTPOINT
<code>mypool</code>	0M	200B	—	<code>/mypool</code>
<code>mypool/myzonefs</code>	8K	5G	8K	<code>/mypool/myzonefs</code>

Creating New File Systems

Administering Oracle Solaris ZFS file systems from the non-global zone is done just like it is in the global zone, although you are limited to operating within the file system that is allocated to the zone (`mypool/myzonefs`). New Oracle Solaris ZFS file systems are always created as a child of this file system because this is the only Oracle Solaris ZFS file system the non-global zone can see. The administrator in the non-global zone can create these file systems. There is no need to involve administrator of the global zone, though the administrator could do so if it were necessary.

1. To create a new file system, use the Oracle Solaris ZFS create command.

```
Global# Oracle Solaris ZFS create mypool/myzonefs/tim
Global# Oracle Solaris ZFS list
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
mypool	504M	197G	99K	/mypool
mypool/myzonefs	197K	5.00G	98.5K	/mypool/myzonefs
mypool/myzonefs/tim	98.5K	5.00G	98.5K	/mypool/myzonefs/tim

The non-global zone administrator can create as many child file systems as s/he wants and each child file system can have its own file systems, and in that way form a hierarchy.

As a demonstration that the non-global zone administrator is limited to the assigned file systems, this example demonstrates trying to break security by creating a file system outside the container's "space".

2. Try to create another file system outside of mypool/myzonefs, using the Oracle Solaris ZFS create command. As you can see, Oracle Solaris ZFS and zones security denies permission for the non-global zone to access resources it has not been allocated and the operation fails.

```
MyZone# Oracle Solaris ZFS create mypool/myzonefs1
cannot create 'mypool/myzonefs1': permission denied
```

Applying Quotas to the File Systems

Typically, to prevent the user consuming all of the space, a non-global zone administrator will want to apply a quota to the new file system. Of course, the child's quota can't be more than 5 GB as that is the quota specified by the global zone administrator to all of the file systems below mypool/myzonefs.

To set a quota on our new file system, use the Oracle Solaris ZFS set quota command.

```
MyZone# Oracle Solaris ZFS set quota=1G mypool/myzonefs/tim
MyZone# Oracle Solaris ZFS list
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
mypool	508M	197G	99K	/mypool
mypool/myzonefs	198K	5.00G	99K	/mypool/myzonefs
mypool/myzonefs/tim	98.5K	1024M	98.5K	/mypool/myzonefs/tim

The administrator of the non-global zone has set the quota of the child file system to be 1G. They have full authority to do this because they are operating on their delegated resources and do not need to involve the global zone administrator.

The Oracle Solaris ZFS property inheritance mechanism applies across zone boundaries, so the non-global zone administrator can specify his/her own property values should s/he wish to do so. As with normal Oracle Solaris ZFS property inheritance, these override inherited values.

Changing the Mountpoint of a File System

Now that the file system is set up and has the correct quota assigned to it, it is ready for use. However, the place where the file system appears (the mountpoint) is partially dictated by what the global zone administrator initially chose as the pool name (in this example, mypool/myzonefs). But typically, a non-global zone administrator would want to change it.

To change the mountpoint, use the Oracle Solaris ZFS set mountpoint command.

```
MyZone# Oracle Solaris ZFS set mountpoint=/export/home/tim mypool/myzonefs/tim
MyZone# Oracle Solaris ZFS list
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
mypool	508M	197G	99K	/mypool
mypool/myzonefs	198K	5.00G	99K	/mypool/myzonefs
mypool/myzonefs/tim	98.5K	1024M	98.5K	/export/home/tim

Note that the mountpoint can be changed for any file system independently.

Setting the Compression Property

The next example demonstrates the compression property. If compression is enabled, Oracle Solaris ZFS will transparently compress all of the data before it is written to disk. The benefits of compression are both saved disk space and possible write speed improvements.

- To see what the current compression setting is, use the Oracle Solaris ZFS get command..

```
MyZone# Oracle Solaris ZFS get compression mypool mypool/myzonefs
mypool/myzonefs/tim
```

NAME	PROPERTY	VALUE	SOURCE
mypool	compression	off	default
mypool/myzonefs	compression	off	default
mypool/myzonefs/tim	compression	off	default

Be aware that the compression property on the pool is inherited by the file system and child file system. So if the nonglobal zone administrator sets the compression property for the delegated file system, it will set it for everything below, as well.

- To set the compression for the file system, use the Oracle Solaris ZFS set command.

```
MyZone# Oracle Solaris ZFS set compression=on mypool/myzonefs
```

- Examine the compression property again in the non-global zone. Note the compression property has been inherited by mypool/myzonefs/tim as with normal Oracle Solaris ZFS administration.

```
MyZone# Oracle Solaris ZFS get compression mypool mypool/myzonefs
mypool/myzonefs/tim
```

NAME	PROPERTY	VALUE	SOURCE
mypool	compression	off	default
mypool/myzonefs	compression	on	local
mypool/myzonefs/tim	compression	on	Inherited from mypool/myzonefs

Taking a Snapshot

One of the major advantages of Oracle Solaris ZFS is the ability to create an instant snapshot of any file system. By delegating a file system to a non-global zone this feature becomes available as an option for the non-global zone administrator.

To take a snapshot named “1st “ of the file system, use the Oracle Solaris ZFS snapshot command.

```
MyZone# Oracle Solaris ZFS snapshot mypool/myzonefs@1st
MyZone# Oracle Solaris ZFS list
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
mypool	512K	197G	99K	/mypool
mypool/myzonefs	198K	5.00G	99K	/mypool/myzonefs
mypool/myzonefs@1st	0K	—	99K	—
mypool/myzonefs/tim	98.5K	1024M	98.5K	/export/home/tim

As with Oracle Solaris ZFS file systems in the global zone, this snapshot is now accessible from the root of the file system in .Oracle Solaris ZFS/snapshot/1st.

Conclusion

Once a zone has been created and a Oracle Solaris ZFS file system has been allocated to it, the administrator for that (non-global) zone can create file systems, take snapshots, create clones, and perform all the other functions of an administrator—within that zone. Yet the global zone, and any other zones, are fully isolated from whatever happens in that zone.

The integration of Oracle Solaris Containers and Oracle Solaris ZFS is just another way that Oracle Solaris 10 is providing cost benefits to customers by allowing them to safely consolidate applications and more easily manage the data those applications use.

For More Information

While this white paper provides the rudimentary steps needed to get Oracle Solaris Container technology and the Oracle Solaris ZFS file system working in collaboration with each other, more sophisticated configurations are possible. For more information regarding Oracle Solaris Containers and Oracle Solaris ZFS, visit <http://www.oracle.com/us/products/servers-storage/solaris/index.html>.

TABLE 1. REFERENCES

MANUALS	
Oracle Solaris ZFS Administration Guide	http://docs.sun.com/app/docs/doc/819-5461
WEB-BASED TRAINING	
Oracle Solaris Containers Learning Center	sun.com/solaris/containers
How to Consolidate Servers and Applications using Oracle Solaris Containers	sun.com/solaris/howtoguides/containers
COMMUNITY RESOURCES	
OpenSolaris Oracle Solaris ZFS Community	opensolaris.org/os/community/Oracle Solaris ZFS
Blogs	http://blogs.sun.com/storage/en_US/entry/storage_blogs5



How to Manage Systems with Oracle Solaris
ZFS in Oracle Solaris Containers
April 2010
Author: Tim Graves, Joost Pronk
Contributing Authors: Paul Steeves

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

Copyright © 2010, Oracle and/or its affiliates. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. UNIX is a registered trademark licensed through X/Open Company, Ltd. 0310

SOFTWARE. HARDWARE. COMPLETE.