Waddles.org - Wad About Me?

Home Culemak Tags Categories About

Search

Q

Home Culemak Tags Categories About



Replicating ZFS Root Disks

2009-11-17 · 2768 WORDS · 13 MINUTE READ ■ OPEN SOURCE SOLARIS · ZFS

Suppose you have a Solaris system already set up with a ZFS root volume and you wish to clone it to another system by replicating the disks. This used to be easy with UFS volumes as you could simply use ufsdump piped to ufsrestore onto a new target disk, install boot blocks then move the new disk over to the target system and boot as normal. With ZFS there are a few extra hurdles incurred by the meta-data but it also saves us from problems like mixing up which disk was the primary mirror and which was the shadow copy. Since we can only replicate from a snapshot, we guarantee that all datasets in the pool contain a consistent set of data from a particular point in time.

Steps to clone a system with ZFS root volumes

- Source System
 - 1. Verify status before starting

- 2. Create recursive snapshot
- 3. Insert and label new disks
- 4. Create partition tables
- 5. Create new zpool
- 6. Replicate pool's dataset
- 7. Replicate remaining datasets
- 8. Set mountpoints
- 9. Set bootfs property
- 10. Export the new pool
- 11. Install boot blocks to both disks
- Target System
 - 1. Transfer disks and boot failsafe
 - 2. Import zpool as rpool
 - 3. Export the zpool
 - 4. Boot to single user mode
 - 5. Unconfigure the system
 - 6. Reconfigure the system
 - 7. Cleanup

Source System

Verify status before starting

```
# zpool status
  pool: rpool
  state: ONLINE
```

scrub: none requested

config:

I	NAME	STATE	READ	WRITE	CKSUM
	rpool	ONLINE	0	0	0
	mirror	ONLINE	0	0	0
	c0t0d0s0	ONLINE	0	0	0
	c0t1d0s0	ONLINE	0	0	0

errors: No known data errors

```
# zpool get bootfs rpool
NAME PROPERTY VALUE SOURCE
rpool bootfs rpool/ROOT/5.10 local
```

zfs list

NAME USED AVAIL REFER MOUNTPOINT

rpool	4.55G	62.4G	95K	/rpool
rpool/ROOT	1.54G	62.4G	18K	legacy
rpool/ROOT/5.10	1.54G	62.4G	1.12G	/
rpool/ROOT/5.10/var	434M	7.58G	434M	/var
rpool/dump	1.00G	62.4G	1.00G	-
rpool/home	31.5K	8.00G	31.5K	/home
rpool/srv	9.60M	62.4G	9.60M	/srv
rpool/swap	2G	64.4G	12.2M	-

Create recursive snapshot

```
# zfs snapshot -r rpool@20091116
```

```
# zfs list
NAME
                          USED AVAIL REFER MOUNTPOINT
rpool
                          4.56G 62.4G 95K /rpool
rpool@20091116
                                       95K -
                             0
                        1.54G 62.4G 18K legacy
rpool/ROOT
rpool/ROOT@20091116
                                        18K -
                         1.54G 62.4G 1.12G /
rpool/ROOT/5.10
rpool/ROOT/5.10@20091116
                          133K - 1.12G -
rpool/ROOT/5.10/var
                          434M 7.58G 434M /var
rpool/ROOT/5.10/var@20091116
                                 - 434M -
                             0
                          1.00G 62.4G 1.00G -
rpool/dump
rpool/dump@20091116
                                 - 1.00G -
rpool/home
                         31.5K 8.00G 31.5K /home
rpool/home@20091116
                                 - 31.5K -
rpool/srv
                        9.60M 62.4G 9.60M /srv
rpool/srv@20091116
                                 - 9.60M -
rpool/swap
                         2.01G 64.4G 12.2M -
rpool/swap@20091116
                                   - 12.2M -
```

Insert and label new disks

/pci@1c,600000/scsi@2/sd@3,0
Specify disk (enter its number): 2

selecting c0t2d0

[disk formatted]

Disk not labeled. Label it now? y

FORMAT MENU:

disk - select a disk

type - select (define) a disk type

partition - select (define) a partition table

3. c0t3d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>

current - describe the current disk
format - format and analyze the disk
repair - repair a defective sector
label - write label to the disk

analyze - surface analysis

defect - defect list management
backup - search for backup labels
verify - read and display labels

save - save new disk/partition definitionsinquiry - show vendor, product and revision

volname - set 8-character volume name

```
!<cmd> - execute <cmd>, then return
  quit
format> disk 3
selecting c0t3d0
[disk formatted]
Disk not labeled. Label it now? y
format> ^D
```

Create partition tables

```
# prtvtoc /dev/rdsk/c0t0d0s2 | fmthard -s - /dev/rdsk/c0t2d0s2
fmthard: New volume table of contents now in place.

# prtvtoc /dev/rdsk/c0t0d0s2 | fmthard -s - /dev/rdsk/c0t3d0s2
fmthard: New volume table of contents now in place.
```

Create new zpool

```
# zpool create newpool mirror c0t2d0s0 c0t3d0s0

# zpool status
  pool: newpool
  state: ONLINE
  scrub: none requested
config:
```

NAME	STATE	READ	WRITE	CKSUM
newpool	ONLINE	0	0	0
mirror	ONLINE	0	0	0
c0t2d0s0	ONLINE	0	0	0
c0t3d0s0	ONLINE	0	0	0

errors: No known data errors

pool: rpool
state: ONLINE

scrub: none requested

config:

NAME STATE READ WRITE CKSUM rpool ONLINE 0 0 0

mirror	ONLINE	0	0	0
c0t0d0s0	ONLINE	0	0	0
c0t1d0s0	ONLINE	0	0	0

errors: No known data errors

# zfs list				
NAME	USED	AVAIL	REFER	MOUNTPOINT
newpool	106K	66.9G	18K	/newpool
rpool	4.56G	62.4G	95K	/rpool
rpool@20091116	0	-	95K	-
rpool/ROOT	1.54G	62.4G	18K	legacy
rpool/ROOT@20091116	0	-	18K	-
rpool/ROOT/5.10	1.54G	62.4G	1.12G	/
rpool/ROOT/5.10@20091116	244K	-	1.12G	-
rpool/ROOT/5.10/var	434M	7.58G	434M	/var
rpool/ROOT/5.10/var@20091116	620K	-	434M	-
rpool/dump	1.00G	62.4G	1.00G	-
rpool/dump@20091116	0	-	1.00G	-
rpool/home	31.5K	8.00G	31.5K	/home
rpool/home@20091116	0	-	31.5K	-
rpool/srv	9.60M	62.4G	9.60M	/srv
rpool/srv@20091116	0	-	9.60M	-
rpool/swap	2.01G	64.4G	12.2M	-
rpool/swap@20091116	0	-	12.2M	-

Replicate pool's dataset

The base dataset contains the /boot/ directory which includes the boot_archive and kernel for booting.

This can also be done over the network using a transport such as ssh or rsh but you must have a target system already running.

rpool	4.56G	62.4G	95K	/rpool
rpool@20091116	0	-	95K	-
rpool/ROOT	1.54G	62.4G	18K	legacy
rpool/R00T@20091116	0	-	18K	-
rpool/ROOT/5.10	1.54G	62.4G	1.12G	/
rpool/ROOT/5.10@20091116	244K	-	1.12G	-
rpool/ROOT/5.10/var	434M	7.58G	434M	/var
rpool/ROOT/5.10/var@20091116	620K	-	434M	-
rpool/dump	1.00G	62.4G	1.00G	-
rpool/dump@20091116	0	-	1.00G	-
rpool/home	31.5K	8.00G	31.5K	/home
rpool/home@20091116	0	-	31.5K	-
rpool/srv	9.60M	62.4G	9.60M	/srv
rpool/srv@20091116	0	-	9.60M	-
rpool/swap	2.01G	64.4G	12.2M	-
rpool/swap@20091116	0	-	12.2M	-

Replicate remaining datasets

```
# for dataset in ROOT ROOT/5.10 ROOT/5.10/var dump home srv swap; do
> zfs send rpool/$dataset@20091116 | zfs receive newpool/$dataset
> done
# zfs list
NAME
                               USED AVAIL REFER MOUNTPOINT
                              2.14G 64.8G 95K /rpool
newpool
                                            95K -
newpool@20091116
                                  0
newpool/ROOT
                              1.12G 64.8G
                                           19K /rpool/ROOT
newpool/ROOT@20091116
                                16K
                                             18K -
newpool/ROOT/5.10
                              1.12G 64.8G 1.12G /rpool/ROOT/5.10
newpool/ROOT/5.10@20091116
                                  0
                                        - 1.12G -
newpool/ROOT/5.10/var
                               434M 64.4G
                                            434M /var
newpool/ROOT/5.10/var@20091116
                                  0
                                            434M -
newpool/dump
                              1.00G 64.8G 1.00G -
newpool/dump@20091116
                                  0
                                           1.00G -
newpool/home
                              31.5K 64.8G
                                           31.5K /rpool/home
newpool/home@20091116
                                           31.5K -
newpool/srv
                              9.60M 64.8G 9.60M /rpool/srv
newpool/srv@20091116
                                  0
                                        - 9.60M -
newpool/swap
                              12.2M 64.8G 12.2M -
newpool/swap@20091116
                                  0
                                        - 12.2M -
                                           97K /rpool
rpool
                              4.56G 62.4G
```

rpool@20091116	18K	-	95K	-
rpool/ROOT	1.54G	62.4G	18K	legacy
rpool/R00T@20091116	0	-	18K	-
rpool/ROOT/5.10	1.54G	62.4G	1.12G	/
rpool/ROOT/5.10@20091116	246K	-	1.12G	-
rpool/ROOT/5.10/var	434M	7.58G	434M	/var
rpool/ROOT/5.10/var@20091116	620K	-	434M	-
rpool/dump	1.00G	62.4G	1.00G	-
rpool/dump@20091116	0	-	1.00G	-
rpool/home	31.5K	8.00G	31.5K	/home
rpool/home@20091116	0	-	31.5K	-
rpool/srv	9.60M	62.4G	9.60M	/srv
rpool/srv@20091116	0	-	9.60M	-
rpool/swap	2.01G	64.4G	12.2M	-
rpool/swap@20091116	0	-	12.2M	-

Set mountpoints

```
# zfs set mountpoint=legacy newpool/ROOT
# zfs set mountpoint=/ newpool/ROOT/5.10
cannot mount '/': directory is not empty
property may be set but unable to remount filesystem
# zfs set mountpoint=/home newpool/home
cannot mount '/home': directory is not empty
property may be set but unable to remount filesystem
# zfs set mountpoint=/srv newpool/srv
cannot mount '/srv': directory is not empty
property may be set but unable to remount filesystem
# zfs list
NAME
                              USED AVAIL REFER MOUNTPOINT
newpool
                              2.14G 64.8G 95K /rpool
                                 0 -
                                           95K -
newpool@20091116
                              1.12G 64.8G 19K legacy
newpool/ROOT
newpool/ROOT@20091116
                                16K -
                                             18K -
                              1.12G 64.8G 1.12G /
newpool/ROOT/5.10
newpool/ROOT/5.10@20091116
                                 0
                                       - 1.12G -
newpool/ROOT/5.10/var
                               434M 64.4G 434M /var
newpool/ROOT/5.10/var@20091116
                                0
                                    - 434M -
```

newpool/dump	1.00G	64.8G	1.00G	-
newpool/dump@20091116	0	-	1.00G	-
newpool/home	31.5K	64.8G	31.5K	/home
newpool/home@20091116	0	-	31.5K	-
newpool/srv	9.60M	64.8G	9.60M	/srv
newpool/srv@20091116	0	-	9.60M	-
newpool/swap	12.2M	64.8G	12.2M	-
newpool/swap@20091116	0	-	12.2M	-
rpool	4.56G	62.4G	96K	/rpool
rpool@20091116	20K	-	95K	-
rpool/ROOT	1.54G	62.4G	18K	legacy
rpool/R00T@20091116	0	-	18K	-
rpool/ROOT/5.10	1.54G	62.4G	1.12G	/
rpool/ROOT/5.10@20091116	246K	-	1.12G	-
rpool/ROOT/5.10/var	434M	7.58G	434M	/var
rpool/ROOT/5.10/var@20091116	659K	-	434M	-
rpool/dump	1.00G	62.4G	1.00G	-
rpool/dump@20091116	0	-	1.00G	-
rpool/home	31.5K	8.00G	31.5K	/home
rpool/home@20091116	0	-	31.5K	-
rpool/srv	9.60M	62.4G	9.60M	/srv
rpool/srv@20091116	0	-	9.60M	-
rpool/swap	2.01G	64.4G	12.2M	-
rpool/swap@20091116	0	-	12.2M	-

Set bootfs property

Failure to complete this step will produce a 'no pool_props' error message when the target system boots.

```
# zpool get bootfs rpool
NAME PROPERTY VALUE SOURCE
rpool bootfs rpool/ROOT/5.10 local

# zpool get bootfs newpool
NAME PROPERTY VALUE SOURCE
newpool bootfs - default

# zpool set bootfs=newpool/ROOT/5.10 newpool
NAME PROPERTY VALUE SOURCE
```

newpool bootfs newpool/ROOT/5.10 local

Export the new pool

```
# zpool export newpool
# zpool status
 pool: rpool
state: ONLINE
scrub: none requested
config:
                 STATE READ WRITE CKSUM
      NAME
      rpool
                ONLINE
                           0
                                 0
        mirror ONLINE
                           0
                                0
                                     0
         c0t0d0s0 ONLINE
                           0
                                0
                                     0
```

c0t1d0s0 ONLINE 0 0

errors: No known data errors

Install boot blocks to both disks

0



On x86 hardware we use grub as a boot loader so the command is different:

```
# installgrub /boot/grub/stage1 /boot/grub/stage2 /dev/rdsk/c0t2d0s0
stage1 written to partition 0 sector 0 (abs 4096)
stage2 written to partition 0, 272 sectors starting at 50 (abs 4146)
```

Target System

Transfer disks and boot failsafe

Make sure you are on the system console (ALOM) of the target system.

```
SC Alert: DISK @ HDD0 has been inserted.
SC Alert: DISK @ HDD1 has been inserted.
{0} ok boot -F failsafe -Z newpool/ROOT/5.10
SC Alert: Host System has Reset
Probing system devices
Probing memory
Probing I/O buses
rsc not found.
rsc not found.
Probing system devices
Probing memory
Probing I/O buses
Rebooting with command: boot -F failsafe -Z newpool/ROOT/5.10
Boot device: /pci@1c,600000/scsi@2/disk@0,0:a File and args: -F failsafe -Z
krtld: Ignoring invalid kernel option -Z.
krtld: Unused kernel arguments: `newpool/ROOT/5.10'.
SunOS Release 5.10 Version Generic_139555-08 64-bit
Copyright 1983-2009 Sun Microsystems, Inc. All rights reserved.
Use is subject to license terms.
Hardware watchdog enabled
Configuring devices.
Searching for installed OS instances...
ROOT/5.10 was found on newpool.
Do you wish to have it mounted read-write on /a? [y,n,?] n
Starting shell.
#
<
                                                                          >
```

On x86 hardware, you should be presented with the grub splash screen where you can select to boot in failsafe mode.

```
GNU GRUB version 0.97 (638K lower / 550848K upper memory)

Solaris 10 10/09 s10x_u8wos_08a X86
Solaris failsafe
```



Grub menu

Import zpool as rpool

This step is necessary otherwise the system will complain that the zpool was last mounted on a different system and refuse to import it. Renaming is not strictly necessary but keeps consistency among various systems.

```
# zpool import -f newpool rpool
cannot mount '/': directory is not empty
cannot mount '/home': failed to create mountpoint
cannot mount '/rpool': failed to create mountpoint
cannot mount '/srv': failed to create mountpoint
cannot mount '/var': directory is not empty
# zfs list
NAME
                            USED AVAIL REFER MOUNTPOINT
                            2.57G 64.4G 95K /rpool
rpool
                                          95K -
rpool@20091116
                               0
rpool/ROOT
                            1.54G 64.4G
                                         19K legacy
rpool/ROOT@20091116
                              16K
                                           18K -
rpool/ROOT/5.10
                            1.54G 64.4G 1.12G /
rpool/ROOT/5.10@20091116
                            3.98M
                                      - 1.12G -
                             434M 64.4G 434M /var
rpool/ROOT/5.10/var
rpool/ROOT/5.10/var@20091116
                                          434M -
                               0
rpool/dump
                            1.00G 64.4G 1.00G -
                                     - 1.00G -
rpool/dump@20091116
                              16K
rpool/home
                            31.5K 64.4G 31.5K /home
```

Export the zpool

Now that we have updated the zpool's meta-data we can export it again and boot it without the system complaining.

```
# zpool export rpool

# init 0
syncing file systems... done
Program terminated
{0} ok
```

Boot to single user mode

If you see messages like this when booting, you have probably failed to replicate the /var dataset. Return the disks to the source system, import the zpool and try zfs send rpool/ROOT/5.10/var again.

```
ERROR: svc:/system/filesystem/minimal:default failed to mount /var/run (see Nov 16 17:52:47 svc.startd[7]: svc:/system/filesystem/minimal:default: Methoc Nov 16 17:52:47 svc.startd[7]: system/filesystem/minimal:default failed fatal Requesting System Maintenance Mode (See /lib/svc/share/README for more information.)

Console login service(s) cannot run

Root password for system maintenance (control-d to bypass):
```

This is what you should see:

```
{0} ok boot -s
SC Alert: Host System has Reset
Probing system devices
Probing memory
Probing I/O buses
rsc not found.
rsc not found.
Probing system devices
Probing memory
Probing I/O buses
Rebooting with command: boot -s
Boot device: /pci@1c,600000/scsi@2/disk@0,0:a File and args: -s
SunOS Release 5.10 Version Generic_141414-01 64-bit
Copyright 1983-2009 Sun Microsystems, Inc. All rights reserved.
Use is subject to license terms.
Hardware watchdog enabled
Booting to milestone " milestone/single-user:default".
Hostname: host01
WARNING: bge0 has duplicate address 192.168.105.104 (in use by 00:03:ba:bf:72
Requesting System Maintenance Mode
SINGLE USER MODE
Root password for system maintenance (control-d to bypass):
single-user privilege assigned to /dev/console.
Entering System Maintenance Mode
Nov 17 12:03:38 su: 'su root' succeeded for root on /dev/console
Sun Microsystems Inc. SunOS 5.10 Generic January 2005
host01:/#
<
```

On x86 hardware, you may need to edit the kernel line in grub's menu to boot to single user mode.

1. Press 'e' to edit the menu item

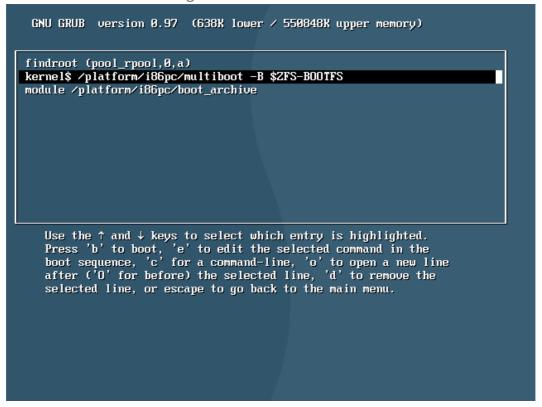
```
GNU GRUB version 0.97 (638K lower / 550848K upper memory)

Solaris 10 10/09 s10x_u8wos_08a X86
Solaris failsafe
```



Edit grub menu item

2. Press down arrow then 'e' again to edit the kernel line.



Edit kernel line

3. Append '-s' to the end of the line

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. 1
grub edit> kernel$ /platform/i86pc/multiboot -B $ZFS-BOOTFS -s
```



Append -s

4. Press Enter to accept the line, then 'b' to boot it.

Unconfigure the system

```
WARNING

This program will unconfigure your system. It will cause it to revert to a " blank" system - it will not have a name or know about other systems or networks.

This program will also halt the system.

Do you want to continue (y/n) ? y stopping NetWorker daemons: nsr_shutdown -q svc.startd: The system is coming down. Please wait. svc.startd: 22 system services are now being stopped. svc.startd: The system is down. syncing file systems... done Program terminated {0} ok
```

Reconfigure the system

```
{0} ok boot

SC Alert: Host System has Reset
Probing system devices
Probing memory
Probing I/O buses
```

```
rsc not found.
rsc not found.
Probing system devices
Probing memory
Probing I/O buses
Rebooting with command: boot
Boot device: bootdisk File and args:
SunOS Release 5.10 Version Generic 141414-01 64-bit
Copyright 1983-2009 Sun Microsystems, Inc. All rights reserved.
Use is subject to license terms.
Hardware watchdog enabled
Hostname: unknown
Configuring devices.
Reading ZFS config: done.
Mounting ZFS filesystems: (6/6)
Select a Language
 0. English
 1. es
 2. fr
Please make a choice (0 - 2), or press h or ? for help: 0
```

Note: If you connected to the ALOM remotely, you should choose **vt100** when prompted for your terminal type.

At the end of the reconfiguration, the system will reboot automatically. It may be necessary to add host entries back into /etc/hosts after the system has booted.

Cleanup

Once you are satisfied the system is running properly, destroy the source system's snapshot which is now on the target system. This can also be done on the source system when you are finished replicating disks.

```
# zpool status
pool: rpool
state: ONLINE
```

scrub: none requested

config:

NAME	STATE	READ	WRITE	CKSUM
rpool	ONLINE	0	0	0
mirror	ONLINE	0	0	0
c0t0d0s0	ONLINE	0	0	0
c0t1d0s0	ONLINE	0	0	0

errors: No known data errors

zfs list

NAME	USED	AVAIL	REFER	MOUNTPOINT
rpool	2.67G	64.3G	95K	/rpool
rpool@20091116	0	-	95K	-
rpool/ROOT	1.65G	64.3G	19K	legacy
rpool/R00T@20091116	16K	-	18K	-
rpool/ROOT/5.10	1.65G	64.3G	1.12G	/
rpool/ROOT/5.10@20091116	93.3M	-	1.12G	-
rpool/ROOT/5.10/var	446M	64.3G	445M	/var
rpool/ROOT/5.10/var@20091116	1.32M	-	434M	-
rpool/dump	1.00G	64.3G	1.00G	-
rpool/dump@20091116	16K	-	1.00G	-
rpool/home	52K	64.3G	31.5K	/home
rpool/home@20091116	20.5K	-	31.5K	-
rpool/srv	9.60M	64.3G	9.60M	/srv
rpool/srv@20091116	0	-	9.60M	-
rpool/swap	12.2M	64.3G	16K	-
rpool/swap@20091116	12.2M	-	12.2M	-

zfs destroy -r rpool@20091116

zfs list

NAME	USED	AVAIL	REFER	MOUNTPOINT
rpool	2.57G	64.4G	95K	/rpool
rpool/ROOT	1.56G	64.4G	19K	legacy
rpool/ROOT/5.10	1.56G	64.4G	1.12G	/
rpool/ROOT/5.10/var	445M	64.4G	445M	/var
rpool/dump	1.00G	64.4G	1.00G	-
rpool/home	31.5K	64.4G	31.5K	/home
rpool/srv	9.60M	64.4G	9.60M	/srv
rpool/swap	16K	64.4G	16K	-

Comments Share

OLDER

Drupal, CCK and Views in a Nutshell

comments powered by Disqus

RECENTS

OPEN SOURCE

MAAS WITH NGINX PROXY PASS 2016-02-23

OPEN SOURCE

ISSUES CONVERTING FROM DRUPAL TO HUGO 2015-12-19

OPEN SOURCE

REPLICATING ZFS ROOT DISKS 2009-11-17

OPEN SOURCE

DRUPAL, CCK AND VIEWS IN A NUTSHELL 2009-10-17

OPEN SOURCE

REBUILDING SANE (CANON PIXMA MP730) FOR UBUNTU 2009-08-24

CATEGORIES

▶ blog-dev (2)

- ▶ open-source (8)
- ▶ software-programming (8)

TAGS

- ▶ blog (2)
- ▶ colemak (5)
- ▶ culemak (2)
- ▶ drupal (2)
- ▶ hugo (1)
- ▶ maas (1)
- ▶ nginx (1)
- ▶ sane (2)
- ▶ solaris (1)
- ▶ ubuntu (3)
- ▶ zfs (1)

TAG CLOUD

blog colemak culemak drupal hugo maas nginx sane solaris ubuntu zfs



© 2016

Powered by Hugo. Theme by PPOffice.