[启用freebsd8.1的ZFS文件系统](http://blog.chinaunix.net/uid-9881170-id-133498.html)

我利用zfs目的是方便将来扩展空间方便，过去，文件系统局限于单个设备，因而其大小以设备的大小为限。由于ZFS 文件系统不局限于特定设备，因此可以轻松、快捷地创建，其创建方法与目录的创建方法相似。通过追加别的空间，ZFS 文件系统会在分配给其所驻留的存储池的磁盘空间中自动增大，可以轻松超过2TB。ZFS 建立在池存储概念的基础上。与典型文件系统映射到物理存储器不同，使用传统文件系统时，每次添加新文件系统都必须编辑/etc/vfstab 文件。ZFS 可根据数据集的属性自动挂载和取消挂载文件系统，从而消除了上述要求，无需管理/etc/vfstab 文件中的ZFS 项。

一，启用，freebsd8.1系统已经集成zfs文件系统功能，无线重新编译内核，直接启动即可:

在/etc/rc.conf中启用ZFS：

echo 'zfs\_enable="YES"' >> /etc/rc.conf

然后可以启动ZFS服务：/etc/rc.d/zfs start

为zfs优化下系统，修改/boot/loader.conf：

vm.kmem\_size="1024M"

vm.kmem\_size\_max="1024M"

vfs.zfs.arc\_max="100M"

vfs.zfs.vdev.cache.size="5M"

创建raidz的ZFS文件系统：

zpool create –f –R /home mailzfs /dev/da1s1

*mailzfs为创建的这个zpool的名字同时也在系统根目录下生成了”/home”文件夹，不需要修改/etc/fstab就可以直接使用，重启后不会消失。这个方式和很好地回避了存储掉线（往往是HBA卡驱动加载的问题，或者是链接延时）导致启动不了系统。*

zfs create mailzfs/data

zfs set mountpoint='/home' mailzfs/data #这个必须有，否则重启后默认挂到“/”上，系统就乱套了。

然后可以看到已经创建的分区：

# zpool list

NAME SIZE USED AVAIL CAP HEALTH ALTROOT

mailzfs 1.94T 72K 1.94T 0% ONLINE -

# df -h

Filesystem Size Used Avail Capacity Mounted on

/dev/da0s1a 9.5G 513M 8.2G 6% /

devfs 1.0K 1.0K 0B 100% /dev

/dev/da0s1g 11G 4.0K 9.9G 0% /home

/dev/da0s1d 47G 12K 44G 0% /tmp

/dev/da0s1e 12G 2.3G 9.0G 20% /usr

/dev/da0s1f 47G 946K 44G 0% /var

mailzfs 1.9T 18K 1.9T 0% /mailzfs

在需要的时候，我还有一块硬盘可以将空间加入，很简单：

zpool add –f mailzfs /dev/da0s1g

但是，有没有人可以告诉我怎么可以把它拿下来，我能加上但拿不下来，除非我destroy。

定期更新下文件系统状态吧：

echo 'daily\_status\_zfs\_enable="YES"' >> /etc/periodic.conf

**二、一些有用的zfs、zpool命令（man zpool）：**

一些有用的zpool命令：

1)       zpool history，看看你都用zpool命令干了些啥。

2)       zpool create zpool\_name raidz c0t0d0 c0t1d0 c0t2d0 c0t3d0 c0t4d0 c0t5d0，创建raidz的zfs文件系统，就是坏掉1个设备也不影响数据，也可以是raidz2,raidz3，坏掉2个、3个盘都没问题了。

3)       zpool create zpool\_name mirror c0t0d0 c0t1d0 mirror c0t2d0 c0t3d0，创建带镜像的zfs文件系统。

4)       zpool create tank /path/to/file/a /path/to/file/b，创建基于文件的zfs文件系统（没试过，有点晕），人家man page也说是实验性的。

5)       zpool destroy -f zpool\_name，小心了，这么一下，就全费费了！

6)       zpool export zpool\_name，感觉就象umount一样，df、mount都看不到了，就是离线了。

7)       zpool status，看看zfs文件系统的状态，如果export了，那就不在线了。

8)       zpool import mailzfs，搞回来了，如果zpool import则可以看到可以搞啥！

9)       zpool create tank mirror c0t0d0 c0t1d0 spare c0t2d0，带热备的，spare，c0t2d0为热备。

10)   zpool remove tank c0t2d0，删掉热备盘，好像删别的不行呀。

11)   zpool upgrade –a，查看本地所有zfs文件系统的pools是不是一样的版本，我运行后：This system is currently running ZFS pool version 14. All pools are formatted using this version 2.

12)   zpool iostat -v pool\_name 5，每5秒统计下IO情况。

13)   zfs list，看zfs自家都有啥。

14)   zfs get mountpoint,compression，查看挂载点和压缩状况

15)   zpool scrub pool\_name，好像是磁盘扫描，能修复一下错误吧，fsck对zfs不好用了。

16)   zfs snapshot pool\_name@20101103，创建一个snapshot，下次恢复到这个状态，但是，哥哥，存在哪里了？快照是文件系统或卷的只读副本，快照不会占用池中的任何附加磁盘空间，快照开销很小。

17)   zfs rollback pool\_name@20101103，恢复到上次20101103的snapshot。

18)   zfs list -t snapshot，查看有哪些snapshot。

19)   zfs list -o space，可以看到详细的空间使用情况。

20)   zfs destroy pool\_name@20101103，删除那个快照。

21)   zfs create pool\_name/home，在pool\_name池里创建了home,这个目录用rm –r删不了的，

22)   zfs set compression=gzip pool\_name/home，启用压缩，节省空间呗。

23)   zfs set compression=off pool\_name/home，取消压缩

24)   zfs set copies=2 pool\_name/home，双备份，但不知道有什么用，不会用，等待进一步学习。

25)   zfs destroy pool\_name/home，删除这个home。

26)   zpool clear pool\_name，重置故障日志设备的错误状态。

27)   zpool get all pool\_name，获得pool\_name的所有属性值。

28)   zfs clone 克隆，一个经典例子，但没有实践过：

在以下示例中，将创建一个名为tank/home/ahrens/bug123 的新克隆，其初始内容与快照tank/ws/gate@yesterday 的内容相同：

# zfs snapshot tank/ws/gate@yesterday

# zfs clone tank/ws/gate@yesterday tank/home/ahrens/bug123

在以下示例中，将从projects/newproject@today 快照为临时用户创建克隆工作区

projects/teamA/tempuser。然后，在克隆工作区上设置属性。

# zfs snapshot projects/newproject@today

# zfs clone projects/newproject@today projects/teamA/tempuser

# zfs set sharenfs=on projects/teamA/tempuser

# zfs set quota=5G projects/teamA/tempuser