Subject: vzfs support
Posted by Eric on Wed, 14 Sep 2005 10:13:51 GMT
View Forum Message <> Reply to Message

Are you going to include something like vzfs?

Subject: Re: vzfs support
Posted by dim on Wed, 14 Sep 2005 11:22:28 GMT
View Forum Message <> Reply to Message

No, it is Virtuozzo feature.

Subject: Re: vzfs support
Posted by TheWiseOne on Wed, 14 Sep 2005 14:14:58 GMT
View Forum Message <> Reply to Message

You're better off without it anyway.

vzfs is an overlay filesystem. Now, unless the vzfs developers have done some real magic, overlay filesystems are always prone to double caching.

As an example in Virtuozzo the data under /vz/private/<veid>/root and then the overlayed vzfs at /vz/root/<veid> is cached twice. I'd like to see this argued, this is somehow possible if the inode betwen the backing and the overlay is shared, but unlikely. To the kernel it does it's smart cache/buffer/swapping based on the files inode.

Reasons like this is why I see my server with 8GB of RAM using 5GB between the buffer/cache.... man would I love to see 2-3GB of that used for application data.

Subject: Re: vzfs support
Posted by dim on Wed, 14 Sep 2005 15:13:36 GMT
View Forum Message <> Reply to Message

What do you mean under 'double caching'?
Which kernel do you use on your server with 8GB RAM?

kernel finds balance between cache buffers/swap itself, but you can try to change it through sysctls. These sysctls are depends on kernel version. For 2.4 kernels it is vm.bdflush, for 2.6 - vm.swappiness, vm.dirty_ratio, vm_dirty_background_ratio.

If you need to avoid much memory used for system buffers - try to tune these parameters.

Subject: Re: vzfs support
Posted by TheWiseOne on Wed, 14 Sep 2005 15:14:12 GMT
View Forum Message <> Reply to Message

I suppose I eat my words on this, it does seem to share the same inode.

Then I can only help that vm.swappiness can help my systems.

```
# ls -ali /vz/private/1/root/etc/inittab /vz/root/1/etc/inittab
451387465 -rw-r--r--    1 root     root         1456 Sep  2 16:31 /vz/private/1/root/etc/inittab
451387465 -rw-r--r--    1 root     root         1456 Sep  2 16:31 /vz/root/1/etc/inittab
```

Subject: Re: vzfs support
Posted by TheWiseOne on Wed, 14 Sep 2005 15:16:52 GMT
View Forum Message <> Reply to Message

I know of the sysctls, under 2.4 it does not work well at all. I just showed it does indeed use the same inode. Traditionally with an overlay fs the inodes are different between the source and mount and since the buffercache identifies by inode each file is cached twice. Whoever wrote vzfs worked around this (it is something able to be done, just not very "mainstream").