

---

Subject: "strong" Disk I/O prioritisation

Posted by [gnuterred](#) on Wed, 09 Apr 2008 11:45:17 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

I want to be able to give absolute I/O priority to one (or a few) VE, and have it so that all I/O for other VEs gets serviced only when idle.

I tried using the "set ... --ioprio" option, as per below, but found that in rough tests, it only slews things slightly in favour of my preferred VE.

Am I doing something wrong? Is this as good as it gets?

As a workaround, I tried using ionice within the VE. Even after adding the sys\_admin and sys\_nice capabilities, I still got

```
# ionice -c3 id
ioprio_set: Operation not permitted
```

I could perhaps live with running ionice within my low priority VEs if I could get it to work.

Detail...

Kernel: Linux sodium 2.6.18-028stab053.tl20080408 (my compile of OVZ's latest vanilla kernel + latest stable patch)

urand is a 512MB cat of /dev/urandom in the root directory of each of VEs 100 and 200 (different file)

By themselves, it takes about 8 seconds to cat the file:

```
tony@sodium:~$ sudo vzctl exec 100 time cat /urand \> /dev/null
```

```
real  0m7.754s
user  0m0.052s
sys   0m0.524s
tony@sodium:~$ sudo vzctl exec 200 time cat /urand \> /dev/null
```

```
real  0m7.803s
user  0m0.044s
sys   0m0.520s
```

I start running one in a constant loop:

```
tony@sodium:~$ while true; do sudo vzctl exec 100 time cat /urand \> /dev/null ; done
```

Then in another terminal I cat the file on the other VE. This shows that IO is roughly shared between VEs at the moment:

```
tony@sodium:~$ sudo vzctl exec 200 time cat /urand \> /dev/null
```

```
real  0m17.745s
```

```
user 0m0.068s
sys 0m0.504s
```

Now I set the priorities:

```
tony@sodium:~$ sudo vzctl set 100 --ioprio 0 --save
Saved parameters for VE 100
tony@sodium:~$ sudo vzctl set 200 --ioprio 7 --save
Saved parameters for VE 200
```

And restart the loop. In the other VE, I run once:

```
tony@sodium:~$ sudo vzctl exec 200 time cat /urand \> /dev/null
```

```
real 0m13.341s
user 0m0.080s
sys 0m0.488s
```

I was hoping this figure would be a lot closer to the original 8 seconds. Can I do better than this in brutally prioritising one VE over another?

My rationale for this is to protect my production webserver (two heavily used phpbb websites) from my other VEs. Currently (linux-vserver, Ubuntu 6.06), simply copying a file of moderate size (more than 500MB) brings the prod webserver to its knees for minutes, which makes me unpopular.