
Subject: Re: [RFC][PATCH 4/7] UBC: syscalls (user interface)

Posted by [dev](#) on Fri, 18 Aug 2006 11:03:31 GMT

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

Andrew Morton wrote:

> On Thu, 17 Aug 2006 16:13:30 +0400

> Kirill Korotaev <dev@sw.ru> wrote:

>

>

>>>I was more thinking about (for example) user land physical memory limit

>>>for that bean counter. If the limits are going down, then the system

>>>call should try to flush out page cache pages or swap out anonymous

>>>memory. But you are right that it won't be possible in all cases, like

>>>for in kernel memory limits.

>>

>>Such kind of memory management is less efficient than the one

>>making decisions based on global shortages and global LRU algorithm.

>

>

> I also was quite surprised that openvz appears to have no way of

> constraining a container's memory usage. "I want to run this bunch of

> processes in a 4.5GB container".

If you mean user memory, then it is possible to set

container limits to 4,5GB. This is what most people care about

and it is not a problem.

Or you mean that you are surprised there are lots of parameters

and there is no a single one limiting the _whole_ memory set of container

memory (sum of kernel memory, user space memory and other resources memory)?

>>The problem here is that doing swap out takes more expensive disk I/O

>>influencing other users.

>

>

> A well-set-up container would presumably be working against its own

> spindle(s). If the operator has gone to all the trouble of isolating a job

> from the system's other jobs, he'd be pretty dumb to go and let all the

> "isolated" jobs share a stinky-slow resource like a disk.

why do you assume that it is always an operator who controls the applications
inside the container?

users can run any application inside and it is systems job to

introduce resource isolation between users, not the operators full-time

job doing monitoring of users.

> But yes, swap is a problem. To do this properly we'd need a way of saying

> "this container here uses that swap device over there".

yep, this is possible with page bean counters as it tracks user pages.

more over, we have an intention of building a system with a single container

memory parameter, but we think this is more user interface question and still requires all the UBC resources accounting.

Thanks,
Kirill
