
Subject: Re: [RFD][PATCH] memcg: Move Usage at Task Move
Posted by [Paul Menage](#) on Wed, 11 Jun 2008 08:04:14 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wed, Jun 11, 2008 at 12:45 AM, KAMEZAWA Hiroyuki

<kamezawa.hiroyu@jp.fujitsu.com> wrote:

>> Is it really such a big deal if we don't transfer the page ownerships
>> to the new cgroup? As this thread has shown, it's a fairly painful
>> operation to support. It would be good to have some concrete examples
>> of cases where this is needed.

>>

> When we moves a process with XXXG bytes of memory, we need "move" obviously.

That's not a concrete example, it's an assertion :-)

>

> I think there is a case that system administrator decides to create `_new_`
> cgroup to isolate some swappy job for maintaining the system.
> (I never be able to say that never happens.)

OK, that seems like a reasonable case - i.e. when an existing cgroup
is deliberately split into two.

An alternative way to support that would be to do nothing at move
time, but provide a "pull_usage" control file that would slurp any
pages in any mm in the cgroup into the cgroup.

>> >

>> > One reason is that I think a typical usage of memory controller is
>> > `fork()->move->exec()`. (by `libcgroup` ?) and `exec()` will flush the all usage.

>>

>> Exactly - this is a good reason **not** to implement move - because then
>> you drag all the usage of the middleware daemon into the new cgroup.

>>

> Yes but this is one of the usage of cgroup. In general, system admin can
> use this for limiting memory on his own decision.

>

Sorry, your last sentence doesn't make sense to me in this context.

If the common mode for middleware starting a new cgroup is `fork()` /
`move` / `exec()` then after the `fork()`, the child will be sharing pages
with the main daemon process. So the move will pull all the daemon's
memory into the new cgroup

> yes. but, at first, I'll try no-rollback approach.

> And can I move memory resource controller's `subsys_id` to the last for now ?

>

That's probably fine for experimentation, but it wouldn't be something we'd want to commit to -mm or mainline.

Paul

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
