
Subject: Re: [-mm PATCH 5/10] Memory controller task migration (v7)
Posted by [Paul Menage](#) on Tue, 28 Aug 2007 20:04:34 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 8/28/07, YAMAMOTO Takashi <yamamoto@valinux.co.jp> wrote:

>
> although i have no good idea right now, something which allows
> to move a process with its thread group leader dead would be better.
>

One way I was thinking of approaching this problem was slightly different:

- every mm always has an "owning" task. Initially that will be the thread that creates the mm
- if the owning thread exits or execs and *isn't* the last user of the mm, then we may need to find a new owner for the mm:

1) My guess is that typically the thread that created the mm will also be the last user of the mm - if this is the case, then in the normal case we don't need to find a new owner.

2) If we do need a new owner, first look amongst the other threads in the process (cheap, should find another user of the mm quickly)

3) next look in the child and parent threads (more expensive, but rarer)

4) if necessary, scan the entire thread list (expensive, but should never be needed in general use)

The advantage of this is that we don't then need to have a memory container pointer in the mm - we can just use the memory container of the mm's owner.

With just a single container type needing to be tied to an mm, this isn't a huge advantage since we're just replacing one pointer (memory container) with another (owning task) and have similar levels of complexity for both. But if we have multiple container subsystems that need to be tied to a particular mm then they can both use the mm owner pointer.

E.g. I want to add a swap container subsystem that restricts which swap devices a group of processes can swap to, and how many pages they can put into swap. And I want to be able to run this independently of the in-memory page accounting subsystem. Having a task owner pointer in the mm allows these to be independent subsystems, and (I believe) isn't any more complex than the work involved to support moving an mm whose thread group leader has exited or execd.

Paul

Containers mailing list

Containers@lists.linux-foundation.org

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