Subject: Re: [PATCH v2 09/11] memcg: propagate kmem limiting information to children
Posted by Glauber Costa on Tue, 21 Aug 2012 09:22:09 GMT
View Forum Message <> Reply to Message

On 08/21/2012 11:54 AM, Michal Hocko wrote:
> On Fri 17-08-12 14:36:00, Glauber Costa wrote:
>> On 08/17/2012 02:35 PM, Michal Hocko wrote:
>>>>> But I never said that can't happen. I said (ok, I meant) the static
>>>>> branches can't be disabled.
>>> Ok, then I misunderstood that because the comment was there even before
>>> static branches were introduced and it made sense to me. This is
>>> inconsistent with what we do for user accounting because even if we set
>>> limit to unlimitted we still account. Why should we differ here?
>>
>> Well, we account even without a limit for user accounting. This is a
>> fundamental difference, no ?
>
> Yes, user memory accounting is either on or off all the time (switchable
> at boot time).
> My understanding of kmem is that the feature is off by default because
> it brings an overhead that is worth only special use cases. And that
> sounds good to me. I do not see a good reason to have runtime switch
> off. It makes the code more complicated for no good reason. E.g. how do
> you handle charges you left behind? Say you charged some pages for
> stack?
>
Answered in your other e-mail. About the code complication, yes, it does
make the code more complicated. See below.

> But maybe you have a good use case for that?
>
Honestly, I don't. For my particular use case, this would be always on,
and end of story. I was operating under the belief that being able to
say "Oh, I regret", and then turning it off would be beneficial, even at
the expense of the - self contained - complication.

For the general sanity of the interface, it is also a bit simpler to say
"if kmem is unlimited, x happens", which is a verifiable statement, than
to have a statement that is dependent on past history. But all of those
need of course, as you pointed out, to be traded off by the code complexity.

I am fine with either, I just need a clear sign from you guys so I don't
keep deimplementing and reimplementing this forever.