
Subject: Re: [PATCH 04/10] memcg: Introduce __GFP_NOACCOUNT.
Posted by [Glauber Costa](#) on Tue, 06 Mar 2012 18:31:55 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 03/06/2012 08:13 PM, Suleiman Souhlal wrote:
> On Tue, Mar 6, 2012 at 2:36 AM, Glauber Costa<glommer@parallels.com> wrote:
>> On 03/04/2012 04:10 AM, Suleiman Souhlal wrote:
>>>
>>> Just a few lines below:
>>>
>>> data = kmalloc_node_track_caller(size, gfp_mask, node);
>>>
>>> -- Suleiman
>>
>> Can't we just make sure those come from the root cgroup's slabs?
>> Then we need no flag.
>
> Do you mean make it so that all kmallocs come from the root cgroup's slabs?
No.

> We would really like to account kmallocs in general (and all the other
> slab types) to the right cgroup...
For the record, even in the scenario where we mark the slabs we want to track, I'd still like to account and limit kmallocs. They are a big source of memory usage.

What I mean, is that if we can specify which cgroup certain kmallocs should be billed against, we could use that to get rid of this GFP flag. And if we really want another non-kmalloc slab to escape accounting, then we can use a SLAB flag instead of a GFP flag.

>
> That said, I'm probably going to concentrate on accounting specially
> marked caches only, for now, since there seems to be a strong
> opposition on accounting everything, even though I don't understand
> this point of view.

Give me a couple of days to sketch something on the lines of what I described above.

I think that would be a good compromise for us here. If it is also okay for Kame, we might have good chances of reaching consensus with that.
